

Hierarchy and Information

Daniel Rogger and Ravi Somani*

September 2019

Abstract

The information public officials use to make decisions determines how public resources are distributed and how effective public policy is. This paper develops a measurement framework for assessing the accuracy of a set of fundamental bureaucratic beliefs and provides experimental evidence on the possibility of ‘evidence briefings’ improving that accuracy. The errors of public officials are large, with over 50 percent of officials making errors that are at least 50 percent of objective benchmark data. The provision of briefings reduces these errors by a quarter of a standard deviation, but in line with theoretical predictions, organizational incentives mediate their effectiveness.

Keywords: information, bureaucrats, public administration, decentralization

JEL Classification: D73, D82, H00, H11, H70, H83, O20

*Rogger: World Bank Research Department, 1818 H St NW, DC 20433, USA, and Institute for Fiscal Studies (IFS); drogger@worldbank.org. Somani: Department of Economics, University College London; ravi.somani.09@ucl.ac.uk. We thank the Ministry of Public Service and Human Resource Development of the Federal Government of Ethiopia for their leadership and support throughout design, implementation and data collection. We are very grateful to Elsa Araya, Verena Fritz and Kerenssa Kay of the World Bank’s Governance Global Practice for collaboration on implementation of the survey that underlines this work. We benefited from discussions with Pierre Bachas, Guadalupe Bedoya, Antonio Cabrales, Aidan Coville, Stefan Dercon, Simeon Djankov, Florian Englmaier, Eliana La Ferrara, Arianna Legovini, David McKenzie, Marcos Vera Hernandez and Imran Rasul, as well as from comments by seminar participants at Bocconi, George Washington University, Paris School of Economics, Kent, Namur, Oxford, Toulouse and the UK’s Department for International Development. The study received ethics approval from Oxford University Blavatnik School of Government’s Departmental Research Ethics Committee (protocol SSD/CUREC1A/BSG C1A 16-010); the registration number at the American Economic Association’s Registry for Randomized Controlled Trials is AEARCTR-0003946. We gratefully acknowledge financial support from the Centre for the Study of African Economies [grant number CUR00090], the UK’s Department for International Development, the International Growth Centre [grant number 32407], and the World Bank’s Development Impact Evaluation i2i fund, Bureaucracy Lab and Governance Global Practice.

1 Introduction

The information public officials use to make decisions determines how public resources are distributed and how effective public policy is (Muralidharan et al., 2016; Dhaliwal and Hanna, 2017; Callen et al., 2018; Dal Bo et al., 2018; Hjort et al, 2019). Errors in that information can reduce social welfare as resources are misdirected and the marginal efficiency of public investments falls. Correspondingly, government circulars and policy-related research frequently attempt to improve the information set of public officials and thereby raise the efficiency of policy making.¹

To date, empirical evidence on the accuracy of information used in public sector decision-making is very limited. This paper contributes to filling that gap by providing descriptive and experimental evidence on the distribution and determinants of information held by public officials inside a large, hierarchical public sector. It develops a measurement framework for assessing the accuracy of a set of fundamental bureaucratic beliefs, provides descriptive evidence on the distribution on information across the hierarchy, and provides experimental evidence on the possibility of ‘evidence briefings’ improving that accuracy.

More precisely, we develop a survey of 1,831 public officials of the Government of Ethiopia that extracts their beliefs about the key variables on which their work depends, such as the size of the population they are serving or sector-specific variables such as the number of primary-age children who are in school. We compare the responses of the surveyed officials to objective benchmark data from administrative and survey data to calculate the scale of errors that these officials make. We thus provide among the first large-scale descriptive evidence on the nature and correlates of information in the public sector.

A minority of the public officials we study make relatively accurate claims about their constituents. Of officials’ assessments of the population they serve, 21% are within 20% of the census-defined population. However, a large proportion of bureaucrats make economically-meaningful mistakes about the basic conditions of local jurisdictions. 47% of officials claim that their district’s population is 50% bigger or smaller than it is. The scale of errors is similar across the

¹This is an explicitly stated aim of many important economics research centres such as: the US-focused Bipartisan Policy Center’s Evidence-Based Policymaking Initiative, which provides “advice and expertise on implementation options and strategies ... as Congress develops legislation and as the Executive Branch devises regulations, policies, and standards to improve the generation of high level, quality evidence.”; the Centre for Economic Policy Research (CEPR) that “initiates, funds and coordinates research activities and communicates the results quickly and effectively to decision makers around the world.”; the Abdul Latif Jameel Poverty Action Lab (J-PAL) which is “a global research center working to reduce poverty by ensuring that policy is informed by scientific evidence.”; and the World Bank, that “generates high-quality and operationally relevant data and research to transform development policy”.

distribution of district populations, implying that public officials are indeed adjusting their estimates towards the underlying state of nature, but doing so with a lot of noise.² Across the socio-economic variables we study, 56% of public officials make errors that are at least 50% of the underlying benchmark data. Restricting our analysis to that set of variables that are the primary responsibility of officials, such as those questions asked to education officials about education, 31% of officials make errors that are at least 50% of the underlying benchmark data.

We find that the errors we report predict resource distributions in budget data, career success and a proxy of bureaucratic productivity in appraisal data, all key elements of government productivity. Such large errors and corresponding consequences are consistent with studies of agents from other settings, such as farmers not optimizing important aspects of the technology they use (Hanna et al., 2014), Indian manufacturing firms having limited information on quality defects, machine downtime, and inventory (Bloom et al., 2012), and professional forecasters making large but predictable mistakes (Ehrbeck and Waldmann, 1996; Andrade and Le Bihan, 2013).

To our knowledge our measurement framework is the first direct measure of individual-level information in related literatures. Fisher et al. (2018) record ‘knowledge-seeking’ activities (proxied by downloads from and visits to a particular website). The finance literature measures the (‘soft’) information held by loan officers at lending institutions. The measures of soft information are either elicited from survey questions that ask directly the extent to which employees feel that they know borrower characteristics (Uchida et al, 2012; Hattori et al, 2015), the predictive power of internal risk ratings relative to ‘hard’ measurable characteristics of firms (Agarwal and Hauswal, 2009, 2010; Brown et al, 2012; Qian et al 2015; Agarwal and Ben-David, 2018, Dal Bo et al, 2018), or measures of the effort exerted by relationship managers to learn about their clients (Liberti, 2018).

Dominant theories of information in the public sector have focussed on informational asymmetries across levels of hierarchy. At the organizational level, models have assessed hierarchy as it pertains to tiers of government (Oates, 1999; Bardhan, 2002, 2016; Mookherjee, 2015). A key rationale for the decentralization of government services to lower tiers of government has been the assumed superior information of decentralized agencies due to their greater proximity to citizens. As Oates (1999) states, “individual local governments are presumably much closer to the people and geography of their respective jurisdictions; they possess knowledge of both local preferences

²We also find that officials’ claims over their own district’s populations are more accurate than their claims over other districts.

and cost conditions that a central agency is unlikely to have’’. Bardhan (2002) simply argues that the cost of access to information is lower relative to central level governments.³

We apply our measurement framework across a representative sample of public officials and provide comparisons of the accuracy of their claims across tiers of government. By limiting our comparisons to claims about the same indicator, at the same percentile of its distribution, for the same district, and controlling for individual factors such as education and experience, we are able to provide the most precise descriptive evidence to date on the relationship between delegation and information (Acemoglu et al., 2007; Aghion et al., 2017).⁴ We find that public officials at the decentralized tier of government do have more accurate information than their centralized colleagues. Decentralized officials make claims there are roughly a third of a standard deviation more accurate than centralized officials.

However, we find substantial variation in the quality of information within tiers of government that correlate strongly with the incentives officials are provided to acquire information. We combine our measures of the accuracy of bureaucratic claims with information about the management practices they operate under to show that our descriptive results closely mirror a series of core theoretical predictions regarding the distribution of information in the public sector. Unlike markets, bureaucracies do not have organic information creation mechanisms, such that individual officials must undertake costly actions to acquire and absorb information. As bureaucratic decision-making is based on team production and decision-making, officials face the threat of being over-ruled by their superiors and an incentive to free ride on the information acquisition of other bureaucrats (Argenziano et al., 2016; Di Pei, 2015; Stiglitz, 2017). Officials will only undertake costly action to acquire information when they work in an incentive environment that rewards that effort and in which the threat of being overruled is limited (Gant et al., 2002; Mookherjee, 2006).

Our empirical results closely match the predictions of this theoretical literature. That decentralized tiers have superior information accords with corresponding theoretical predictions (Mookher-

³It is not necessarily the case that the local government should have superior information. Under the revelation principle, it is feasible that the central government can perfectly extract the information of the local government (Myerson, 1981; Mookherjee, 2006). Or as Cremer et al. (1994) argues, in acquiring information “there is no reason why central government cannot use any of the techniques available to local government”.

⁴Using our measures, we can assess the importance on the precision of public officials’ estimates of basic personal characteristics, such as an official’s education or experience, local environmental features, such as the poverty rate, or by their proximity to the local environment we are questioning them on. Demographic and environmental variables and proxies of the remoteness of the officials from their constituents do not explain a substantial proportion of the variation in errors. Rather, the institutional environment in which officials work explains the dominant portion of the variation we document. For example, the size of the coefficient on our measure of decentralization in our core descriptive regression is 41 times larger in absolute terms than the coefficient on the local poverty rate.

jee, 2006; Acemoglu et al., 2007). Moreover, we find that management practices related to information acquisition substantially mediate the impact of decentralization on the accuracy of claims. In those organizations in which management practices are least likely to reward information acquisition, there is no difference between the accuracy of claims at different tiers of government. These results imply that unconditional decentralization, without effective corresponding management practices or more broadly incentives for information acquisition, does not improve the information environment.

Theories of information in the public sector have also focussed on hierarchy at the individual-level, between a manager and subordinate (Aghion and Tirole, 1997; Garicano, 2000; Dessein, 2002). We find no difference in the accuracy of information across managers and employees. This is consistent with theory, since in our setting control rights within the organization are vested in management, limiting the incentive for subordinates to acquire additional information beyond organization-wide incentives.⁵ Together our two main descriptive results provide support for theories in which superior information at decentralized levels, either across tiers of government or layers of management, is a result of greater control rights and complementary incentives rather than proximity.

Building on this descriptive investigation of the incentives that encourage information acquisition in the public sector, we undertake a complementary experimental investigation of a reduction in the costs of accessing information. Precisely, we assess whether the accuracy of bureaucratic beliefs can be improved by evidence briefings that are a frequent means by which government or academic research is communicated to public officials. Such briefings have long been used within public service but are increasingly a focus of research entities such as the Centre for Economic Policy Research, JPAL/IPA, and the World Bank Group. For example, the number of evidence briefings produced per year by the Campbell Collaboration, a focal coordinator in the “use of systematic reviews and other evidence synthesis for evidence-based policy”, has increased steadily over the past decade (Campbell Collaboration, 2019).

There is surprisingly little evidence on whether such briefings do in fact inform public officials and the conditions under which they are effective. Elliott and Popay (2000) present qualitative evidence from the UK that organizational factors appear to significantly mediate the impact of research in healthcare organizations. Callen et al (2018) document changes in frontline outcomes

⁵Correspondingly, the manager has no incentive to keep information from the implementing agent.

from the installation of a new information technology in Pakistan. This paper builds on an existing body of non-experimental work on evidence briefings (Weiss and Bucuvalas, 1980) as well as experimental research on the effects of evidence briefings using email-based experiments on a selected sample of individuals already registered to be interested in research evidence and policy briefs (Masset et al, 2013), with selected participants in research dissemination workshops (Vivalt and Coville, 2018; Nellis et al, 2019), with politicians in Brazil and in Pakistan (Hjort et al, 2019; Liaqat, 2019), and a select group of development policy professional (Banuri et al, 2017). The present study is the first to provide individual-level measures of bureaucratic beliefs on the current state of service delivery in their jurisdiction and test the direct effect of evidence briefings on such beliefs with a representative sample of bureaucrats. Combining our experiment with rich data on individual and organizational characteristics, we provide evidence on the conditions under which information acquisition by public officials occurs, providing a potential explanation for the conflicting results of the existing literature.

We provide an information package that exactly mimics circulars (internal communications) of the Ethiopian government to a random sub-set of the officials that we study. The information package contains all the information we test for in our survey. Lowering the marginal cost of information acquisition dramatically and symmetrically in this way should yield superior information overall (Aghion and Tirole, 1997).⁶

We find an average treatment effect equivalent in magnitude to the descriptive impact of de facto delegation. The errors made by treated officials are 0.28 of a standard deviation lower than control officials. Our results are consistent with evidence from a range of settings that reducing the cost of becoming informed increases information acquisition. Hastings and Weinstein (2008) find that providing a simple information packet to parents regarding local school test results increases the proportion of parents that factor this information into their school choice. Jessoe and Rapson (2014) find that electricity consumers become more responsive to price changes when presented with high-frequency information about their residential electricity usage through a simple in-home display. Banerjee et al. (2018) show that mailing Indonesian households information on their eligibility for rice subsidies resulted in a significant increase in the amount of subsidy that they

⁶This finding is echoed in Gabaix et al. (2006), where information acquisition is a decreasing function of both the cost of acquisition and its complexity, and Garicano (2000), in which cheap information is widely shared within an organization. Myatt and Wallace (2012) study how much costly attention players in a coordination game should pay to signals that vary in accuracy (how precisely it identifies the state) and clarity (how easy it is to understand), finding that the clarity of signals is of primary importance even if they have poor underlying accuracy.

received, implying that they absorbed and acted on the information.⁷ Our experiment extends these results to the public administration, in which free-riding concerns bring in to question whether even large decreases in the cost of acquiring information have impacts on its equilibrium level (Aghion and Tirole, 1997). We provide evidence that it does.

However, the treatment effect arises totally from organizations with below-average management practices, reinforcing the argument that organizational incentives are key mediating factors in how new information is received and absorbed. While external interventions can reduce the scale of errors made by public officials, suitable organizational incentives ensure self-acquisition of that information.

This result is of significance to researchers who aim to update the beliefs of public officials. Organizational incentives to acquire information determine the efficacy of interventions and technologies aimed at improving the information of public-sector agents. Our experimental findings highlight that when the public sector is designed appropriately, the rationale for compensating interventions - which are often ad hoc and focused on short term goals - is weakened.⁸ The paper thus relates to the nascent literature on how the design of bureaucratic organizations impacts on state capabilities (reviewed by Finan et al., 2017), and on the importance of these capabilities for development outcomes (Besley and Persson, 2011; Acemoglu and Robinson, 2012; Pepinsky et al., 2017).

The rest of this paper proceeds as follows. Section 2 details the approach we took to measuring information in the public sector. Section 3 presents detailed descriptives of the errors we observe officials making. Section 4 provides descriptive evidence on the determinants of information in hierarchy. Section 5 outlines our experimental design and presents results. Section 6 concludes.

⁷In the Banerjee et al study, the more information that was provided on the mailing, the larger the treatment effect, implying that it was the information, rather than simply the presence of a reminder that was driving the results.

⁸The results echo those of Olken (2007) in which top-down government monitoring reduces corruption substantially more than grass-roots monitoring, Muralidharan and Sundararaman (2011) where effectively designed incentive schemes improved student learning outcomes far more than unconditional grants, and Pomeranz (2015) in which pre-existing incentives for tax compliance substitute for government audits.

2 Measuring Information

2.1 Institutional Background

Ethiopia is Africa's second most populous country, home to 100 million people or 10% of the population of Sub-Saharan Africa. Its ethnically diverse population is dispersed across over a million square kilometers, such that population characteristics vary widely across space.

Like other developing countries, government expenditures represent a significant fraction of GDP (18%) and the public sector is a large employer of formal workers. Corruption is less prevalent than in most of Sub-Saharan Africa, but government effectiveness is seen as relatively weak overall. Never fully colonized, Ethiopia's government is organized as a developmental state, with a relatively vertical and rigid hierarchy.

The state is organized in a federal system, with three major tiers of government: federal, regional, and district (woreda). Each tier has its own staff of civil servants who must define and refine government policy within their tier, direct budgetary and other resources, and manage the implementation of public services. Though the specific mix and focus of these tasks differs across tiers, each tier undertakes these tasks to a degree, and must make choices informed by the information they have on citizen needs. For example, roughly half of all public resources are spent by the Federal Government, with the rest being directed by the regional governments or spent directly by district administrations (World Bank, 2016b). None of the officials we study implements public services themselves, at any of the tiers, but rather administers public policy and resources.

Across sectors, the structure of the vertical hierarchy is relatively constant. Ministry organizations focus more on agenda setting for the sector, and the development of governing policies, but also direct resources to service providers and thus focus on managing implementation. Regional organizations develop policies for their region in line with federal stipulations, and guide resources across districts. District governments focus on service delivery, while translating policies into guidelines for their district. Each tier therefore makes decisions on how to distribute resources depending on where they believe there is most significant need. This is true in both the development of policy, which dictates the areas of the sector that should receive resources, as well as in implementation, which assigns resources geographically. Federal and regional organizations co-ordinate closely in their development of policy and operational activities, and can be seen as a

coherent ‘centralized’ body that provides joint guidance to district governments.

In the analysis we present below, we conceptualize hierarchy along two key dimensions of the public service. The relationship between central government agencies and district governments is such that many basic decision rights lie with the district government. Day-to-day decisions made by members of a district organization cannot be overruled by the central government. The relationship between central and district government agencies therefore accords closely to the theoretical literature’s notion of de facto ‘agent authority’ (Aghion and Tirole, 1997). In contrast, within an organization a manager has the ability to overrule the agents she employs on business decisions. Thus, the hierarchical relationships within an organization accord more closely to the notion of de facto ‘principal authority’. By contrasting the differences in information between principals (central tiers of government or managers) and agents (district governments or employees) under these two regimes allows us to assess the relationship between the nature of authority and information acquisition.

We also use the fact that different organizations in the Ethiopian context are managed differently to investigate how environments that differ in terms of how they value the acquisition of information are associated with our observed measure of information acquisition. As the culture of an organization becomes more empirical, we can interpret this as an increase in the individual rewards for information use. We can thus assess the relationship between mediating incentives and information acquisition.

2.2 Survey of Public Officials

The implementation of any of the three core tasks of public administration outlined above requires information on the characteristics and thus needs of the citizens they serve. To elicit an understanding of the information public officials have about their environment and to understand the incentive environments in which they operate, between June and September 2016 we held face-to-face interviews with 1,831 public administrators in 382 organizations across the three tiers of Ethiopia’s government. Table A1 lists the organizations included in our study, and Figure A1 provides a map of the district governments we sampled. Our analysis focuses on the agriculture, education, health, revenue, and trade sectors only. For each of these five sectors, we sample the main ministry, all the corresponding regional offices, and a geographically representative sample of

corresponding district offices from across the country.⁹ Within each organization, we interviewed senior managers and a representative sample of their staff. We limit our scope to the professional grades of technical and administrative officers, excluding grades that cover cleaners, drivers, secretaries, etc.

Table 1 provides descriptives for the 382 organizations we study. All officials work within a relatively standard structure, with a manager overseeing levels of hierarchy below him within a clearly defined organizational structure. As is common, the more centralized an authority is, the larger it is in terms of both the number of managers and the number of employees. Federal organizations also have a higher number of employees per director than either regional or district government offices.

In terms of bureaucrat characteristics, around 20% are women (at managerial grades, 8% are women, at non-managerial grades 24% are women); 82% of Ethiopian bureaucrats have some form of university education, with 11% having a postgraduate degree (at managerial (non-managerial) grades, 17% (10%) have a post-graduate degree). As in other state organizations, bureaucrats enjoy stable employment once in service: the average bureaucrat has 13 years in service, with their average tenure in the current organization being 7 years.

Across tiers, bureaucrats are similar in that they are in their mid-thirties, have been in their current post for roughly similar amounts of time, and have all worked in roughly 3 service organizations. However, officials at regional organizations are more senior and more likely to be a manager. Officials at district organizations are less likely to have a university education, and only 1% of district officials have a master's degree. They are also more likely to be male, with 18% of district officials female compared with 23% of regional officials and 27% of federal employees.

To develop and enumerate a questionnaire that was relevant for the Ethiopian setting, we worked closely with the Ministry of Public Service and Human Resource Development and employed ex-civil servants within our enumeration teams to facilitate navigation of the public service. As will be described in more detail below, we used the survey to ask a range of questions about citizen characteristics and the public service rules to capture the knowledge of public officials with regards their working environment. Beyond our focus on collecting bureaucrat knowledge, we surveyed officials on their basic characteristics and the nature of management practices utilized at their organization. The implementation of the survey was successful across the organizations we

⁹Our survey, and therefore analysis, necessarily excluded the majority of the Somali region for security reasons.

visited, with 99.5% of public officials sampled agreeing to be interviewed and 98.2% of interviews being classified by the enumerator as having gone ‘somewhat well’ (26.4%) or ‘very well’ (71.7%).

In conjunction with the survey, we also collected data on the performance of public officials as assessed in their annual appraisal. Each year, public officials are evaluated by their direct manager on the tasks that they were expected to contribute to. For example, tasks might include ‘Monitor and provide support to the team preparing the budget’ and ‘Support the team to prepare soft and hard copy documents of the budget’. Managers evaluate the quality of contributions bureaucrats make to the tasks they were involved in and produce an overall ‘performance’ score. In addition to this performance-related score, public officials are evaluated on their ‘attitude’ to work, which intends to measure their office behavior and alignment to the organization (Abagisa, 2014; Tereda, 2014). For the year 2016, we collected the performance, attitude, and total scores (which are a weighted average of performance and attitude scores) for each official from a subset of the organizations we visited for which they were available.

2.3 Subjective Assessments of Public Officials

We used the public official’s survey to ask each official direct questions on the socio-economic conditions of the citizens they serve and on civil service rules that govern their workplace. Respondents had to answer based on their existing knowledge of their constituencies and workplace rather than by referring to external reference documents. Focussing on the ‘tacit’ knowledge of public officials is consistent with classic conceptions of public sector tasks whose formal measurement is infeasible or prohibitively costly, thus requiring public officials to use personal judgement for their implementation. Interpreting legislation into specific policies and implementing them will likely never be a fully data-driven task.

This approach is supported by existing data (World Bank, 2012; Masaki et al., 2017) and our own survey evidence as to what public official’s most important sources of information are. Our survey probed how officials gathered information to use in their work, and the three most frequently cited sources were ‘formal field visits’, ‘discussions with frontline colleagues’, and ‘informal interactions with colleagues’. Only 12.8% of officials state that they use management information systems (MIS) as their primary source of information. As Duflo (2017) argues, policymakers “tend to decide on [policy design details] based on hunches, without much regard for evidence.” The Data Appendix provides further details of our survey questions relating to the sources of information on

which officials base their decision-making, but the approach taken here is that the tacit knowledge and beliefs of public officials will always play a substantive role in a nation's governance.

The knowledge questions we asked are listed in Table A2 and relate to both district characteristics and civil service rules. Those relating to socio-economic characteristics of citizens correspond to the core variables of interest in the national 'Growth and Transformation Plan', and sector-specific plans, such as the 'Education Sector Development Program IV: 2010 - 2015'. Those relating to the public service correspond to the rules outlined in the 'Ethiopian Federal Civil Servants Proclamation No. 515/2007'. Thus, all the variables we asked about are fundamental quantities relating to the core work of the public officials we interviewed. For example, the core mission of education officers is to implement the education sector development programs, focusing on the topics it prioritizes.

There are three sets of questions. The first set asks basic demographic questions about districts such as population and the proportion of the population that lives in rural areas. These questions are asked to respondents across all sectors. The second part asks sector-specific questions enumerated to sector-officials only. For example, respondents working in the education sector were asked the following questions about the citizens in their constituency: "What do you think the primary (1-8) enrolment numbers are for [district name] according to official administrative data?"; "What do you think the primary (1-8) pupil-teacher ratio is for [district name] according to official administrative data?"; "What do you think the primary (1-8) pupil-class ratio is for [district name] according to official administrative data?"; and, "What do you think the primary (1-8) pupils-per-school ratio is for [district name] according to official administrative data?"¹⁰ The third set ask questions about the rules of employment of the civil service, such as the number of hours a public official is expected to work in a regular working week, the annual leave allowance in days in the first year of service, and so on. These are once again asked to all officials. All indicators asked about were continuous in nature.

Since the three tiers of government serve constituencies of differing sizes, we asked federal officials which region they knew best, and then asked them questions about the 25th, median and 75th percentile districts within that region by socio-economic indicator. Regional officials were asked about the corresponding 25th, median and 75th percentile districts within the region where they worked. District officials were asked about the local government in which they worked. Thus,

¹⁰District-level sector-specific information was not available for the Addis Ababa City Administration, Dire Dawa City Administration and Harar Regional Administration, hence these administrations are excluded from the analysis.

all officials made statements about specific districts in their constituency, and in particular for the federal officials, those they stated they knew best.¹¹

For each question, public officials were permitted to state, “Don’t Know”, though the enumerator did not suggest this as a response. In 23% of questions, the respondent stated that they simply did not know and could not provide a valid guess. This is in contrast to the fact that the quantities assessed are the core variables required to make policy, planning and distribution decisions in accordance with the national sector plans.

2.4 Objective Assessments of Citizen Characteristics

For each indicator, we have administrative or survey data that can benchmark the responses by officials. The specific sources for each variable are outlined in Table A2 and are the latest Census, the latest available service delivery data from the education and health management information systems, and the Woreda and City Benchmarking Survey (WCBS). Each data set is created by sourcing information directly from service providers and aggregating it at the center of government. The sources we use are generally regarded as the highest quality large-scale measurement undertaken in Ethiopia on the topics outlined above. For example, in education, the Education Management Information System is comprised of school-level data that are processed through each tier of government, officials of which undertake quality checks to ensure that the data are as accurate as possible. The WCBS data was collected in collaboration with the World Bank and the Ministry of Public Service and Human Resource Development as a monitoring tool over five separate waves of data collection. All of this data is available to the officials we study through the relevant directorates of their tier of government at the minimum, and in some cases online.¹²

We take the administrative and survey data as a proxy for the underlying state of the world. We use these variables to benchmark the claims of public officials. For each indicator, we subtract the corresponding value in the benchmark data from the relevant claims in our data set, such that a positive error is an overestimate.

¹¹The exact phrasing of questions depended on the tier at which the respondent worked. Within each indicator, federal/regional level respondents are asked about the value of the 25th percentile district in their (chosen) region, the 50th percentile district in their (chosen) region, and the 75th percentile district in their (chosen) region.

¹²One proxy of the quality of the benchmarking data is their stability over time. By assessing the extent to which indicators vary within districts over data collection rounds, we aim to understand whether the data reflect a noisy collection process. Table A2 presents summary statistics and tests for the presence of a random walk for each of the indicators used in the analysis where panel data are available, and the Appendix describes the results. We find evidence that the benchmarking data is consistent over time.

Our core analysis aims to bring together the full set of claims in a unified framework. However, individual errors are not comparable across indicators. Making an error of 0.5 in a district’s pupil-class ratio is large for that indicator, but tiny for the population indicator. To jointly assess the scale of errors across multiple indicators, we undertook a number of transformations of the raw error data. First, we took absolute values of the errors so that deviations of different signs are comparable.¹³ Second, we trimmed the data at the 95th percentile. In a very small number of cases, errors were so large that it seemed infeasible for even the most uninformed bureaucrat to have made such an error, and it looked likely to be a data entry problem. Either way, the interpretation of the remaining data is based on the errors of 95% of the claims of public officials in our study.

Our core transformation is to create z-scores of each of the variables by subtracting the mean and dividing by the standard deviation of all the errors related to a particular indicator. This scales the claims by an empirical measure of a ‘normalized’ measure of error for that indicator. The interpretation of the z-score is a unitless measure of how many standard deviations of the full distributions of error away from the benchmark data an official’s assessment is. Thus, it is a measure of whether an individual official made a ‘large’ error relative to the empirical distribution of errors we observe for that indicator. Comparisons across indicators therefore take into account the empirical distribution of errors for each indicator. They compare the scale of an official’s error relative to the errors made by all other officials in our data set, indicator by indicator.

Defining z-scores in this way assumes that the tails of the raw errors are roughly symmetric for each indicator, such that an error 10% below and above the truth are equally likely. However, the distribution of errors of a number of the indicators we study are right-skewed, such that we may be underestimating the severity of an error in the sense that few people would make it. To test the robustness of our core measure to this criticism, we define a ‘skewness-adjusted’ z-score, where we test for skewness of the distribution of a raw indicator and replace skewed indicators with their log. We then take the z-score of these transformed variables in a similar way to the above measure.

Finally, we move away from distributional concerns by taking the ordinal rank of absolute errors within an indicator and divide this by the maximum rank, thereby creating a ‘rank-proportion’ score. This can be interpreted as the percentile of error made by an official relative to other officials

¹³The probability of a negative (as opposed to a positive) error in the data we use for our core analysis is 0.404, such that the likelihood of over- and under-estimates is relatively equal. We have also transformed the errors to be percentages of the underlying benchmark data, and then created the various measures outlined below. Taking this approach does not qualitatively change our core results.

in our data set. The Data Appendix provides further details on how the error variables we use in the analysis were created.

2.5 Measuring the Incentive Environment

To understand the incentive environments in which public officials in the organizations we are studying make decisions on information absorption, we also collected data on management practices. Following recent efforts to collect data on the management practices of public administrations (Rasul and Rogger, 2018; Rasul et al., 2018) we used a public sector version of the World Management Survey (WMS; Bloom and Van Reenen, 2007; Bloom et al., 2012) to elicit measures of the management practices under which public officials operate.

The WMS evaluation tool elicits management practices through a semi-structured interview covering 7 topics: flexibility, incentives, monitoring, roles, staff involvement, staffing, and targeting. Table A3 details each of the 19 management-related questions, by topic, as well as the 1-5 scoring grid used by our enumerators for each question. These questions provide a holistic assessment of practices related to the topic.¹⁴

We take the ‘Management Practice: Monitoring’ index as a proxy for the organization’s value of the information acquired by officials and the ‘Management Practice: Other’ index, which aggregates the non-monitoring sub-topics of the WMS, as a measure of alignment. Beyond the theoretical precedents in Aghion and Tirole (1997) and related literature, the validity of this particular division of the management scores is supported by a factor analysis of the seven topic scores. In Table A4, we undertake 2-factor principal component analysis using both the organization and individual manager as units of analysis. In both cases, we find that the first factor is positively correlated with all elements of the WMS. However, by far the most divergent element of the second factor is monitoring. It has a correlation of -0.37 with the second factor, while the second most divergent element of the WMS is targeting with a correlation of -0.17. This indicates that the monitoring element of the WMS is picking up aspects of management that are distinct to those in the first

¹⁴To provide a sense of the holistic (rather than specific) nature of these questions, we go through one example: a question relating to management practices relating to monitoring was, “In what kind of ways does your Directorate track how well it is delivering services? Can you give me an example?” Enumerators could then score responses on a continuous 1-5 scale, where for indication the scoring grid described a score of one as corresponding to circumstances where the “Directorate does not track performance”; a score of three corresponded to, “Directorate tracks a number of performance indicators. These are seen and reviewed by senior management only”; and a score of five corresponded to, “Full set of indicators are tracked formally and continuously. Reviews are conducted regularly and involve representatives of all directorate staff groups. The results of the review are formally communicated to all staff.”

factor. The Appendix provides further details on how we constructed our management measures.

Figure A2 plots the distributions of the ‘Management Practice: Monitoring’ and ‘Management Practice: Other’ indices across the organizations we use in our core regressions.¹⁵ Both indices exhibit substantial variation, indicating that de facto management practices vary substantially across the Ethiopian Public Service. For practices related to the organization’s value of the information acquired by officials, the 75th percentile organization has a ‘Management Practice: Monitoring’ score that is 141% higher than the 25th percentile organization. For practices related to the alignment of preferences of agents within an organization, the 75th percentile organization has a ‘Management Practices: Other’ score that is 135% higher than the 25th percentile organization. This variation occurs despite organizations operating under the same broad political environment, the same federal administration, sharing the same national history, and working under the same governing rules over budget allocations, wages, non-wage benefits, promotion schedules and other contracting arrangements. This variation across organizations is consistent with evidence from other management surveys of the public sector (Bloom et al., 2015; Rasul and Rogger, 2018; Rasul et al., 2019).

We do not find evidence that district or public servant characteristics predict the management scores. In Table A5, we regress our aggregated indices of organization-level management practices on district and organization-average individual characteristics. We also regress average tenure in the service in the organization and the proportion of civil servants with an undergraduate education or higher on district characteristics to assess whether there is evidence that observably less experienced/educated civil servants select into or work in districts with more difficult operating environments. We find no evidence of such endogeneity on these observable margins.

3 The Errors of Officialdom

The left-hand panel of Figure 1A provides histograms of both the benchmark data and public officials’ claims regarding the population of the districts that they serve. The histograms are divergent at many points in the distribution, indicating that in general public officials are making errors about the size of the underlying populations they are serving. The right-hand panel of

¹⁵The ‘Monitoring’, and ‘Other’ indices are positively correlated with each other, with a coefficient of 0.62. Importantly, the partial correlations of these two dimensions of management can still be separately estimated from each other.

Figure 1A provides a histogram and density plot of the scale of errors made by public officials in their estimates of the populations they serve.

The distribution of errors in Figure 1A is perhaps as would be expected. On this specific indicator, the average error is close to 0, and the distribution is close to being normally distributed. However, our data allow us to provide estimates of the scale of the standard deviation in the errors of public officials. The dashed vertical line represents the median population of a district in Ethiopia (120,398) to give a sense of the magnitude of the errors for this indicator. Many officials' claims are relatively accurate; 21 percent of public officials guess within 20% of the benchmark data. On the other hand, the scale of error of a substantial fraction of officials' claims is large; 16% of claims by officials over or underestimate the size of the claimed district by more than the population of the median district. 37% of officials overestimate or underestimate the population of the district by more than half the population of the median district. And 47% of officials claim that their district's population is 50% bigger or smaller than it is.¹⁶

We find evidence that public officials adjust their estimates towards the underlying state of nature. The scale of errors is similar across the distribution of district populations, centred on the truth. The left-hand panel of Figure 1B exhibits the scale of errors made by public officials in their claims about districts that have the 25th percentile smallest population in their region, the median, and the 75th percentile. The right-hand panel of Figure 1B exhibits the corresponding proportional errors. That is, the proportion of the benchmark population made up by the claim. In each case, over 46% of officials claim that they are serving a population that is at least 50% smaller or 50% larger than it is. Thus, across the full range of district sizes in Ethiopia, a sizeable fraction of public officials is making substantial errors as to the number of citizens they are serving.

These population errors have real effects on the distribution of the government's budget. The distribution of the national budget across the three tiers of government and across districts is largely determined by a formula based on population, the poverty rate and a district's ability to generate own-source revenue (World Bank, 2012; Assefa, 2015; World Bank, 2016b). In Table A6, we use the most recent budget data available to regress the log of the approved budget of the district on the claims that civil servants make about its population. Conditional on the log of

¹⁶A simple test of distributional knowledge is to ask federal and regional officials whether they can correctly rank the 25th, 50th and 75th percentile districts in their region of expertise in terms of population. In our data, only 18% of officials state the correct ordering for their region (this is true for both federal and regional officials separately). 47% of these central officials cannot accurately state which of the 25th and 75th percentile districts in their region has the larger population.

variables on which the budget is distributed, and even on further characteristics of the districts (remoteness, ethnic fractionalization, the population, the percentage of rural inhabitants, and the poverty rate), we see that population estimates are significant predictors of budget allocations at the 5% level when clustered at the organization and 10% level when clustered at the district. These results imply that districts whose officials overestimate their population by 1-percent receive a 2-percent larger budget over and above the budget specified by the formula.

The distribution of errors among other variables than population is similarly large. The estimates of 16% (23%) of public officials regarding the characteristics of their constituents are at least a standard deviation (half a standard deviation) away from the objective benchmark data we collect. The mean absolute error in education bureaucrats' estimates of primary enrollment numbers is 38% of the true enrollment figures. Across all of the education variables, the mean absolute error is 37% of the benchmark data. The mean error in estimates of the proportion of pregnant women who attended ANC4+ during the current pregnancy (the 'antenatal care rate') was 35% of the benchmark data. Agriculture officials overestimate the number of hectares in their district that are recorded as used for agricultural purposes by almost a factor of 2. As noted above, these magnitudes imply substantial misallocation of resources across space and subject matter. If bureaucrats think that antenatal care is a larger issue than it is, they will devise policies that prioritize corresponding investments beyond those that are justified by the underlying state of the citizenry.

Figure 2 provides cumulative distribution functions of the three aggregated measures of error that we use in this study. Figure 2A presents the distribution of the unadjusted z-scores, with separate plots for Ethiopia's three tiers of government. Since these represent z-scores of absolute errors, a lower z-score implies a more accurate claim. The figure highlights substantial variation in errors across government that echoes the variation in the raw indicators described above and we observe officials making the full distribution of claims. Figure 2B plots the distribution of z-scores where adjustments are made to skewed underlying distributions when necessary. Conditional on differences in the ranges of the distribution produced by the two different methods, the graph is qualitatively similar to that of the unadjusted z-scores. The adjusted scores have a more even distribution of extreme values, as would be expected. Figure 2C plots the rank-proportion transformation, and while district errors continue to be substantially smaller than those at higher tiers, some distinctions arise between the regional and federal government levels. Overall, the three

measurement methods provide highly consistent messages.

We find evidence that these errors correlate with an individual official’s career success. By regressing the data we have collected on individual appraisal scores, we can assess how the magnitude of errors an individual is likely to make in their claims about citizen characteristics correlate with their career progress and a proxy of individual-level productivity.¹⁷ Table A7 presents the results. Column 1 indicates that unconditionally, a standard deviation increase in absolute error is associated with a reduction in appraisal scores of 0.02 of a standard deviation. Controlling for tier (column 2), sector (column 3), indicator (column 4), percentile (column 5), and region (column 6) leads to a coefficient of -0.028, which is significant at the 5% level. This relationship holds if one controls for individual controls (column 7), though the characteristics of officials account for some of the variation, and for differences in organizational characteristics (columns 8 and 9). Thus, information acquisition is positively correlated with incentives in the public sector, and manager’s assessments of official’s productivity.¹⁸

4 Basic Descriptives on the Determinants of Information

With evidence that the tacit beliefs of bureaucrats relate to resource distribution, productivity and career success, we now turn to providing broad descriptive evidence on the determinants of superior information in public hierarchies. The descriptive evidence is some of the most detailed available on the determinants of information held by public officials, and provides a framework in which to undertake the experimental analysis.

We undertake an ANOVA analysis of our error data using variables related to the individual characteristics of bureaucrats, the hierarchical structure of the bureaucracy, and the marginal costs and benefits of information acquisition. The results are presented in Table A8.

We find that individual characteristics of public officials have little explanatory power in this setting. It is true that public officials’ estimates are more accurate in areas we would expect them to have greater expertise. When asked to estimate the population of the median district in

¹⁷To ensure that we take account of differences in scoring systems across organizations, we convert appraisal scores into organization-specific z-scores, and cluster at the district government-sector level. Our sample in this regression is all claims at organizations at which we have appraisal data; 3,963 claims at 135 organizations.

¹⁸In terms of the impact of basic characteristics on appraisal scores, managers perform significantly better on the evaluations, while females achieve significantly lower scores. Education and tenure are not significantly associated with the evaluations at the usual levels. Interestingly, admitting to not knowing the characteristics of the citizenry does not have a significant relationship with an official’s appraisal scores in any of the specifications presented.

their region, they are 10 percentage points less accurate than in estimates of their own district. Similarly, officials are on average 31 percentage-points more accurate in assessing sector-specific variables relevant to their main work tasks relative to estimates of general demographic variables for the district. However, the median claim over sector-specific variables is still 70% larger or smaller than the underlying benchmark data.

Similarly, we find that environmental features, such as specific features of the population and the proximity to local constituents, only explain 1.3% of the variation in errors that we observe. Thus, our findings provide little support to theories in which decentralized agents (such as local governments in a multi-tier system) are assumed to have superior information due to their proximity to local conditions (Oates, 1999; Bardhan, 2002).

Rather, the results in Table A8 imply that structural variables related to control rights, such as the tier of government and organizational incentives explain the dominant proportion of variation in the errors. We therefore focus on the importance of hierarchical structure and organizational incentives in what follows.

4.1 Formal Authority as an Incentive to be Informed

If a manager has de facto authority over an employee, this limits the agent's incentive to invest in information acquisition due to concerns that she will be overruled and the temptation to free-ride on the principal's efforts (Aghion and Tirole, 1997). Relationships in which the agent has formal authority should be characterized by their having superior information on local conditions. In the Ethiopian government, this would imply claims by district officials will be more accurate than those by central government officials. Figure 2A presents graphical evidence that officials at the district tier of government do indeed make lower errors at almost every point in the distribution. A Kolmogorov-Smirnov test of distributional equality is rejected with a p-value of 0.00. Regional and federal organizations have almost identical distributions of errors, though regional organizations are slightly more prone to outliers.

Conversely, within an organization a manager's de facto power limits the non-manager's incentive to absorb new information, reducing the likelihood of an agent having superior knowledge. To investigate the relationship between both types of hierarchy and information empirically, we use an OLS regression of the following form,

$$y_c = \gamma_1 \mathbf{1}\{\mathbf{District} = \mathbf{1}\} + \gamma_2 \mathbf{1}\{\mathbf{Manager} = \mathbf{1}\} + X_c + \epsilon_c \quad (1)$$

where y_c is a dependent variable relating to the scale of the error of the claim, such as the z-score of the absolute error, $\mathbf{1}\{\mathbf{District} = \mathbf{1}\}$ is a dummy that takes the value 1 if the claim is made by a district official, $\mathbf{1}\{\mathbf{Manager} = \mathbf{1}\}$ is a dummy that takes the value 1 if the officer making the claim is a manager, and X_c is a set of sector, district, indicator and percentile fixed effects, individual control variables and noise controls.¹⁹ Individual-level controls are tenure in the civil service of the respondent making the claim; tenure in position; tenure in organization; an indicator for whether the respondent is female; an indicator for whether the respondent holds an undergraduate degree; and an indicator for whether the respondent holds a master’s degree. Given the proximity of the responsibilities of federal and regional governments, as described above, we aggregate these into a single ‘centralized’ tier. When we run the regressions with all three tiers of government, the federal coefficient is not significantly different to the regional at the usual levels, as will be documented below.²⁰

The regressions are at the claim level, so comparisons can be made across individual percentile-indicator pairings, providing the most precise comparison we can offer. In our preferred specification, the variation that identifies the core coefficients is that between two sets of claims, made about the same indicator, at the same percentile of the corresponding distribution, in the same sector, about the same district by individuals at the two levels of hierarchy, conditional on their individual characteristics. Thus, our specification ensures a relatively tight comparison between levels of the hierarchy. We restrict our sample to those claims that are made about the same districts to ensure the variation in our regressions arises only from coherent comparisons. That means the claims used for comparison are made by 594 federal and regional government officials, and 297 district officials. Since there may be correlations in the sectoral information about a particular district, we cluster at this level throughout our analysis.

Table 2 presents the results. In columns 1 and 2, the dependent variable is a binary that takes the value 1 if the individual official stated that they did not know the answer to a question, and

¹⁹Noise controls are composed of the time of day of the survey, day-of-survey fixed effects, enumerator fixed effects, an indicator for the enumerator’s subjective assessment of the quality of the interview, an indicator for the duration of the interview (decile fixed effects), controls for the average tenure and gender of the managers interviewed.

²⁰The fact that there is little difference between the federal and regional tiers is also evidence against physical distance or remoteness playing a critical role in the acquisition of information in this setting (Agarwal and Hauswald, 2010; Landier et al., 2007).

thus could not make a claim. We see in column 1 that district officials are 17 percentage points less likely to answer ‘Don’t Know’, while there is almost no difference between managers and non-managers in readiness to provide an estimate. Controlling for sector, district, indicator, percentile and individual controls in column 2 of Table 2 does not change these results. This implies that if you were to enter offices at different tiers of government, you would face a richer information environment at the district level even conditional on the level of errors. The error results should therefore be interpreted as ‘the perceived state of knowledge within an office’, and imply they represent a lower bound on the difference in knowledge between the two tiers.

Columns 3 to 5 of Table 2 present unconditional versions of (1) where the dependent variable is the z-score of errors across socio-economic variables. In column 3 we include a district government dummy only. We see that across all claims those made by district officials have significantly lower errors than those of central officials. In contrast, column 4 includes only a manager dummy and indicates that there is no difference in the magnitude of errors made by managers and their employees. In column 5 we include both dummies, and find that the results do not change.

This pattern of results is reflected in regressions using variation from within the sector (column 6), constraining comparisons to be based on claims about a common district (column 7), focused on common indicators (column 8) and on common percentiles of those indicators (column 9). Column 10 presents our preferred specification, where we also condition on characteristics of the individual official making the claim. It presents some of the most precise empirical evidence we know confirming a classic assumption in organizational theory; that agents have superior information about local conditions. However, in these data we find that agents with formal authority (those acting in district organizations) have superior information, while those who do not (those acting under a principal within the same office) have no better information than their seniors.²¹ The superior information of agents is a product of a particular form of decentralization rather than a rationale for it. In terms of the geography of information in the public sector, these results imply that information is spread relatively evenly across layers of the hierarchy within an office, but that significant asymmetries occur across offices.²²

²¹In our preferred specification, the interaction between district government and manager is not significant at the usual levels, with a coefficient of -0.07 and a p-value of 0.58 (the coefficient on district government is -0.307 with a p-value of 0.004 and the coefficient on manager is 0.053 with a p-value of 0.52). Thus, at both central and district tiers of government, there is a common information environment within the organization.

²²We find evidence of superior information at the district government level across the entire distribution of claims. Table A9 presents the results from quantile regressions of the error on our measures of hierarchy and fixed effects. The district government indicator is significantly negative at each decile of the distribution of errors.

The coefficient on the district government binary is -0.33, implying that claims made by district officials are on average 0.33 standard deviations closer to the benchmark data than central governments across the z-scores of variables related to citizen characteristics. As can be seen by comparison to the coefficients in Table A8, such institutional features dwarf the association between greater tenure in the service or increased formal education and information acquisition. The coefficient on district government is also an order of magnitude more important than if an MIS system exists within the corresponding organization. If we remove district fixed effects to allow for estimation of the relationship between district poverty rates and errors, we see that the coefficient on the poverty rate (measured from 0 to 100) is 0.008 with a p-value of 0.05. Thus, officials make greater errors in poorer districts, as we might expect. A one percentage-point increase in poverty is associated with a 0.008 standard deviation increase in the absolute error. The size of the coefficient on district government in Table 2 is 41 times larger in absolute terms, such that the quality of information in district organizations is akin to estimating over the 25th percentile of poor districts rather than the 66th.²³

4.2 Heterogeneity by Management Practices

The impact of those control rights can be mitigated by organizational incentives (Aghion and Tirole, 1997). Recent empirical evidence has shown that management affects the distribution of assets across firms such as human capital (Bloom and Van Reenen, 2007; Qian et al., 2015; Bender et al., 2018) and physical capital (Bloom and Van Reenen, 2010). It therefore seems plausible that it would have substantive impacts on the distribution of information.

The diversity in management practices across Ethiopia’s public organizations, and specifically in terms of how highly information is valued as an organizational asset, can be interpreted as variation in the rewards to officials of acquiring information. Organizations that value the use of information, captured by the ‘Management Practices: Monitoring’ index, create a high relative reward for obtaining information. In environments in which organizations value information, captured in the ‘Monitoring’ index, we should expect to see lower errors. Conversely, management practices that raise the alignment of the principal and agents’ interests may lead to a strategic substitution of effort in acquiring information.

²³To assess the robustness of these core results to our methods of measurement, we present a series of robustness checks in Table A10 and discuss them in the Appendix.

We can extend our analysis in (1) to include measures of management practices and assess the consistency of these ideas with our data. To do so, we broaden the variation across which we make comparisons by switching from district fixed effects to regional fixed effects. Our reasoning is that since management practices vary at the organization-level, within a district we constrain the variation to be across 5 organizations. By extending our estimating variation to the region, we are able to capitalize on variation across an average of 15 organizations while still constraining the comparisons of errors within a restricted geographic and institutional environment. We also follow standard practice in studies of the World Management Survey indices, and include measures of organizational capital, in this setting with respect to investments that assist the flow of information.²⁴

To ensure the results of Table 2 tolerate these changes, we present our preferred specification with these modifications in column 8 of Table A10. We see that the coefficient is -0.28 and the p-value is 0.09.

Table 3 presents the results. In Column 1, we include the management indices, the manager dummy, and only those controls equivalent to our preferred specification in Table 2. We see that the coefficients on the management indices accord with theory exactly. Increasing the value an organization places on the use of information, proxied by the ‘Management Practices: Monitoring’ index, decreases errors substantially. The corresponding coefficient is -0.17, and the p-value 0.014. Similarly, the coefficient on the ‘Management Practices: Other’ index is 0.28, and has a p-value of 0.005. Managers continue to have information equivalent to their employees. These results are slightly strengthened by the inclusion of noise controls and proxies of the organization’s investments in capital related to information acquisition, as described in column 2. Combining the indicator of de facto authority, the dummy for district government, and the management practices that proxy incentives leaves the qualitative results consistent with the theoretical literature. Once again, these results indicate the primacy of incentives for determining the distribution of information in the public sector.

In column 4 of Table 3, we undertake a first assessment of whether the association between

²⁴Capital control variables are the proportion of managers in the organization that state that a management information system is in place; the organization average response to “In what proportion of projects would you say information flows effectively between individuals in the same organization?”; the organization average response to “In what proportion of projects would you say information flows effectively between organizations?”; the organization average response of managers to “How many personnel do you manage?”; the organization average response of employees to “How many people would you say regularly give you tasks as part of your formal work duties?”; and, measures of alignment related to the sharing of information.

de facto authority and information varies by the nature of management practices. We include an interaction between the monitoring management practice index and the district government binary. It is negative, and large in magnitude relative to the baseline association with de facto authority, with a p-value of 0.17. We take this heterogeneity analysis further in Figure 4 by plotting the distribution of coefficients of the district government binary as we vary the decile of management practice. The figure makes clearer that the relationship between de facto authority and information is mediated by incentives. At the lowest levels of management practices observed in our setting, where incentives to acquire information are weakest, providing agents with de facto authority has zero association with the relative magnitude of errors. At the highest level of management practices, the corresponding coefficient is -0.66 with a p-value of 0.04. These results imply that unconditional decentralization, without effective corresponding management practices or more broadly incentives for information acquisition, does not improve the information environment. This finding could help reconcile the divergent findings on the impact of decentralization on service delivery (documented in surveys such as Bardhan (2002) or Crawford and Hartmann (2008)), as management practices are typically not measured in such studies.²⁵

Returning to Table 3, we assess whether management practices have a differential effect by whether an official is a manager. Organizational monitoring practices have slightly stronger effects for managers, but the coefficient for non-managers is still negative and significant at the 5% level, while the coefficient relating to the interaction is insignificant at the usual levels.

We can also assess the baseline correlations of management practices that align principals and agents. When the objectives of the principal and agent are closely aligned, this increases the incentives to ‘free-ride’, as each trusts the recommendations of the other, reducing overall investments in information acquisition (Aghion and Tirole, 1997; Myatt and Wallace, 2012). In line with this prediction, alignment in our setting, as measured by the ‘Management Practices: Other’ variable, is positively associated with errors across the columns in Table 3. The coefficient in column 3 is 0.32 and significant at the 1% level, thus large compared with our other coefficients of interest.²⁶ We do not find any evidence of these results varying across the hierarchy. The

²⁵In the setting of our study, we do find evidence that management practices vary across tiers. Table A11 reports on organization-level regressions where our management indices are the dependent variables, the district government binary is the explanatory variable of interest, and a range of controls are presented. We find that once differences in individual characteristics and sector fixed effects are conditioned on, the coefficient on district government is large (0.41) and significant at the 5% level, indicating that on average, district governments have more of a culture of monitoring. This finding is strengthened once we condition on differences in organizational investments in capital related to information acquisition. We also find weaker evidence for other management practices.

²⁶We do not find equivalent baseline effects using our secondary measure of alignment, which directly asks officials

coefficient on an interaction of the manager binary and the ‘Other’ index is 0.00, with a p-value of 0.97. Our results therefore imply that increases in alignment have a broadly negative association with the incentive to acquire information as officials aim to free-ride on their colleagues efforts.

The qualitative conclusions from the results in this section are robust to an alternative grouping of management practices in which monitoring and incentives are aggregated together and all other practices are aggregated separately, to the exclusion of organizational capital variables, and to the inclusion of controls for district environmental characteristics, as shown in Table A12.

5 Experimental Evidence on Information in Hierarchy

5.1 Experimental Design

Our descriptive analysis has implied the importance of organizational incentives in determining the accuracy of bureaucratic beliefs, and in particular those that increase the marginal benefits of information acquisition. We capitalized on the relative feasibility of an experiment in this setting to test the impacts of a reduction in marginal costs on the accuracy of bureaucratic beliefs. This allows us to separately identify the effect of reducing access costs from other changes that come with delegating authority or decentralising activities. We aimed to design an intervention that would *symmetrically* reduce principal and agent costs of information acquisition *dramatically*. Even in the presence of free-riding, a large symmetric reduction in costs of information acquisition can increase the general accuracy of agent beliefs in a hierarchy (Aghion and Tirole, 1997; Garicano, 2000; Gabaix et al., 2006).²⁷

Our intervention is based on a common approach to improving the information of public officials - the evidence briefing. Such briefings are now commonplace as a tool of government or research organizations whose aim is to inform public officials. By condensing research findings, evaluations or survey data into focused communications, evidence briefings aim to inform the decision-making of public officials by reducing the cost of interacting with the raw analysis.²⁸ Government circulars

how aligned they feel. The corresponding coefficient is 0.03 and insignificant at the usual levels.

²⁷Reducing the marginal cost for the principal (agent) only will increase the investment of the principal (agent) but decrease the investment of the agent (principal) due to the strategic substitutability between the two.

²⁸The prevalence of conditions to produce an evidence or ‘policy’ briefing as a requirement of research funding awards implies the importance of the medium in the profession’s approach to influencing public policy. For example, such a requirement is included in most of the UK’s Economic and Research Council grants (<https://esrc.ukri.org/funding/guidance-for-grant-holders/information-for-new-grant-holders/>) and International Growth Centre grants (<https://www.theigc.org/wp-content/uploads/2016/12/IGC-Outputs-Guideline.pdf>).

have been widely used throughout history to provide information to a body of public officials. Recently, forms of evidence briefing have begun to be routinely produced by economics research institutions as diverse as the Bipartisan Policy Center, the Centre for Economic Policy Research, the Institute for Fiscal Studies, JPAL/IPA, and the World Bank Group. The use of evidence briefings is part of a broader effort to make public policy ‘evidence-based’ by informing the beliefs of public officials (Weiss and Bucuvalas, 1980; Smith, 1996; Masset et al, 2013; Pawson, 2006, Parkhurst, 2017).

Prior to the enumeration of the survey, a random sub-sample of the organizations we study were sent the official administrative data corresponding to our policy variables of interest as a government circular. A set of federal organizations received data for all regions for all indicators, highlighting the median district for each indicator for each region. Similarly, a random sub-sample of district jurisdictions were sent the official administrative data for their whole region for all indicators, also highlighting the median district for each indicator. Since the information package to the district-level jurisdictions had to pass through the regional offices, this meant that all regional officers were treated. Table A13 presents a comparison of treatment and control groups across key characteristics of individual officials and the organizations in which they work. Treatment and control groups look balanced across individual and organizational characteristics, differing on a single variable at the usual levels of significance (years in position).

For the treated district jurisdictions, the information package was sent to the Regional Bureau of Public Service with instructions to pass on the data to the randomly selected districts. This primarily occurred through the Woreda Office of Public Service, which then disseminated the information to the offices within the district jurisdiction. Since all regions were asked to communicate with the district administrations of the arrival of the survey team in advance (to prepare the logistics, in terms of office space, in terms of timing etc), this rules out concerns that the letter itself, and not the information within the letter, is driving the treatment effects of the experiment.²⁹ Letters and data packages were typically sent two weeks before the survey from May 2016 onwards. The first surveys were conducted in mid-June and lasted until mid-September.

The information package was designed to accord exactly with standard communications of the Ethiopian Public Service. The package, and accompanying letter, was designed in close collabo-

²⁹Beyond the chief executive of the organization (a level of hierarchy above those managers we include in our study) there is no ex-ante reason to believe that other officials would connect the survey and the circular. It is standard practice to send circulars from central government in this context for circulation amongst staff. Such circulars are frequently empirical in nature and our design was aimed at mimicking such standard documents.

ration with the Ministry of Public Service and Human Development, and was sent through the standard bureaucratic channels. Thus, it was representative of broader service communications, and as far as we understand it would have been received as such. The accompanying letter stated that, “we have assembled the most up-to-date administrative data available to provide an understanding of the context your office works within. We would like to provide you with this information as an input to improved service delivery within your organization.”³⁰ Since the package was shared as a circular, the expectation was for it to be shared widely among staff within a district office.

The data packages make information on core variables of national policy available to a random subset of Ethiopian officials in a form they are used to at close to zero cost. The package is a clear reduction in the marginal cost of acquiring information. Officials must then decide to read and absorb the information, and make that decision in the wider incentive environment we have documented above.

5.2 Experimental Results

Table 4 presents the results of the intervention on the errors made by officials in their claims about the socio-economic characteristics of the citizenry they serve. The regression framework has the same structure as (1), the sample is as in previous tables, and the dependent variable is once again a function of the absolute errors made by officials. We include a binary treatment variable that takes the value 1 if the organization received the data package, and 0 if it did not.

Column 1 of Table 4 presents the unrestricted treatment effect, which does not restrict the comparison of claims to be within indicator, percentile, district, and so on. We see that it is positive and statistically insignificantly different from zero at the usual levels. Throughout Table 4, we present p-values of the treatment effect with two distinct approaches to clustering. In the square and curly brackets below each treatment coefficient we follow the rest of the paper in clustering at the district government-sector level. These clusters correspond to the coherent sector teams within the district-government in which many information distribution choices would most likely be made. However, we note that the treatment is at the district level, and therefore we also present the p-value of the treatment coefficient from the same regression but clustered at the district level. In neither case is the fully unconditional treatment effect in column 1 significant at

³⁰Since the information provided in the circular was a set of statistics about the jurisdictions themselves, they do not suffer from issues of replicability or population validity, a concern with the extrapolation of evidence from other settings (Pritchett and Sandefur, 2015).

the usual levels.

In column 2 of Table 4, we condition on the set of controls used in previous tables, where the treatment effect relates to a comparison within indicator, percentile, sector and region, with individual and organizational controls. The impact of the data package is now negative and significant at the 10% level (a p-value of 0.08 using our standard clustering approach; a p-value of 0.06 clustering at the level of treatment). Its magnitude is roughly half of the coefficient on the district government dummy in Table 2, and large compared to the control variables used for comparisons above. It seems that officials have taken the opportunity of the reduction in the cost of information acquisition and increased their knowledge of the constituencies in which they work.³¹

Given that we have documented the importance of de facto authority, we control for our core explanatory variables from Table 2 in column 3. The treatment effect jumps once we control for these critical drivers of variation in claim error. The coefficient, -0.28, has a p-value of 0.004 when using our standard clustering approach and 0.003 when clustering at the treatment level. Once we control for treatment, the coefficient on the indicator for whether claim is made by a district official increases to -0.44, and the coefficient on manager continues to be small, which is evidence that the experiment is not driving our results in Tables 2 and 3. A final set of controls is the addition of the management indices assessed in Table 3, ‘Management Practices: Monitoring’ and ‘Management Practices: Other’. The treatment effect continues to be large, with a coefficient of -0.23, and corresponding p-values of 0.04 with our standard clustering approach and 0.03 when clustering at the treatment level. The management indices both have the expected signs and are significant at the 5% level. Thus, controlling for the key drivers of information acquisition, we see that a substantial reduction in transaction costs can improve bureaucratic knowledge.

The treatment effect is robust to using the different measurement techniques for aggregating absolute error. Using proportional rank of error as the dependent variable, shown in column 5, we find a treatment effect of -0.074, with corresponding p-values of 0.03. The effect weakens slightly when we use the skewness-adjusted z-score, but is still significant at the usual levels (column 6). Finally, we note that since the data package was sent by sectoral officials at the regional level, we should also test the robustness of the effect to any correlations between activities within the corresponding regional office. We therefore cluster at the region-sector level and the p-value raises

³¹Table A14 adds in each set of controls one by one to show how the results change between Table 4 Columns 1 and 2.

to 0.048 (column 7).

Together, the results support the prediction that substantial reductions in the cost of acquisition for officials can improve the accuracy of their information. This could be related to external efforts to improve the knowledge of public policymakers by presenting key findings of research in digestible, ‘low-cost’ ways, and by ‘translating’ research findings into a medium of communication that reduces the transaction costs to public servants of absorbing relevant messages. The magnitude of the coefficients in Table 4 are large relative to the descriptive impacts of incentives documented in Tables 2 and 3, and substantially larger than many contextual factors, such as making claims over higher poverty districts. Providing a data pack to an organization improves the accuracy of its officials’ claims by a similar magnitude to if they were estimating the conditions of the 5th percentile of poor districts rather than the 90th percentile.

However, as we have seen in previous tables, the conditions under which public servants work are highly variable, and estimates relevant to the ‘average’ bureaucrat may not provide a full picture of the impacts of our treatment. In Table 5, we therefore assess the heterogeneity of the treatment effect along the critical margins for information acquisition outlined throughout the paper. Column 0 presents our baseline treatment effect from Table 4 for comparison.

Column 1 presents the baseline specification in column 0 with the addition of an interaction between the treatment binary and the district government binary. The coefficient is not significant at the usual levels, with a corresponding p-value of 0.14. Similarly, column 2 presents the baseline specification with the addition of an interaction between the manager and treatment indicators. The interaction is once again not significant at the usual levels, and the p-value 0.37. There is thus no evidence that treatment varied substantially by the level of authority. As expected, in both authority regimes, a large decrease in the cost of information acquisition increases the equilibrium level of knowledge.³²

Where we should expect to find results is in the incentives associated with information. In equilibrium, agents weigh up the marginal benefit of acquisition with the cost. Where benefits are low, small changes in the cost of acquisition can have large impacts on the equilibrium calculus for an agent, and our treatment should have large effects. Where the benefits of holding knowledge are high, information holdings will already be substantial, and thus small changes in costs will lead to limited additional information absorption. We test these ideas in column 3 of Table 5,

³²Interestingly, there is no evidence of a shift in the distribution of information between managers and employees. This is consistent with the treatment having been applied consistently within offices.

where we interact our treatment effect with the ‘Management Practices: Monitoring’ index. The results closely match the theoretical prediction. At the mean level of management, we observe a similar treatment effect to the baseline. However, there is a large positive interaction effect, with a coefficient of 0.34, significant at the 1% level. As management practices related to monitoring improve, the mean treatment effect is attenuated. This can be seen clearly in Figure 5A, where we plot the treatment effect at the deciles of the monitoring management practice index based on the specification of column 3 in Table 5. At the lowest levels of management observed in our data, the treatment effect is extremely large. The coefficient is -0.87 (p-value 0.00), which is almost three times the average effect of a claim being made by a district official. However, from the 60th percentile of management practices related to monitoring upwards, the treatment effect is indistinguishable from zero at the 5% level. The best managed organizations incentivize their officials such that they already have sufficiently precise information that our intervention has no effect.

In column 4 of Table 5, we interact the treatment binary with the ‘Management Practices: Other’ index. Once again, the interaction is positive and significant at the 1% level. The treatment effect at the mean level of the management index is negative, implying that once again, a higher score on the management index attenuates our treatment effect towards zero. This can be clearly seen in Figure 5B, in which we plot the treatment effect at the deciles of the ‘Other’ management index, with all the controls of the specification in column 4 of Table 5. The figure presents a distribution of coefficients similar to that for monitoring, with substantial treatment effects at low levels of management, and a null effect at the 60th percentile onwards. Once again, this is consistent with theory. At low levels of staff alignment, agents will be more responsive to reductions in costs of information acquisition as there is little crowding out from free-riding incentives. However, as the organization implements management practices that align actors’ incentives more closely, the incentive to absorb information lessens and agents aim to free-ride on each other’s efforts. Overcoming free-riding concerns with a reduction in acquisition costs becomes harder. These results are robust to alternative groupings of management, to the exclusion of organizational capital variables, and to the inclusion of local government characteristics, as presented in Table A15.

Together, these interactions are consistent with the idea that organizational incentives are important mediators of the data package treatment. Where incentives for information absorption

are strong, officials will source information relevant for their core work without external intervention. Where incentives are weak, there is room for external intervention to improve bureaucratic knowledge. In these poorly managed environments, external provision of low-cost information can have impacts on the magnitude of officials' errors equivalent to or greater than the design of incentives. However, improving the underlying incentives for information acquisition would yield superior information outcomes without external intervention.

6 Conclusions and Further Research

This paper develops a framework for measuring the accuracy of public officials' beliefs about their constituents and applies it to a large, hierarchical public sector.³³ We provide descriptive and experimental evidence on the scale of the errors in these beliefs and their determinants. We find that many public officials make large errors about the characteristics of the citizens they serve, and that these errors correlate with the distribution of government budgets and assessments of individual productivity. Both our descriptive and experimental results underline the importance of incentives for information acquisition in public sector hierarchies, in line with theoretical predictions from the control rights literature. As such, we add to a burgeoning literature on bureaucrats and their importance in government productivity (such as Dal Bo et al. (2013) and Bertrand et al. (2016)).

Our descriptive evidence compares errors made about the same socio-economic indicator, referring to the same district, at the same percentile of the corresponding distribution by individuals at two levels of hierarchy in the same sector, conditional on their individual characteristics. It therefore presents the most precise empirical evidence we know confirming a classic assumption in organizational theory; that decentralized agents have superior information about local conditions. By combining these results with data on the characteristics of officials, their organizations and the environmental features of their constituencies, we provide detailed descriptive evidence on what the conditions under which this empirical finding holds.

In particular, a large literature on decentralization invokes the assumption that decentralized agencies have superior information relative to their centralized counterparts (Oates, 1999, 2005; Bardhan and Mookherjee, 2006). While we do find that decentralized agencies in our data have superior information, this is only true for organizations that are characterized by management

³³While the empirical setting of this paper is the public sector, it is reasonable to assume that the same results would be observed where the same hierarchical conditions exist in the private sector.

practices that induce decentralized officials to acquire information. The difference in errors between centralized and decentralized agencies without these management practices is zero. The finding that the impact of decentralization on information is mediated by management practices provides one potential explanation for the divergent findings of the decentralization literature on service delivery (Bardhan, 2002; Crawford and Hartmann, 2008). Additionally, we do not find any corresponding evidence that employees have superior information relative to managers. In our setting, we find a common informational environment within organizations, but not between them.

Our descriptive analysis implies the importance of organizational incentives in determining the accuracy of bureaucratic beliefs, and in particular those that increase the marginal benefits of information acquisition. We capitalize on the relative feasibility of an experiment to test the impacts of a reduction in marginal costs on the accuracy of bureaucratic beliefs through the provision of an evidence briefing to a random sub-sample of the officials we study. We make precise the incentive environments in which evidence briefings have impacts on bureaucratic beliefs. We find that such evidence briefings, now a common tool of economics researchers who aim to influence public policy, can improve the accuracy of officials' beliefs. However, once again this result is significantly mediated by organizational incentives for information acquisition. Thus, whilst there is much to learn from existing work on the impact of evidence provision to public officials, our results imply that average impacts across the public service may mask important heterogeneity in the distribution of that learning.

Our paper therefore provides a framework for understanding bureaucratic constraints to knowledge diffusion in the public sector. While Brynjolfsson and McElheran (2016) document the rapid adoption of data-driven decision making in the private sector in the United States, World Bank (2016a) documents the slow rate of adoption of similar projects in public sectors across the world. Much policy-focused research aims to translate its findings into a form digestible by public officials. Our results imply that embedding knowledge into public organizations is a function of the de facto authority and corresponding incentive environment in which a public official works as much as it is a function of limiting the costs of absorption.

References

- [1] ABAGISA, J. (2014) “Challenges and Prospects of Balanced scorecard Implementation in the Development Bank of Ethiopia,” Mimeo.
- [2] ACEMOGLU, D., P.AGHION, C.LELARGE, J.VAN REENEN AND F.ZILIBOTI (2007) “Technology, Information, and the Decentralization of the Firm,” *Quarterly Journal of Economics* 122 (4): 1759-1799.
- [3] ACEMOGLU, D. AND J. A.ROBNSON (2012) “Why Nations Fail: The Origins of Power, Prosperity and Poverty,” New York: Crown Business
- [4] AGHION, P. AND J.TIROLE (1997) “Formal and Real Authority in Organizations,” *Journal of Political Economy* 105 (1): 1–29.
- [5] AGHION, P., N.BLOOM, B.LUCKING, R.SADUN AND J.VAN REENEN (2017) “Turbulence, Firm Decentralization and Growth in Bad Times,” National Bureau of Economic Research Working Paper No. 23354.
- [6] ANDRADE, P. AND H.LE BIHANA (2013) “Inattentive Professional Forecasters,” *Journal of Monetary Economics* 60 (8): 967-982.
- [7] ARGENZIANO, R., S.SEVERINOV AND F.SQUINTANI (2016) “Strategic Information Acquisition and Transmission,” *American Economic Journal: Microeconomics* 8 (3): 119-55.
- [8] AGARWAL, S. AND I.BEN-DAVID (2018) “Loan Prospecting and the Loss of Soft Information,” *Journal of Financial Economics* 129 (3): 608-628.
- [9] AGARWAL, S. AND R.HAUSWALD (2010) “Distance and Private Information in Lending,” *Review of Financial Studies* 23 (7): 2757-2788.
- [10] AGARWAL, S. AND R.HAUSWALD (2009) “Authority and Information,” Mimeo.
- [11] ASSEFA, D. (2015) “Fiscal Decentralization in Ethiopia: Achievements and Challenges,” *Public Policy and Administration Research*, 5 (8): 27-39.
- [12] BANDIERA, O., M.BEST, A.KHAN AND A.PRAT (2018) “Incentives and the Allocation of Authority in Organizations: A Field Experiment with Bureaucrats,” Mimeo: London School of Economics.
- [13] BANURI, S., S.DERCON AND V. GAURI (2017) “Biased policy professionals,” World Bank Policy Research Working Paper 8113.

- [14] BANERJEE, A., R.HANNA, J.KYLE, B. A.OLKEN AND S.SUMARTO (2018) “Tangible Information and Citizen Empowerment: Identification Cards and Food Subsidy Programs in Indonesia,” *Journal of Political Economy* 126 (2): 451-491.
- [15] BARDHAN, P. (2002) “Decentralization of Governance and Development,” *Journal of Economic Perspectives* 16 (4): 185-205.
- [16] BARDHAN, P. AND D.MOOKHERJEE (2006). “Decentralisation and Accountability in Infrastructure Delivery in Developing Countries,” *Economic Journal*, 116 (508): 101-127.
- [17] BENDER, S., N.BLOOM, D.CARD, J.VAN REENEN AND S.WOLTER (2018) “Management Practices, Workforce Selection, and Productivity,” *Journal of Labor Economics* 36 (S1): S371-S409.
- [18] BERTRAND, M., R.BURGESS, A.CHAWLA AND G.XU (2016) “The Costs of Bureaucratic Rigidity: Evidence from the Indian Administrative Service,” University of Chicago Working Paper.
- [19] BESLEY, T. AND T. PERSSON (2011) “Pillars of Prosperity: The Political Economics of Development Clusters,” Princeton: Princeton University Press.
- [20] BLOOM, N. AND J.VAN REENEN (2007) “Measuring and Explaining Management Practices Across Firms and Countries,” *Quarterly Journal of Economics* 122 (4): 1351-1408.
- [21] BLOOM, N., B.EIFERT, A.MAHAJAN, D.MCKENZIE AND J.ROBERTS (2013) “Does Management Matter? Evidence from India,” *Quarterly Journal of Economics* 128 (1): 1–51.
- [22] BLOOM, N., R.SADUN AND J.VAN REENEN (2012) “The Organization of Firms Across Countries,” *Quarterly Journal of Economics* 127 (4): 1663-1705.
- [23] BLOOM, N., C.PROPPER, S.SEILER AND J.VAN REENEN (2015) “The Impact of Competition on Management Practices in Public Hospitals,” *Review of Economic Studies* 82 (2): 457–489.
- [24] BROWN, M., W.SCHALLER, S.WESTERFIELD AND M.HEUSLER (2012) “Information or Insurance? On the Role of Discretion in Relationship Lending,” University of St. Gallen Working Paper.
- [25] BRYNJOLFSSON, E. AND K.MCELHERAN (2016) “The Rapid Adoption of Data-Driven Decision-Making,” *American Economic Review* 106 (5): 133-39.
- [26] CAMPBELL COLLABORATION (2019) “The Campbell Collaboration Online Library,” Available at <https://campbellcollaboration.org/library.html>(Accessed 4th February 2019)
- [27] CALLEN, M., S.GULZAR, A.HASANAIN, M.Y.KHAN AND A.REZAEI (2018) “Data and Policy Decisions: Experimental Evidence from Pakistan,” Stanford University Working Paper.
- [28] CRAWFORD, G. AND C.HARTMANN (2008) “Decentralisation in Africa: A Pathway Out of Poverty and Conflict?” Amsterdam: Amsterdam University Press.
- [29] CRAWFORD, V. P. AND J.SOBEL (1982). “Strategic Information Transmission,” *Econometrica: Journal of the Econometric Society*: 1431-1451.

- [30] CREMER, J., A.ESTACHE AND P.SEABRIGHT (1994) “The Decentralization of Public Services: Lessons from the Theory of the Firm,” World Bank Policy Research Working Paper 1345.
- [31] D’AGOSTINO, R.B., A.J.BELANGER AND R.B.D’AGOSTINO JR (1990) “A Suggestion for Using Powerful and Informative Tests of Normality,” *American Statistician* 44 (4): 316-321.
- [32] DAL BO, E., F.FINAN AND M.A.ROSSI (2013) “Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service,” *Quarterly Journal of Economics* 128 (3): 1169-1218.
- [33] DAL BO, E., F.FINAN, N.Y.LI AND L.SCHECHTER (2018) “Government Decentralization Under Changing State Capacity: Experimental Evidence from Paraguay,” National Bureau of Economic Research Working Paper No. 24879.
- [34] DESSEIN, W. (2002) “Authority and Communication in Organizations,” *Review of Economic Studies* 69 (4): 811-838.
- [35] DHALIWAL, I. AND R.HANNA (2017) “The Devil is in the Details: The Successes and Limitations of Bureaucratic Reform in India,” *Journal of Development Economics* 124: 1-21.
- [36] DI PEI, H. (2015) “Communication with Endogenous Information Acquisition,” *Journal of Economic Theory* 160: 132-149.
- [37] DUFLO, E. (2017) “Richard T. Ely Lecture: The Economist as Plumber,” *American Economic Review* 107 (5): 1-26.
- [38] DUGGAN, M. AND F.M.SCOTT MORTON (2006) “The Distortionary Effects of Government Procurement: Evidence from Medicaid Prescription Drug Purchasing,” *Quarterly Journal of Economics* 121 (1): 1-30.
- [39] EHRBECK, T. AND R.WALDMANN (1996) “Why are Professional Forecasters Biased? Agency Versus Behavioral Explanations,” *Quarterly Journal of Economics* 111 (1): 21-40.
- [40] ELLIOTT, H. AND J.POPAY (2000) “How Are Policy Makers Using Evidence? Models of Research Utilisation and Local NHS Policy Making,” *Journal of Epidemiology & Community Health* 54 (6): 461-468.
- [41] FINAN, F., B.OLKEN AND R.PANDE (2017) “The Personnel Economics of the Developing State,” *Handbook of Economic Field Experiments* 2: 467-514
- [42] FISHER, J.R.B., J.MONTAMBAULT, K.P.BURFORD, T.GOPALAKRISHNA, Y.J.MASUDA, S.M.W.REDDY, K.TORPHY AND A.I.SALCEDO (2018) “Knowledge Diffusion Within a Large Conservation Organization and Beyond,” *PLoS ONE* 13(3).
- [43] GABAIX, X., D.LAIBSON, G.MOLOCHE AND S.WEINBERG (2006) “Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model,” *American Economic Review* 96 (4): 1043-1068
- [44] GANT, J., C.ICHNIOWSKI AND K.SHAW (2002) “Social Capital and Organizational Change in High-Involvement and Traditional Work Organizations,” *Journal of Economics & Management Strategy* 11 (2): 289-328.

- [45] GARICANO, L. (2000) “Hierarchies and the Organization of Knowledge in Production,” *Journal of Political Economy* 108 (5): 874-904.
- [46] HANNA, R., S.MULLAINATHAN AND J.SCHWARTZSTEIN (2014) “Learning Through Noticing: Theory and Evidence from a Field Experiment,” *Quarterly Journal of Economics* 129 (3): 1311–1353.
- [47] HARRIS, R.D.F. AND E.TZAVALLIS (1999) “Inference for Unit Roots in Dynamic Panels Where the Time Dimension is Fixed,” *Journal of Econometrics* 91 (2): 201–226
- [48] HASTINGS, J.S. AND J.M.WEINSTEIN (2008) “Information, School Choice, and Academic Achievement: Evidence from Two Experiments,” *Quarterly Journal of Economics* 123 (4): 1373-1414.
- [49] HATTORI, M., K.SHINTANI AND H.UCHIDA (2015) “The Repository of Soft Information Within Bank Organizations,” *Journal of Money, Credit and Banking* 47 (4): 737-770.
- [50] HJORT, J., D. MOREIRA, G. RAO AND J. F. SANTINI (2019) “How Research Affects Policy: Experimental Evidence from 2,150 Brazilian Municipalities,” National Bureau of Economic Research Working Paper No. w25941.
- [51] IM, K.S., M.H.PESARAN AND Y.SHIN (2003) “Testing for Unit Roots in Heterogeneous Panels,” *Journal of Econometrics* 115 (1): 53–74.
- [52] JESSOE, K. AND D.RAPSON (2014) “Knowledge Is (Less) Power: Experimental Evidence from Residential Energy Use,” *American Economic Review* 104 (4): 1417-38.
- [53] LANDIER, A., V.B.NAIR, AND J.WULF (2007). “Trade-Offs in Staying Close: Corporate Decision Making and Geographic Dispersion,” *Review of Financial Studies* 22 (3): 1119-1148.
- [54] LIAQAT, A. (2019) “Representation through Information: Politician Responsiveness to Citizen Preferences,” Mimeo.
- [55] LIBERTI, J.M. (2018) “Initiative, Incentives, and Soft Information,” *Management Science* 64 (8): 3714-3734.
- [56] MASAKI, T., S.CUSTER, A.ESKENAZI, A.STERN AND R.LATOURELL (2017) “Decoding Data Use: How Do Leaders Source Data and Use it to Accelerate Development?” Williamsburg, VA: AidData at the College of William & Mary.

- [57] MASSET, E., M. GAARDER, P. BEYNON AND C. CHAPOY (2013) “What is the impact of a policy brief? Results of an experiment in research dissemination,” *Journal of Development Effectiveness*, 5 (1): 50-63.
- [58] MOOKHERJEE, D. (2006) “Decentralization, Hierarchies, and Incentives: A Mechanism Design Perspective,” *Journal of Economic Literature* 44 (2): 367-390.
- [59] MOOKHERJEE, D. (2015) “Political Decentralization,” *Annual Review of Economics* 7: 231-249.
- [60] MURALIDHARAN, K. AND V.SUNDARARAMAN (2011) “Teacher Performance Pay: Experimental Evidence from India,” *Journal of Political Economy* 119 (1): 39-77.
- [61] MURALIDHARAN, K., P.NIEHAUS AND S.SUKHTANKAR (2016) “Building State Capacity: Evidence from Biometric Smartcards in India,” *American Economic Review* 106 (10): 2895-2929.
- [62] MYATT.D, P. AND C.WALLACE (2012) “Endogenous Information Acquisition in Coordination Games,” *Review of Economic Studies* 79 (1): 340-374
- [63] MYERSON, R.B. (1981) “Optimal Auction Design,” *Mathematics of Operations Research* 6 (1): 58-73.
- [64] NELLIS, G., T. DUNNING, G. GROSSMAN, M. HUMPHREYS, S.HYDE, C. MCINTOSH AND C. REARDON (2019) “Learning about cumulative learning: An experiment with policy practitioners,”Mimeo.
- [65] OATES, W.E. (1999) “An Essay on Fiscal Federalism,” *Journal of Economic Literature* 37 (3): 1120-1149.
- [66] OATES, W.E. (2005) “Toward A Second-Generation Theory of Fiscal Federalism,” *International Tax and Public Finance* 12 (4): 349-373.
- [67] OLKEN, B.A. (2007) “Monitoring Corruption: Evidence from a Field Experiment in Indonesia,” *Journal of Political Economy* 115 (2): 200-249.
- [68] PAWSON, R. (2006) “Evidence-Based Policy: A Realist Perspective,” London: Sage Publications.

- [69] PARKHURST, J. (2017) "The Politics of Evidence: From Evidence-Based Policy to the Good Governance of Evidence," Routledge Studies in Governance and Public Policy, Abingdon: Routledge.
- [70] PEPINSKY, T.B., J.H. PIERSKALLA AND A.SACKS (2017) "Bureaucracy and Service Delivery," *Annual Review of Political Science* 20: 249-268.
- [71] PETERSEN, M.A. (2004) "Information: Hard and Soft," Mimeo.
- [72] POMERANZ, D. (2015) "No Taxation Without Information: Deterrence and Self-Enforcement in the Value Added Tax," *American Economic Review* 105 (8): 2539-69.
- [73] PRITCHETT, L. AND J.SANDEFUR (2015) "Learning from Experiments When Context Matters," *American Economic Review* 105 (5): 471-75.
- [74] QIAN, J., P.E.STRAHAN AND Z.YANG (2015) "The Impact of Incentives and Communication Costs on Information Production and Use: Evidence from Bank Lending," *Journal of Finance* 70 (4): 1457-1493.
- [75] RASUL, I. AND D.ROGGER (2018) "Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service," *Economic Journal* 128: 413-446.
- [76] RASUL, I., D.ROGGER AND M.WILLIAMS (2019) "Management and Bureaucratic Effectiveness: Evidence from the Ghanaian Civil Service," Mimeo, University College London.
- [77] ROYSTON, P. (1992) "Comment on sg3. 4 and an Improved D'Agostino Test," *Stata Technical Bulletin* 1(3).
- [78] SMITH, A.F.M. (1996) "Mad Cows and Ecstasy: Chance and Choice in an Evidence-Based Society," *Journal of the Royal Statistical Society* 159 (3): 367-383.
- [79] SOMANI, R. (2018) "Information and Public Service Delivery: Evidence from an E-Governance Initiative," Mimeo, University College London.
- [80] STEIN, J.C. (2002) "Information Production and Capital Allocation: Decentralized Versus Hierarchical Firms," *Journal of Finance* 57 (5): 1891-1921.
- [81] STIGLITZ, J.E. (2017) "The Revolution of Information Economics: The Past and the Future," National Bureau of Economic Research Working Paper No. 23780.
- [82] TEREDA, A. (2014) "Design and Implementation of Balanced Scorecard in the Development Bank of Ethiopia," Mimeo.
- [83] UCHIDA, H., G.F.UDELL AND N.YAMORI (2012) "Loan officers and relationship lending to SMEs," *Journal of Financial Intermediation* 21 (1): 97-122.
- [84] VIVALT, E. AND A. COVILLE (2019) "How do policymakers update," Mimeo.

- [85] WEISS, C. H. AND M. J. Bucuvalas (1980) “Social Science Research and Decision-Making,” Columbia University Press.
- [86] WORLD BANK (2012) “Better Results from Public Sector Institutions,” Washington, DC: World Bank
- [87] WORLD BANK (2016a) “World Development Report 2016: Digital Dividends,” Washington, DC: World Bank.
- [88] WORLD BANK (2016b) “Ethiopia Public Expenditure Review,” Washington, DC: World Bank.

Appendix

What is the Appropriate Conception of Information in the Public Sector?

Errors in bureaucratic beliefs matter because they are used as the basis of public policy decisions. In our representative survey of public officials, we asked public servants, “What are the three most important sources of information you use to find out about the state of service delivery in your jurisdiction?” along the following margins: ‘Informal interactions e.g. through co-workers’, ‘Management Information Systems’, ‘Formal field visits’, ‘Reports from the frontline’, ‘Information from higher-tier institutions’, ‘Formal citizen/user interactions’, or ‘Media sources’. Officials were asked to rank their top three sources or specify that others are more important.³⁴

In our setting, information is dominantly sourced from inside the public service rather than externally from the media. Only 8.93% of public officials in our survey responded that media sources were one of their top three sources of information, and 2.12% of officials selected it as their top source of information. This finding was consistent across tiers, with 8.54% of federal, 8.56% of regional and 9.18% of district officials stating media was one of their top-3 information sources, across sectors and across grades. That public officials are heavily reliant on information produced by the public service is consistent with the findings of AidData’s ‘Decoding Data Use’ project

³⁴Only 4% of officials stated that one of their top three sources of information was not included in our survey options.

(Masaki et al., 2017). Similarly, a review of Ethiopia's 'Sector Development Program IV: 2010-2015' documents for the five sectors studied in this paper finds almost no references to external sources.

The most frequently cited source of information was 'Formal field visits', with 63% of officials stating that this was a key source of information. Discussions with frontline colleagues, and informal interactions with colleagues were the second and third most cited source of information, with 51.9% and 45.9% of officials stating their significance respectively. A similar pattern is observed restricting the selections to the top cited source of information for each individual: formal field visits are most frequently cited (22.3%), followed by informal interactions with colleagues (19.6%) and then reports from the frontline (17.2%).

These responses indicate that the dominant form of information used for decision making in the Ethiopian public sector is in terms of knowledge transmitted or gained through personal experience, widely known as the 'tacit' knowledge of individuals. Together, information gained through the direct experience of the individual - from field visits or interacting with colleagues or citizens - is one of the top three sources of information collection for 90% of the officials in our data. Only 12.8% of officials state that they use management information systems (MIS) as their primary source of information. Field visits and informal interactions are 3 times more likely to be in the top 3 most important sources of information than MIS, and 10 times more likely than external media sources.³⁵ These findings for Ethiopia's public service are consistent with wider findings from the World Development Report 2016 that highlighted how public officials across the developing world fail to capitalize on digital databases available to them.³⁶

This lack of use of MIS does not seem to be due to a lack of skills. Our survey asked the extent to which officials agreed with the statement, "I possess the necessary skills to make best use of the technology that is available to me." Two-thirds of respondents stated that they 'slightly' or 'strongly' agreed with the statement. Similarly, education indicators explain less than 1% of

³⁵The reliance on tacit knowledge is in contrast to the fact that many public officials believe relevant information is available from other sources. We asked officials how long it would take to receive statistical data as an input to their work from the relevant directorate, and the median response was 1 day (the mean was 1.8 days and the 90th percentile was 2 days). Averaging bureaucrat responses to the organization level, the 90th percentile slowest organization delivered the data within 3 days.

³⁶Many public officials believe that the tacit information they have is sufficient for the successful implementation of their daily work. Our respondents stated that in 70.9% of the projects they worked on, "I have necessary information regarding the current state of service delivery in my jurisdiction/my area of work". This was consistent across tiers (federal, 73.6%; regional, 72.9%; district, 69.4%), sectors (Agriculture, 71.1%; Education, 75.3%; Health, 74.3%; Revenue, 66.1%; Trade, 67.6%) and grades (Grades 1-5, 71.1%; Grades 6+, 70.7%).

the variation in the probability that officials state that they use informal interactions in their top 3 most important sources of information. The lack of use of MIS also does not seem to be due to variation in facilities such as hours of electricity an organization typically receives, the quality of the phone network or internet access. An aggregate index of the quality of facilities at the organizations we study also explains less than 1% of the variation in the likelihood of relying on tacit knowledge.³⁷ Rather, the culture of information use leans strongly towards a reliance on tacit knowledge.

Measuring the Errors of Officialdom

This section describes in detail the creation of the error variables, their relative scale, and a descriptive assessment of their basic determinants. For each individual surveyed, we obtained their subjective assessments ('claims') of demographic characteristics of their constituency, of the sector-specific variables relating to their constituency, and of the variables relating to the public service rules.

Where individuals stated that they did not know the answer at all, a value of 0 was entered to signal that the respondent did not know the answer to a particular question at all. Of the 891 individuals who make up our common-claims sample used in the core analysis, 275 (31%) of individuals respond 'Don't Know' to at least one of the questions asked. 181 individuals (20%) respond 'Don't Know' to all of the questions in the information module. There is, therefore, selection into the sample of respondents who provide non-zero estimates. Regressing individual characteristics on a binary variable that takes the value 1 if a respondent answers 'Don't Know' to all questions, we find evidence that district officials are less likely to do so. We make two comments regarding the selection into the error sample: firstly, it seems a reasonable assumption that the probability of responding 'Don't Know' is positively associated with the most extreme (latent) errors, and hence our estimates of the difference across tiers can be interpreted as a lower bound. Secondly, the error data provide consistent estimates of the accuracy of information that is present within a public office in Ethiopia – for example, if one entered a public office at random in Ethiopia and asked about a particular service delivery indicator, the information that one would

³⁷The index of facilities is an organization-level average of z-scores of the facilities questions asked to managers only: how many hours of electricity an organization receives per typical working day, the state of the phone network, the extent of internet access, the access staff have to computers, the access staff have to vehicles, and the staff's ability to write memos/spreadsheets/powerpoint.

receive would be reflected within the error data we use for our regressions. To ensure that these 0 claims do not affect the analysis on the errors, a missing value is inserted in place of the 0 for the error variables. Furthermore, a ‘Don’t Know’ binary indicator is created that takes the value 1 whenever a respondent does not know the answer to a question.

Our benchmark data is regarded as the best available in the Ethiopian context. One proxy of the quality of the benchmarking data is their stability over time. By assessing the extent to which indicators vary within districts over data collection rounds, we aim to understand whether the data reflect a noisy collection process. For example, if the ratio of pupils per school has a high variance from one year to the next within the same district, this could be seen as a warning sign as to the quality of measurement. Table A2 presents summary statistics and tests for the presence of a random walk for each of the indicators used in the analysis where panel data are available. Column 4 shows the estimated autoregressive parameter from a Dickey Fuller Regression, controlling for district fixed effects and column 5 shows the estimated autoregressive parameter from a Dickey Fuller Regression, controlling for district fixed effects and district-specific linear time trends. The estimated coefficients are less than 1, which indicate stability in the trend, for all indicators except for three: the amount of land used for agricultural purposes; the amount of land used for pastoral purposes; and the number of business licenses issued. In addition, column 6 reports the z-statistics and column 7 the associated p-value from a unit-root test based on Harris and Tzavalis (1999) where there are sufficient data to do so. Column 8 reports the fixed-N critical value at the 5% level and column 9 the fixed-N test statistic from Im, Pesaran and Shin (2003), allowing for linear time trends. These tests are appropriate for panel data settings with short timeframes. The results of the tests provide significant evidence against the existence of a unit-root process, and therefore a random walk, for the tested series. These results provide support to the notion that the benchmarking data are consistent over time.

For each statistic we study, we subtract the corresponding benchmark data sourced from survey or administrative data from the claim of the respondent. For questions related to the rules of the service, which are uniform across all respondents regardless of sector and tier, the claim is subtracted from the benchmark data and the absolute value generated. For the demographic and sector claims, the format of these differ slightly by tier. District officials were asked directly about their district, and the relevant benchmark indicators subtracted from these responses. The regional officials were asked about the 25th, 50th and 75th percentile districts of individual variables

separately. The federal officials were asked, “Which region, out of the following options, do you work on most/think that you know best/feel most comfortable answering questions about?” They were then asked about the 25th, 50th and 75th percentile districts of individual variables within that region. This ensures that each official is responding to questions on the districts that they are most familiar with, and work most closely on. Federal officials across sectors chose the full distribution of regions, and the distribution of choices closely matches population and district numbers.

The raw errors vary significantly, as documented in the main text. The distribution of the errors suggests that civil servants, on average, tend to overestimate the value of the indicator, with the average error of all but three of the indicators we study being positive (the three indicators with a negative actual error are: the proportion of rural inhabitants; the primary pupil-class ratio; and the hectares of land used for pastoral purposes). Since it is difficult to compare errors across different indicators, the overall skewness coefficient for the z-score of the actual error is 0.4. A test of skewness relative to the normal distribution is rejected at the usual levels. Individually, the indicators tend to be positively skewed (have distributions with longer right tails). Six of the 20 indicators have a negative skewness coefficient: proportion of rural inhabitants (-3.02); the antenatal care rate (-0.49); the rate of births delivered by a skilled attendant (-0.22); the share of households living on subsistence agriculture (-0.73); the proportion of income from regional block grants (0.64); and the proportion of income from own sources (-0.56).

In the core analysis, we aim to compare the claims of officials at multiple tiers of government relating to the same district. We therefore focus on districts for which we have subjective assessments from each tier of government. This requires a district to be one of the 25th/50th/75th percentiles for at least one of the indicators we study, since these are the districts over which federal and regional officials made assessments. We therefore drop roughly a third (24) of the 66 districts on which we collected data to ensure that each claim has counterparts at all three tiers of government.

We create a data set where the unit of observation is a claim of a bureaucrat about a district. As described, each federal and regional official is asked about three districts for each indicator (the 25th/50th/75th percentiles of that indicator), and so each bureaucrat-indicator-percentile is an observation for these officials. The claim of the district official about a specific indicator in her district is a single observation. Each response is associated with the corresponding district census

ID.

In a very small number of cases, there seems to have been a data entry issue, where additional digits were included or excluded in the recording of responses to non-fractional variables. This leads us to some errors that are off by a very significant factor (of 94 at the 95th percentile of errors). These are not reflective of the wider variation in the data. Two-thirds of absolute errors are less than 1 in proportional terms. The most extreme values seem outside the realm of feasibility for even the least informed bureaucrat. Since a plausible explanation is that these outliers reflect data entry issues rather than outsized errors, we therefore censor the top 5% of raw errors for each of the variables. The interpretation of the paper is therefore analysis of the lowest 95% of errors that public officials make, rather than of the top 5%.

Absolute values are then taken of all errors. At this stage, we have a data set that is not comparable across, but only within indicators. Making an error of 0.5 in the percentage of rural inhabitants is large for that indicator, but tiny for the population indicator. We therefore undertake a series of transformations of the absolute errors to make them comparable across indicators. We take three approaches.

Our core measure is created by constructing z-scores that are the value of the absolute error minus the mean value across the distribution of absolute errors within the indicator, all divided by the standard deviation of absolute errors across the distribution of absolute errors within the indicator. The interpretation of this z-score is a unitless measure of how many standard deviations away from the benchmark data an official's assessment is. Comparisons across indicators therefore take into account the empirical distribution of errors for each indicator. They compare the scale of an official's error relative to the errors made by all other officials in our data set.

One issue with the above measure is that it does not take into account potential skewness of the underlying indicators. Table A2 provides information on the Pearson moment coefficient of skewness for each of the indicators. It is evident that the non-fraction indicators exhibit right-skew, with positive coefficients larger than 1 in all cases. For non-fraction indicators, a test for skewness relative to the normal distribution is conducted based on D'Agostino et al. (1990) and Royston (1992). The test statistic is approximately normal under the null hypothesis and allows for standard hypothesis tests. Thus, if the p-value associated with the test statistic is less than or equal to 0.05, we classified the variable as significantly skewed and replace the indicator with its log, consistent with Duggan and Morton (2004). For those that are not significantly skewed, no

transformation is applied. The ‘skewness-adjusted’ z-score is calculated on this set of values.

A third approach is to use the ordinal rank of the absolute error within the entire distribution of absolute errors within the indicator (again consistent with Duggan and Morton, 2004). We rank all errors according to size, give each a corresponding rank, and divide these ranks by the maximum rank within the distribution of absolute errors within the indicator. The rank-proportion scores are therefore free of distributional assumptions. They can be interpreted as the percentile of error made by an official relative to other officials in our data set.

Measuring the incentive environment

To construct our management indices, we followed the best-practices laid out in the World Management Survey literature (such as in Bloom and Van Reenen, 2007; Bloom et al., 2012). Management interviews were ‘double blind’ so: (i) managers were not told in advance they were being scored or shown a score grid; (ii) enumerators were given no information on the organization’s performance. Only managers (directors) were enumerated the World Management Survey and the measures of organizational management practices we use for our core analysis average management scores over the most senior divisional-bureaucrat reports where there are multiple directors.

The scores on each question are converted into normalized z-scores (so are continuous variables with mean zero and variance one by construction) and are increasing in the commonly understood notion of ‘better management’. For example, we assume greater monitoring corresponds to better management practice. We then take unweighted means of the underlying z-scores to aggregate the questions on monitoring shown in Table A3 and separately all other management practices.

?????We complement our measure of alignment with a second measure by asking civil servants directly about the extent of alignment they feel to the organization and their daily activities. In the civil servants survey, we asked officials, “To what extent do you believe that the organization’s mission is aligned to your own mission?” and “How consistent with your mission are the various tasks and activities assigned to you on a day-to-day basis?” To measure the extent to which the organization as a whole and the individual tasks assigned to the civil servant are consistent with her own mission, we create an alignment index that is the mean of the organization average response to the former of these questions and the proportion of employees in the organization that state “I am contributing to fulfilling that mission on an everyday basis” to the latter.

Basic Determinants of Bureaucrat Errors

What are the descriptive determinants of these errors? Table A8 undertakes complementary ANOVA and OLS analyses where the dependent variable is the z-score of errors across variables outlined above. We focus on variables related to the hierarchical structure of the bureaucracy, and the marginal costs and benefits of information acquisition. Column 1 of Table A8 presents partial sum of squares for these variables with corresponding F-statistic p-values presented in parentheses. Column 2 provides coefficients for the corresponding variables within a formal regression framework for consistency with our later tables.

The first panel tests whether the hierarchical structure of government matters for the quality of officials' information along three margins. We see that tier of government is of substantial importance, and has a higher partial sum of squares than any of the other 'structure' variables. Its p-value is 0.00, which is consistent with the notion that the structure of formal authority is a critical determinant of information absorption. In column 2, where the coefficients are from a regression of the variables listed in Table A8 on the z-score of errors, we see that the district government binary has a p-value of 0.02, previewing future results.

We can also assess these unconditional results graphically. Figure A3 presents the average errors of centralized and district officials in their claims about each of the 42 districts on which we focus our core analysis. The set of claims over which these averages are made are the same as in our regressions such that they are balanced across variables and percentiles. The scatterplot is ranked by the average error of claims made about each district, and we can immediately see that some districts are fundamentally more difficult to make assessments about. Corresponding to the results in Table 2, we see that the unconditional means of claims by district officials are typically below those of officials in centralized organizations across the distribution. In 75% of cases, the district officials' assessments are more accurate than those of the corresponding central government officials.

We also test the extent to which being of managerial rank is associated with the scale of an official's errors, and find that the sum of squares is very low and the associated p-value is 0.83. The corresponding conditional coefficient in column 2 is 0.01, with a p-value of 0.86. We also find little evidence that there are substantial differences in errors across sectors, consistent with the idea that it is the structure of incentives, rather than the topic of focus, that matters for information acquisition.

Though not shown in the table, we have also investigated the relationship between environmental features and the scale of errors made. These are lower in districts with a lower poverty rate, smaller population, and those that are less remote. However, these features explain only 1.3% of the variation we observe, and once again highlight the primacy of incentives in determining the quality of information a public official has. As indicated in Figure 3, the mean error varies substantially across claimed about districts. Therefore, there must be unobserved aspects of individual districts that make information costlier to acquire.

It seems natural to hypothesize that agents would have a greater information advantage under de facto agent-authority in environments where information is costlier to acquire. However, increases in costs to both agents typically have an ambiguous effect on the relative distribution of information. For example, differences in the shape of the cost function at the specific levels of investment will play a role in determining the effect on the relative distribution. It is therefore an empirical question as to whether the differential between principal and agent increases as the contracting environment becomes more difficult. To assess this in our setting, Table A9 presents the results of a quantile regression of our preferred specification over the quantiles of absolute error. The coefficient on the district dummy is negative and significant at the 1% level across all percentiles of absolute error, implying that district officials' claims are closer to the benchmark data than centralized officials' claims at all levels of absolute error. There is no clear pattern across the percentiles in terms of the size of coefficients, and 6 of the 9 are between -0.20 and -0.27. Thus, the relationship between formal authority and information in this setting is not mediated by the scale of the error to which it relates, but rather formal authority has a broadly positive relationship with information acquisition.

The marginal costs of information acquisition may be determined by features of the variables we test officials on, the characteristics of the responding official, their organization's infrastructure for disseminating information, and features of the local environment. The second panel of Table A8 outlines a corresponding series of variables that increase or decrease the marginal costs of information acquisition. First we focus on features of the claims, and see that the specific indicator on which officials make claims explains a significant portion of the variation in overall error rates, as one might expect. The percentile of an indicator (whether it is the 25th, 50th or 75th percentile district in the data) is similarly associated with the scale of error, with the average being a more familiar quantity to officials than other points in the distribution. These results imply

the importance of controlling for indicator features in our core regressions.

Being a more experienced or educated bureaucrat may lower the cost of information acquisition. However, the coefficients in column 1 corresponding to bureaucratic characteristics indicate that individual demographics have little relationship with the officials' overall likelihood of making errors. The corresponding regression coefficients are all insignificant at the usual levels. The sum of squares is higher for organizational features, consistent with theory. For example, Aghion and Tirole (1997) argues that "it is always optimal for the firm to be in a situation of overload" where managers credibly commit to lower monitoring of employees through increased span of control. A higher span of control therefore leads to lower errors on average among officials, as employees increase their investment in information. We see that the corresponding coefficient on span of control in column 2 of Table A8 is negative, implying that operating under a higher span of control leads to officials making lower errors in their claims.

To ensure consistency with the main analysis, we include district fixed effects in Table A8. To explore the notion that simple proximity to local constituents (such as being an official at a local-level organization) is driving differences in information quality, we can drop these fixed effects and include measures of 'remoteness' of the district: the average travel time of constituents to the nearest urban area, the average elevation of the district, the average slope of the district, and the road density of the district. The partial sum of squares of all of these variables combined is 2.76. While non-trivial (and accounted for in our main specifications with district government fixed effects), these variables combined account for less variation in errors than variables capturing management practices or decentralization. In the OLS regression similar to column 2 of table A8, removing district fixed effects and including these variables, none are statistically significant at usual levels and a joint Wald test of their significance gives a p-value of 0.546. These results are consistent with the argument put forward in Cremer et al. (1994) that simple proximity is unlikely to be a key driver in differences in information across tiers of government.

The final panel of Table A8 assesses factors related to the marginal benefits of information acquisition. We split our measures of management practices into those related to monitoring and all others. Management practices related to monitoring explain the largest portion of the variance in errors. The coefficient on monitoring-related management practices in column 2 of Table A8 is negative, as we would expect, and significant at the 5% level.

Alignment between the mission of an organization and its employees is a feature of the principal-

agent relationship frequently discussed in the literature. Table A8 includes our complementary measures of alignment. The measure based on the management practices approach explains substantially more variation than that based on direct questioning. However, the coefficients on neither variable are significant at the usual levels in column 2 with p-values of 0.138 and 0.431 respectively.

Robustness of Core Results

To assess the robustness of the core descriptive results to our methods of measurement, we present a series of robustness checks in Table A10. The first column of Table A10 re-runs our preferred regression, but disaggregating the central tier of government binary into the federal and regional tiers, with region as the excluded category. The coefficient on district government is -0.31 with a p-value of 0.00, equivalent to that in our preferred regression, and the federal government binary has a value of 0.04 and is insignificant at the usual levels, implying that there is no statistical difference between the errors made by federal and regional officials. As described above, the federal and regional organization structure is equivalent and functionally distinct to the district. In columns 2 to 5 of Table A10, we return to a binary district variable and vary the level at which we cluster the standard errors in our preferred regression. Taking the conservative perspective that information regarding a specific district may be freely shared among officials across government, we cluster at the ‘claimed about’ district in column 2. The district coefficient continues to be significant at the 1% level. Similarly, we may be concerned that errors are correlated across the respondent’s claims, within their organization, or across the region they work in through their regional sector offices (who provide them with regulations and technical assistance). We therefore cluster at the respondent level, the organization-of-respondent level, and the region-sector level in columns 3, 4 and 5 of Table A10 respectively. In each case, the coefficient on the district government binary is significant at the 1% level. We can also rerun our preferred specification using proportional rank as an alternative measure of the magnitude of errors. Column 6 of Table A10 presents the result, which implies a district official makes an error that is 10 percentage points lower than a central bureaucrat. Returning to comparison with the poverty rate, this magnitude is equivalent to 37 times the magnitude of the coefficient on the poverty rate using this dependent variable, and thus akin to making claims over the 25th percentile of poor districts rather than the 62nd. As outlined in section 2.4, we may be concerned about using absolute errors in the presence of skewness of the

underlying indicators. We therefore re-run our preferred specification on the ‘skewness-adjusted z-score’ and find that the coefficient on district government is even larger in magnitude with a coefficient of -0.39, and a p-value of 0.00. Distinct measurement approaches produce qualitatively similar results in this setting.

Table 1: Characteristics of Civil Servants

Means and standard deviations

	(1) All	(2) Federal	(3) Regional	(4) Local
Organizational Characteristics				
Number of organizations	382	5	54	323
Number of bureaucrats per organization	5.7	57	8.7	4.4
Number of managers per organization	1.2	6.8	2.6	0.9
Number of employees per organization	3.6	49	5.4	2.6
Span of control (employees per manager)	7.8	28	16	6.0
Official's Characteristics				
Number of bureaucrats	1831	280	431	1120
Age	35.3	35.87	38.1	34.07
	[8.85]	[9.1]	[8.74]	[8.58]
Years in position	2.7	2.29	2.98	2.69
	[2.74]	[2.88]	[2.87]	[2.64]
Years in organization	7.34	5.8	7.97	7.48
	[7.13]	[6.47]	[7.27]	[7.18]
Years in civil service	13.13	12.87	15.49	12.29
	[8.88]	[8.98]	[9]	[8.65]
Number of service organizations worked in	1.7	2.1	2.19	1.41
	[2.07]	[2.22]	[2.46]	[1.81]
Grade	7.93	7.53	10.89	6.9
	[13.06]	[9.92]	[17.97]	[11.21]
Education [undergraduate degree=1]	0.82	0.96	0.9	0.76
	[0.38]	[0.19]	[0.3]	[0.43]
Education [masters degree=1]	0.11	0.35	0.22	0.01
	[0.32]	[0.48]	[0.42]	[0.12]
Gender [female=1]	0.2	0.27	0.23	0.18
	[0.4]	[0.44]	[0.42]	[0.38]
Manager	0.25	0.12	0.32	0.26
	[0.44]	[0.33]	[0.47]	[0.44]

Notes: Standard deviations are in parentheses. The unit of observation for the first panel is the civil servant and for the second panel the organization. Only employees and managers responded to the information module and hence only these individuals are used for the statistics in the first panel. These statistics are based on the survey sample and not on administrative data. The span of control is taken from the survey question "How many personnel do you manage?"; the average is taken for an organization if there is more than one manager; and the organization-level statistics are reported (each organization has an equal weight in the 'By Organization' statistics). 'Manager' refers to managers within the organization. Grade is the official civil service grade of the professional civil servant, ranging from 1 to 17 in the sample. The 2013-14 National Civil Service Human Resource Statistics Abstract (Ministry of Civil Service) records the overall female employee percentage to be 35%; this statistic includes frontline staff. Figures are rounded to two significant figures.

Table 2: Hierarchy and Information

Dependent Variable: A binary indicator of whether respondent doesn't know in columns (1) and (2); Z-score of errors across all policy domains in columns (3) to (10)

OLS Estimates

Standard Errors: Clustered at the district government-sector level

	(1) Don't know Unconditional	(2) Don't know Conditional	(3) Unconditional Decentralized	(4) Unconditional Manager	(5) Unconditional Hierarchies	(6) Sector FEs	(7) District government FEs	(8) Indicator FEs	(9) Percentile FEs	(10) Individual Characteristics
District Government [Yes=1]	-0.17	-0.17	-0.27		-0.27	-0.28	-0.30	-0.31	-0.31	-0.33
	[0.034]	[0.033]	[0.084]		[0.084]	[0.084]	[0.086]	[0.087]	[0.087]	[0.087]
Manager [Yes=1]	0.011	0.0040		0.011	0.023	0.020	0.014	0.016	0.015	0.034
	[0.031]	[0.029]		[0.066]	[0.066]	[0.067]	[0.065]	[0.066]	[0.066]	[0.069]
Sector fixed effects	No	Yes	No	No	No	Yes	Yes	Yes	Yes	Yes
District government fixed effects	No	Yes	No	No	No	No	Yes	Yes	Yes	Yes
Indicator fixed effects	No	Yes	No	No	No	No	No	Yes	Yes	Yes
Percentile fixed effects	No	Yes	No	No	No	No	No	No	Yes	Yes
Individual controls	No	Yes	No	No	No	No	No	No	No	Yes
Adjusted R-squared	0.022	0.13	0.012	-0.00066	0.012	0.0099	0.059	0.067	0.072	0.069
Observations [clusters]	2144 [98]	2144 [98]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]
Central government respondents [percentage]	594 [67%]	594 [67%]	429 [62%]	429 [62%]	429 [62%]	429 [62%]	429 [62%]	429 [62%]	429 [62%]	429 [62%]
District government respondents [percentage]	297 [33%]	297 [33%]	258 [38%]	258 [38%]	258 [38%]	258 [38%]	258 [38%]	258 [38%]	258 [38%]	258 [38%]

Notes: Clustered standard errors are in parentheses. Standard errors clustered at the district government-sector interacted level. All columns report OLS estimates. The unit of analysis in all columns is the claim. The sample used for the analysis is the set of respondents that make claims about the same districts. The dependent variable is a binary indicator for whether the respondent could not provide a claim in columns (1) and (2); the z-score of absolute error across all policy domain items in columns (3) to (12). Sector fixed effects are indicators for whether the claim is within an agriculture, education, health, revenue, or trade organization. Percentile fixed effects are indicators for the 75th, 50th or 25th percentile for the respective indicator for the respective region. District government fixed effects identify the district that the respondent is making claims about. Indicator fixed effects control for the service delivery indicator within which the claim is made. Manager is equal to one if the claim is from a director. Individual characteristics included in the specification: tenure in the civil service of the respondent making the claim; tenure in the position; tenure in the organization; an indicator for whether the respondent is female; an indicator for whether the respondent holds an undergraduate degree; an indicator for whether the respondent holds a masters degree. Figures are rounded to two significant figures.

Table 3: Information, Incentives and Alignment

Dependent Variable: Z-score of errors across all policy domains in all columns

OLS Estimates

Standard Errors: Clustered at the district government-sector level

	Incentives				
	(1) Management practices	(2) Noise controls	(3) Combined	(4) Interaction Decentralized	(5) Interaction Manager
District Government [Yes=1]			-0.27 [0.16]	-0.34 [0.16]	-0.27 [0.16]
Manager [Yes=1]	0.011 [0.070]	-0.026 [0.076]	-0.029 [0.076]	-0.021 [0.077]	-0.035 [0.076]
Management Practice: Monitoring	-0.17 [0.067]	-0.21 [0.067]	-0.20 [0.067]	-0.14 [0.080]	-0.17 [0.067]
Management Practice: Other	0.28 [0.097]	0.31 [0.11]	0.32 [0.11]	0.32 [0.10]	0.31 [0.11]
District Government x Management Practice: Monitoring				-0.19 [0.14]	
Manager x Management Practice: Monitoring					-0.10 [0.090]
Sector, indicator, percentile fixed effects	Yes	Yes	Yes	Yes	Yes
District government fixed effects	No	No	No	No	No
Region fixed effects	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes
Capital control variables and noise controls	No	Yes	Yes	Yes	Yes
Adjusted R-squared	0.034	0.061	0.063	0.064	0.063
Observations [clusters]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]
Central government respondents [percentage]	429 [62%]	429 [62%]	429 [62%]	429 [62%]	429 [62%]
District government respondents [percentage]	258 [38%]	258 [38%]	258 [38%]	258 [38%]	258 [38%]

in the first column and standard errors clustered at the district government-sector level. The dependent variable is the z-score of the organization average response of employees to "How many people would you say regularly give you tasks as part of your formal work duties?"; alignment (the z-score of the row mean of the organization-average response to "To what extent do you believe that your organization's mission is aligned to your own mission?" and the organization proportion of selecting "I am contributing to fulfilling that mission on an everyday basis" to "How consistent with your mission are the various tasks and activities assigned to you on a day-to-day basis?"). Noise controls include the time of day of the survey, day-of-survey fixed effects, enumerator fixed effects, an indicator for the enumerator's subjective assessment of the quality of the interview; an indicator for the duration of the interview (decile of duration fixed effects); controls for the average tenure of the managers in the organization and the proportion of female managers in the organization. Figures are rounded to two significant figures.

Table 4: Information Experiment

Dependent Variable: Z-score of errors across policy domains in columns (1) to (4) and (7); rank of error (proportion) in column (5); and skewness-adjusted z-score in column (6)

OLS estimates in all columns

Standard Errors: Clustered at the district government-sector level in columns (1) to (6) and at the region-sector level in column (7)

	(1) Treatment only	(2) Conditional Treatment Effect	(3) Baseline with treatment	(4) Management practices	(5) Rank Error (Proportion)	(6) Skewness-Adjusted Z-Score Error	(7) Clustering at Region-Sector Level
Treated District [Yes=1]	0.053	-0.15	-0.28	-0.23	-0.074	-0.2	-0.23
	[0.087]	[0.085]	[0.098]	[0.11]	[0.032]	[0.11]	[0.11]
<i>P-value of coefficient clustered at district government-sector level</i>	{0.54}	{0.078}	{0.0048}	{0.036}	{0.022}	{0.086}	{0.048}
<i>P-value of coefficient clustered at district level</i>	(0.51)	(0.057)	(0.0031)	(0.033)	(0.024)	(0.053)	(0.033)
District Government [Yes=1]			-0.44	-0.41	-0.11	-0.39	-0.41
			[0.17]	[0.17]	[0.048]	[0.17]	[0.16]
Manager [Yes=1]			-0.021	-0.026	-0.011	-0.0074	-0.026
			[0.077]	[0.076]	[0.020]	[0.066]	[0.080]
Management Practice: Monitoring				-0.18	-0.018	-0.027	-0.18
				[0.069]	[0.024]	[0.088]	[0.064]
Management Practice: Other				0.28	0.038	0.090	0.28
				[0.11]	[0.038]	[0.13]	[0.14]
Sector, region, percentile, indicator fixed effects	No				Yes		
Individual controls	No				Yes		
Capital control variables and noise controls	No				Yes		
Adjusted R-squared	-0.00018	0.057	0.062	0.066	0.092	0.10	0.066
Number of observations [clusters]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [38]

Notes: All columns report OLS regression coefficients and standard errors in brackets. Standard errors clustered at the district government-sector interacted level in columns (1) to (6) and at the region-sector interacted level in column (7). P-values are reported in curly brackets. P-values associated with clustering at the district level are reported in round brackets. The unit of analysis in all columns is the claim. The sample used for the analysis is the set of respondents that make claims about the same districts. The dependent variable is the z-score error for any claim made within any policy domain in columns (1) to (4) and column (7). The dependent variable is the rank of the absolute error in terms of the proportion of the total number of claims for the indicator in column (5). The dependent variable is the skewness-adjusted z-score in column (6). The tiers of government are grouped into District Government and Centralized Government which incorporates both Federal and Regional tiers, hence the omitted category for tier in the regressions is Centralized Government. Treated district is an indicator equal to one if the district received an information package. Sector fixed effects are indicators for whether the claim is within an agriculture, education, health, revenue, or trade organization. Percentile fixed effects are indicators for the 75th, 50th or 25th percentile for the respective indicator for the respective region. Indicator fixed effects are indicator variables for which specific variable the claim is being made over. Management practice: Monitoring is the organization-average z-score for the monitoring practice of management as per Bloom and Van Reenen (2007); Management practice: Other is the mean of the organization average z-scores for targeting, incentives, roles, flexibility, staffing, and staff involvement. Manager is equal to one if the claim is from a manager. Individual characteristics included in the specification: tenure in the civil service of the respondent making the claim; tenure in the position; tenure in the organization; an indicator for whether the respondent is female; an indicator for whether the respondent holds an undergraduate degree; an indicator for whether the respondent holds a masters degree. Capital control variables include the proportion of managers in the organization that state that a management information system is in place; the organization average response to "In what proportion of projects would you say information flows effectively between individuals in the same organization?"; the organization average response to "In what proportion of projects would you say information flows effectively between organizations?"; the organization average response of managers to "How many personnel do you manage?"; the organization average response of employees to "How many people would you say regularly give you tasks as part of your formal work duties?"; alignment (the z-score of the row mean of the organization-average response to "To what extent do you believe that your organization's mission is aligned to your own mission?" and the organization proportion of selecting "I am contributing to fulfilling that mission on an everyday basis" to "How consistent with your mission are the various tasks and activities assigned to you on a day-to-day basis?"). Noise controls include the time of day of the survey, day-of-survey fixed effects, enumerator fixed effects, an indicator for the enumerator's subjective assessment of the quality of the interview; an indicator for the duration of the interview (decile of duration fixed effects); controls for the average tenure of the managers in the organization and the proportion of female managers in the organization. Figures are rounded to two significant figures.

Table 5: Heterogeneous Treatment Effects

Dependent Variable: Z-score of errors across policy domains

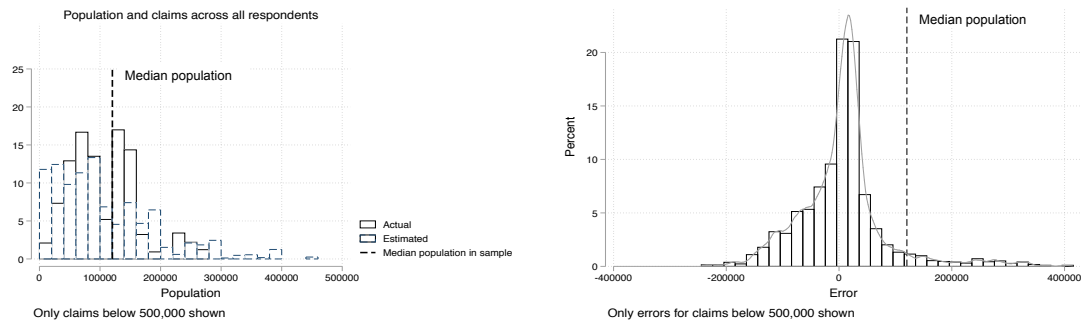
OLS estimates in all columns

Standard Errors: Clustered at the district government-sector level

	(0) Baseline treatment effect	(1) Interaction: district government	(2) Interaction: manager	(3) Interaction: monitoring	(4) Interaction: other management practices
Treated District [Yes=1]	-0.23 [0.11]	-0.025 [0.19]	-0.20 [0.12]	-0.26 [0.10]	-0.34 [0.11]
District Government [Yes=1]	-0.41 [0.17]	-0.089 [0.28]	-0.41 [0.17]	-0.46 [0.16]	-0.43 [0.16]
Manager	-0.026 [0.076]	-0.032 [0.075]	0.068 [0.13]	-0.018 [0.076]	-0.015 [0.076]
Management Practice: Monitoring	-0.18 [0.069]	-0.18 [0.073]	-0.18 [0.069]	-0.43 [0.11]	-0.18 [0.065]
Management Practice: Other	0.28 [0.11]	0.30 [0.12]	0.28 [0.11]	0.30 [0.11]	0.023 [0.14]
Treated District x District Government [Yes=1]		-0.37 [0.25]			
Treated District x Manager			-0.12 [0.13]		
Treated District x Management Practice: Monitoring				0.34 [0.13]	
Treated District x Management Practice: Other					0.38 [0.14]
Sector, region, percentile, indicator fixed effects			Yes		
Individual controls			Yes		
Capital control variables and noise controls			Yes		
Adjusted R-squared	0.066	0.067	0.065	0.071	0.071
Number of observations [clusters]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]

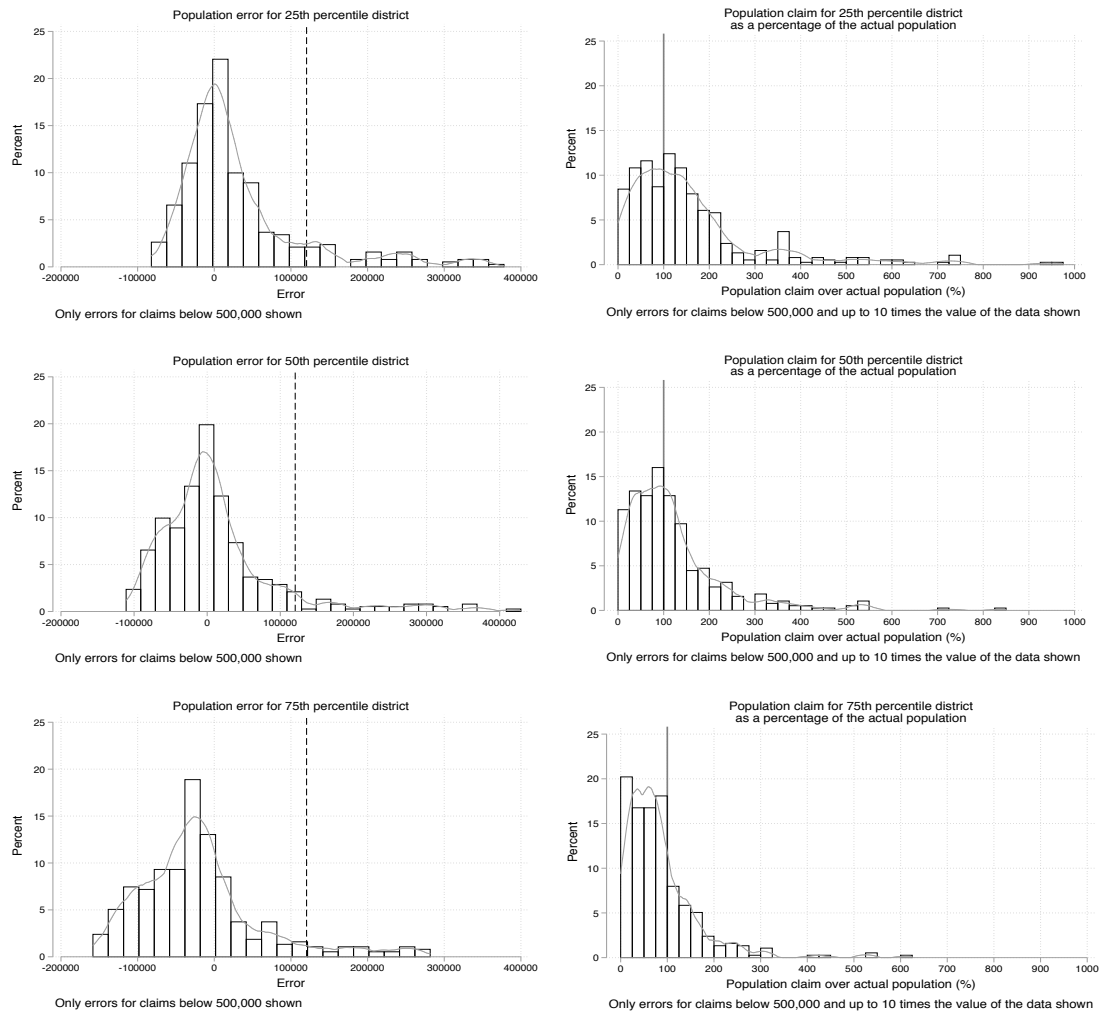
Notes: All columns report OLS regression coefficients and standard errors in brackets. Standard errors clustered at the district government-sector interacted level. The unit of analysis in all columns is the claim. The sample used for the analysis is the set of respondents that make claims about the same districts. The dependent variable is the z-score error for any claim made within any policy domain. The tiers of government are grouped into District Government and Centralized Government which incorporates both Federal and Regional tiers, hence the omitted category for tier in the regressions is Centralized Government. Treated district is an indicator equal to one if the district received an information package. Sector fixed effects are indicators for whether the claim is within an agriculture, education, health, revenue, or trade organization. Percentile fixed effects are indicators for the 75th, 50th or 25th percentile for the respective indicator for the respective region. Indicator fixed effects are indicator variables for which specific variable the claim is being made over. Management practice: Monitoring is the organization-average z-score for the monitoring practice of management as per Bloom and Van Reenen (2007); Management practice: Other is the mean of the organization average z-scores for targeting, incentives, roles, flexibility, staffing, and staff involvement. Manager is equal to one if the claim is from a manager. Individual characteristics included in the specification: tenure in the civil service of the respondent making the claim; tenure in the position; tenure in the organization; an indicator for whether the respondent is female; an indicator for whether the respondent holds an undergraduate degree; an indicator for whether the respondent holds a masters degree; an indicator if the respondent is a manager. Capital control variables include the proportion of managers in the organization that state that a management information system is in place; the organization average response to "In what proportion of projects would you say information flows effectively between individuals in the same organization?"; the organization average response to "In what proportion of projects would you say information flows effectively between organizations?"; the organization average response of managers to "How many personnel do you manage?"; the organization average response of employees to "How many people would you say regularly give you tasks as part of your formal work duties?"; alignment (the z-score of the row mean of the organization-average response to "To what extent do you believe that your organization's mission is aligned to your own mission?" and the organization proportion of selecting "I am contributing to fulfilling that mission on an everyday basis" to "How consistent with your mission are the various tasks and activities assigned to you on a day-to-day basis?"). Noise controls include the time of day of the survey, day-of-survey fixed effects, enumerator fixed effects, an indicator for the enumerator's subjective assessment of the quality of the interview; an indicator for the duration of the interview (decile of duration fixed effects); controls for the average tenure of the managers in the organization and the proportion of female managers in the organization. Figures are rounded to two significant figures.

Figure 1A: All Respondents' Population Estimates and Actual Population



Notes: Figure 1A shows the histograms of actual populations (black) and population claims (blue dashed) for all claims on the left-hand panel; the right-hand panel show histograms of actual errors for all claims. Actual errors are the claim of the respondent minus the actual population value. The black dashed line represents the median actual population in the sample of districts. All claims refers to claims over the 75th percentile, 50th percentile, and 25th percentile district in terms of population within the (chosen) region for the federal and regional respondents; and claims over own districts for district level respondents. The sample includes all employees and managers in the Ethiopian Civil Servants Survey that were asked the information module. In both panels, observations corresponding to population claims over 500,000 are not included for presentation purposes.

Figure 1B: Federal and Regional Respondent Population Estimates and Actual Population



Notes: Figure 1B left-hand panels show histograms of actual errors for the 25th percentile district, 50th percentile district, and 75th percentile district. Actual errors are the claim of the respondent minus the actual population value. The black dashed line represents the median actual population in the sample of districts. The sample is all federal and regional employees and managers in the Ethiopian Civil Servants Survey that were asked the information module. The right-hand panel shows histograms of the population claim divided by the actual population of the district multiplied by 100 to convey the size of the population claim relative to the data in percentage terms. The solid gray line at 100 represents an accurate claim. In both panels, observations corresponding to population claims over 500,000 are not included for presentation purposes. In the right-hand panel, observations where the claim is more than 10 times (1000% of) the actual data are not shown.

Figure 2: Distributions of Errors Across Tiers of Government

Figure 2A: CDF of Errors Using Z-Scores

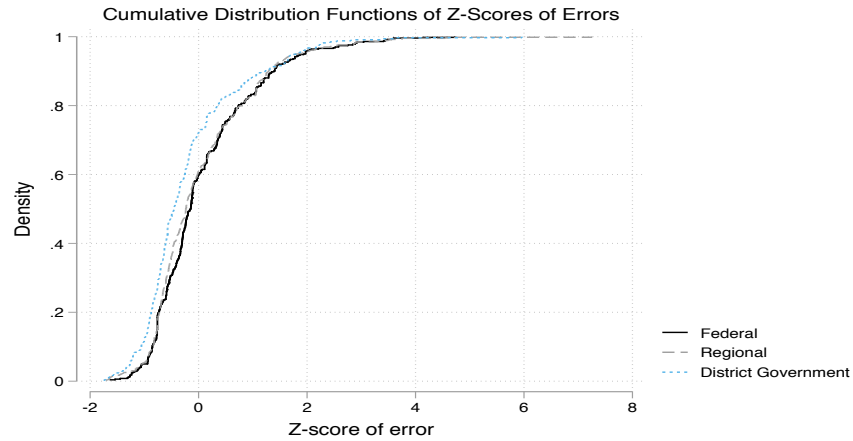


Figure 2B: CDF of Errors Using Skewness Adjusted Z-Scores

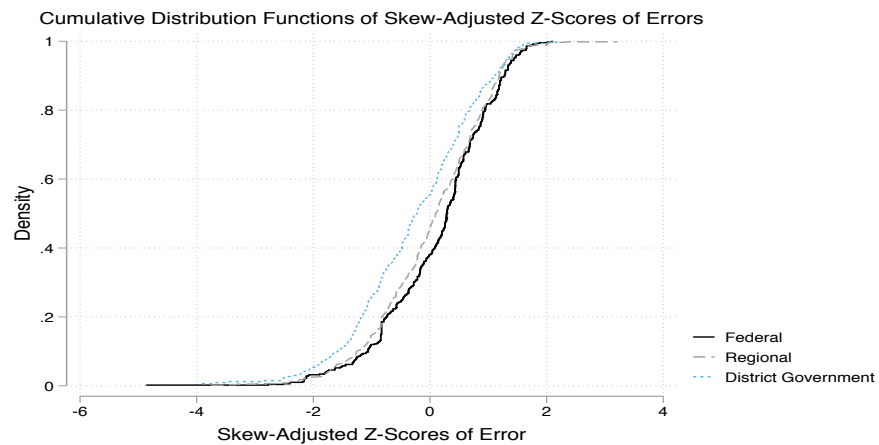
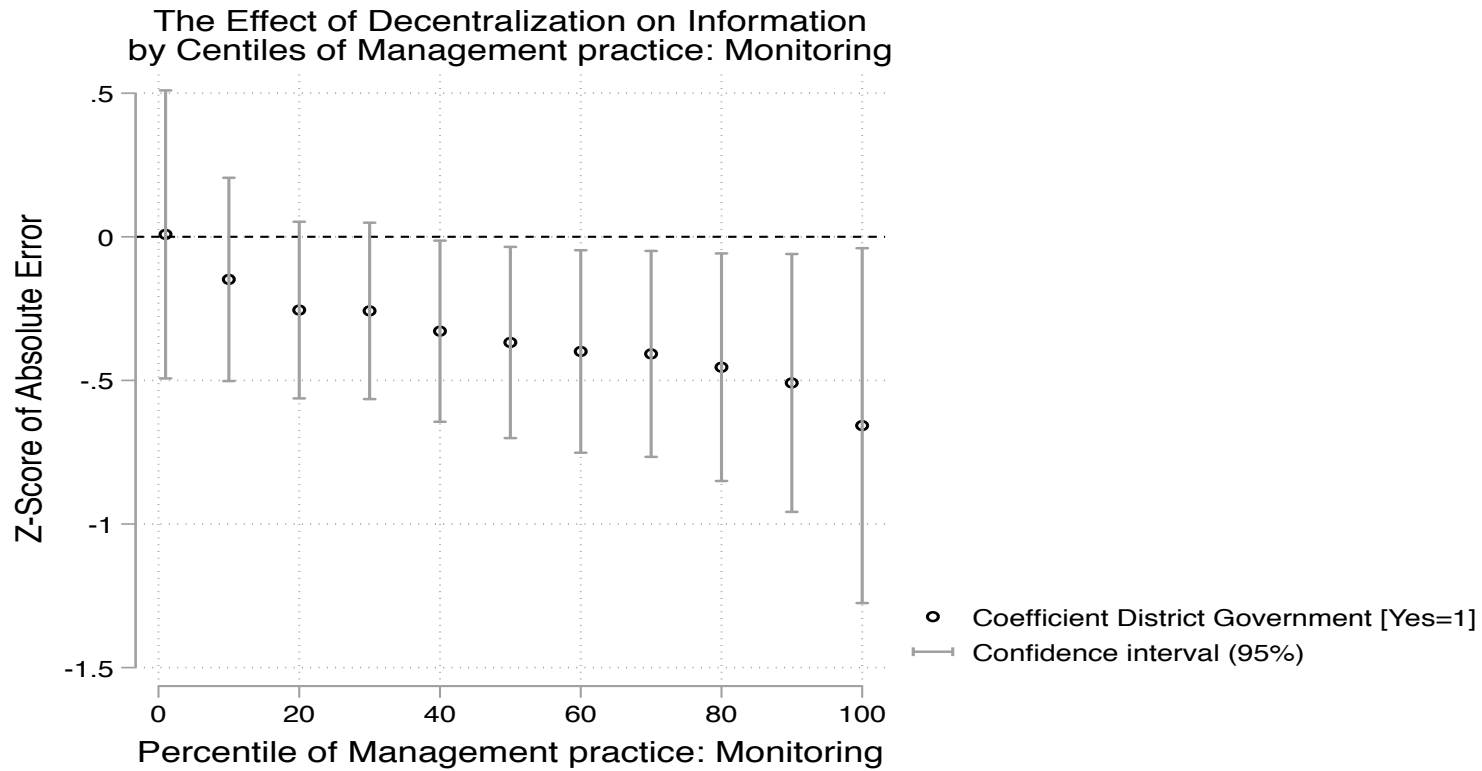


Figure 2C: CDF of Errors Using Ranks



Notes: The figure shows the cumulative distribution functions of the federal, regional, and district government level respondent's errors across all policy domains. The top panel shows the errors in terms of z-score within the indicator. The middle panel shows the error in terms of skewness-adjusted z-score within the indicator, the skewness-adjusted z-score is the standard z-score in cases of variables with a 0 to 1 support and for variables that are not significantly skewed according to the test from D'Agostino, Balanger, and D'Agostino (1990), for those variables that are not 0 to 1 in support and are significantly skewed, the log of the variable is taken and the z-score of this is the skewness-adjusted z-score. The bottom panel shows the errors in terms of rank within the indicator as a proportion of the maximum rank. The sample used in the analysis is the set of respondents making claims about the same district. The black solid line refers to federal claims; the gray dashed line to regional claims; and the blue dotted line to district government claims.

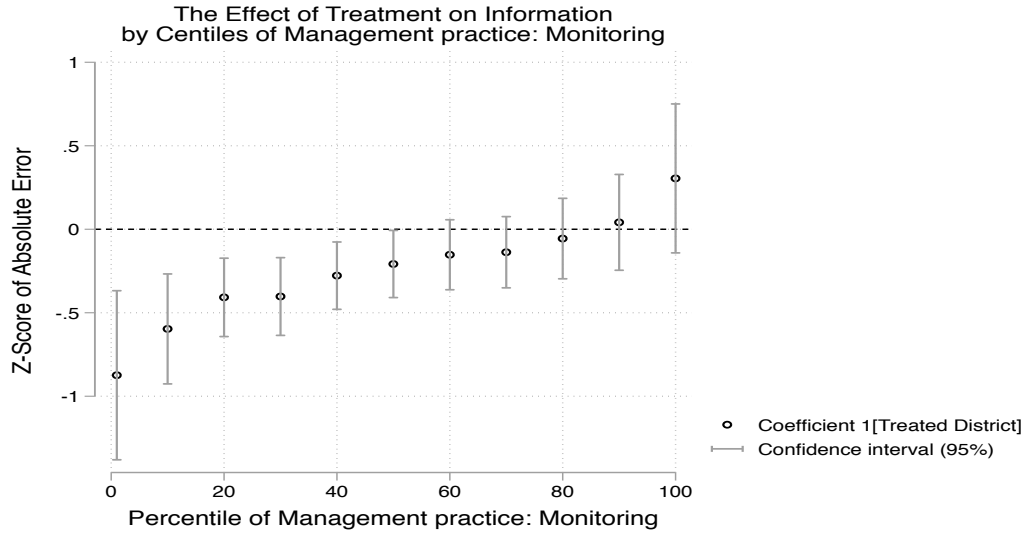
Figure 3: Incentives, Decentralization and Information



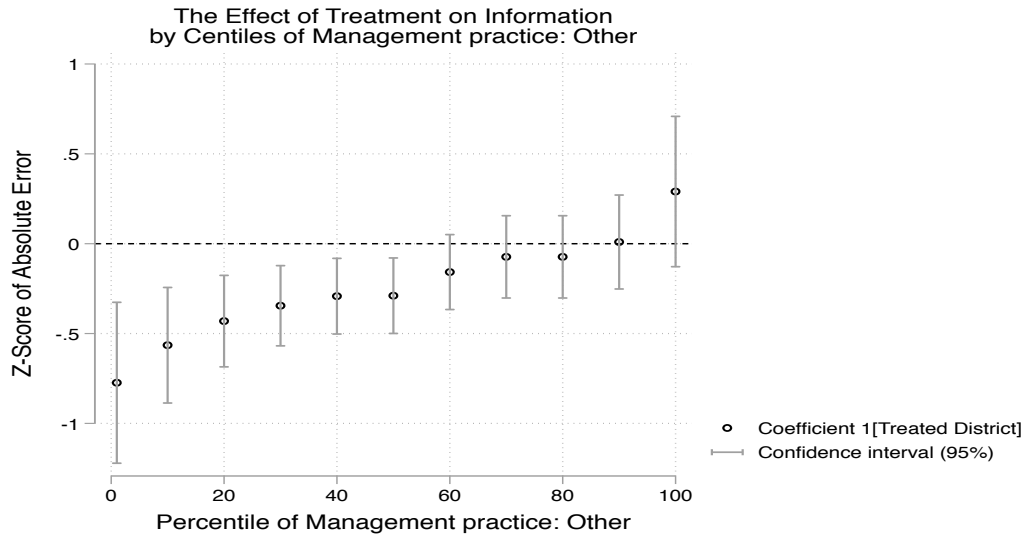
Notes: The figure shows the OLS coefficient on District Government [Yes=1] for each decile of Management practice: Monitoring, in a regression also conditional on Management practice: Other; Individual controls; Sector fixed effects; Indicator fixed effects, Percentile fixed effects, Marginal cost controls and Noise controls. 95% confidence intervals are show in the gray lines, with standard errors clustered at the district government-sector interacted level.

Figure 4: Heterogeneous Treatment Effects by Management practices

4A: Treatment and Management practice: Monitoring



4B: Treatment and Management practice: Other



Notes: The figure shows the effect of the treatment on the z-score of errors across all policy domains by different percentiles of Management practice: Monitoring in the top figure; and by different percentiles of Management practice: Other in the bottom figure. The sample used for the analysis is the set of respondents that make claims about the same districts. Management practices and errors are in z-scores. Treated district is a binary indicator equal to one if the district received an information package. Regressions are conditional on Management practices; Individual controls; Sector fixed effects; Indicator fixed effects. Percentile fixed effects, Marginal cost controls and noise controls. Region fixed effects; and an indicator for whether the claim is from a District Government. 95% confidence intervals are shown in the gray lines, with standard errors clustered at the district government-sector level. Management practice: Monitoring is the organization average of the monitoring topic, constructed from the z-scores of each individual item noted in table A3. Management practice: Other is the organization average of the average across the z-scores of targeting, incentives, roles, flexibility, staffing, and staff involvement, detailed in table A3.

TA1: List of Surveyed Organizations

Tier of Governance	Region	Organization Name
Federal	-	Federal Ministries of Agriculture; Education; Health; Revenue; and, Trade
Regional	Addis Ababa	Addis Ababa City Administration Bureaus of Education; Health; Revenue; and, Trade
Regional	Afar	Afar Regional Bureaus of Agriculture; Education; Health; Revenue; and, Trade
Regional	Amhara	Amhara Regional Bureaus of Agriculture; Education; Health; Revenue; and, Trade
Regional	Benishangul Gumuz	Benishangul Gumuz Regional Bureaus of Agriculture; Education; Health; Revenue; and, Trade
Regional	Dire Dawa	Dire Dawa City Administration Bureaus of Agriculture; Education; Health; Revenue; and, Trade
Regional	Gambella	Gambella Regional Bureaus of Agriculture; Education; Health; Revenue; and, Trade
Regional	Harar	Harar Regional Bureaus of Agriculture; Education; Health; Revenue; and, Trade
Regional	Oromia	Oromia Regional Bureaus of Agriculture; Education; Health; Revenue; and, Trade
Regional	SNNPR	SNNPR Regional Bureaus of Agriculture; Education; Health; Revenue; and, Trade
Regional	Somali	Somali Regional Bureaus of Agriculture; Education; Health; Revenue; and, Trade
Regional	Tigray	Tigray Regional Bureaus of Agriculture; Education; Health; Revenue; and, Trade
District	Afar	Afar Awash Fentale Agriculture, Education, Health, Revenue, and Trade Offices
District	Afar	Afar Telalak Agriculture, Education, Health, Revenue, and Trade Offices
District	Afar	Afar Teru Agriculture, Education, Health, Revenue, and Trade Offices
District	Amhara	Amhara Awabel Agriculture, Education, Health, Revenue, and Trade Offices
District	Amhara	Amhara Basona Worana Agriculture, Education, Health, Revenue, and Trade Offices
District	Amhara	Amhara Borena (Former Debresina) Agriculture, Education, Health, Revenue, and Trade Offices
District	Amhara	Amhara Chefa Gula Agriculture, Education, Health, Revenue, and Trade Offices
District	Amhara	Amhara Dejen Agriculture, Education, Health, Revenue, and Trade Offices
District	Amhara	Amhara Enarj Enawaga Agriculture, Education, Health, Revenue, and Trade Offices
District	Amhara	Amhara Gidane Agriculture, Education, Health, Revenue, and Trade Offices
District	Amhara	Amhara Jabitahnan Agriculture, Education, Health, Revenue, and Trade Offices
District	Amhara	Amhara Jile Timuga Agriculture, Education, Health, Revenue, and Trade Offices
District	Amhara	Amhara Kutaber Agriculture, Education, Health, Revenue, and Trade Offices
District	Amhara	Amhara Simada Agriculture, Education, Health, Revenue, and Trade Offices
District	Benishangul Gumuz	Benishangul Gumuz Dibate Agriculture, Education, Health, Revenue, and Trade Offices
District	Benishangul Gumuz	Benishangul Gumuz Yasso Agriculture, Education, Health, Revenue, and Trade Offices
District	Gambella	Gambella Gambella Zuria Agriculture, Education, Health, Revenue, and Trade Offices
District	Gambella	Gambella Abobo Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Ale Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Amigna Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Arsi Negelle District government Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Babile Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Bako Tibe Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Begi Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Dedessa Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Digluna Tijo Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Gida Ayana (Gida Kiremu) Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Goro Gutu (Goro) Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Guduru Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Haro Maya Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Hitosa Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Jardega Jarte Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Jeldu Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Kofale Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Mesela Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Midaga Tola Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Nono Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Seru Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Siraro Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Tikur Enchini Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Wadera Agriculture, Education, Health, Revenue, and Trade Offices
District	Oromia	Oromia Were Jarso Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Amaro Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Analimo Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Basketo Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Benatsemay Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Bona Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Chere Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Dale Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Decha Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Doyo Gena Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Gomibora Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Hawassa Zuriya Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Kucha Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Shebedino Agriculture, Education, Health, Revenue, and Trade Offices
District	SNNPR	SNNPR Wenago Agriculture, Education, Health, Revenue, and Trade Offices
District	Somali	Somali Afdem Agriculture, Education, Health, Revenue, and Trade Offices
District	Somali	Somali Erer District government Agriculture, Education, Health, Revenue, and Trade Offices
District	Somali	Somali Harshin Agriculture, Education, Health, Revenue, and Trade Offices
District	Somali	Somali Jijiga Zuria District government Agriculture, Education, Health, Revenue, and Trade Offices
District	Somali	Somali Kebri Beyah Agriculture, Education, Health, Revenue, and Trade Offices
District	Tigray	Tigray Erob Agriculture, Education, Health, Revenue, and Trade Offices
District	Tigray	Tigray Gulo Mekeda Agriculture, Education, Health, Revenue, and Trade Offices
District	Tigray	Tigray Hintalo Wajerat Agriculture, Education, Health, Revenue, and Trade Offices
District	Tigray	Tigray Tahtay Koraro Agriculture, Education, Health, Revenue, and Trade Offices
District	Tigray	Tigray Wereilehi Agriculture, Education, Health, Revenue, and Trade Offices

Table A2: Description of Survey and Administrative Data

Summary statistics

	Source	(1) Mean	(2) Standard deviation	(3) Skewness	(4) Autogressive parameter from DF regression, no time trends	(5) Autogressive parameter from DF regression, linear time trends	(6) Z Statistic from Harris Tsavalis Unit Root Test (linear time trends)	(7) P-value from Harris Tsavalis Unit Root Test (linear time trends)	(8) Fixed-N exact critical value for the Im-Pesaran-Shin test at the 5% level (linear time trends)	(9) Fixed-N Im-Pesaran-Shin Test statistic	(10) N
Demographic and Socio-economic Characteristics											
What do you think is recorded as the population of [INSERT WOREDA] according to official administrative data in the Census of 2007 Gregorian Calendar (1999/2000 Ethiopian Calendar)?	Census 2007	99,605	61,868	1.19							740
What do you think is recorded as the percentage of rural inhabitants of [INSERT WOREDA] according to official administrative data in the Census of 2007 Gregorian Calendar (1999/2000 Ethiopian Calendar)?	Census 2007	0.84	0.26	-2.66							726
What do you think is recorded as the unemployment rate of [INSERT WOREDA] according to official administrative data in the Census of 2007 Gregorian Calendar (1999/2000 Ethiopian Calendar)?	Census 2007	4.07	5.63	2.21							740
Education indicators											
What do you think the primary (grades 1-8) enrolment numbers are for [INSERT WOREDA] according to official administrative data?	EMIS 2006/07-2013/14	21,929	14,020	1.06	0.61	0.05	-12.36	0.00	-2.34	-2.63	705
What do you think the primary (grades 1-8) pupil-section ratio is for [INSERT WOREDA] according to official administrative data?	EMIS 2006/07-2013/14	57.17	26.71	5.32	0.41	-0.03					705
What do you think the primary (grades 1-8) pupils-per-school ratio is for [INSERT WOREDA] according to official administrative data?	EMIS 2006/07-2013/14	550	229	1.61	0.57	0.16	-5.46	0.00	-2.34	-2.79	705
What do you think the primary (grades 1-8) pupil-teacher ratio is for [INSERT WOREDA] according to official administrative data?	EMIS 2006/07-2013/14	97.61	354	15.1	-0.15	0.02					467
Health indicators											
What do you think the Proportion of pregnant women who attended ANC4+ during the current pregnancy is for [INSERT WOREDA] according to official administrative data (the baseline of the 2007 Core Plan)?	HSDP Core Plan 2008/09-2013/14	0.51	0.29	0.08	-0.13	-0.53	-6.81	0.00			688
What do you think the Contraceptive Acceptance Rate is for [INSERT WOREDA] according to official administrative data (the baseline of the 2007 Core Plan)?	HSDP Core Plan 2008/09-2013/14	0.59	0.29	-0.23	-0.19	-0.52	-6.38	0.00			688
What do you think the Rate of Births Attended by Skilled Health Personnel is for [INSERT WOREDA] district government according to official administrative data (the baseline of the 2007 Core Plan)?	HSDP Core Plan 2008/09-2013/14	0.33	0.22	2.62	-0.19	-0.60	-8.26	0.00			688
What do you think the Proportion of infants fully immunized is for [INSERT WOREDA] according to official administrative data (the baseline of the 2007 Core Plan)?	HSDP Core Plan 2008/09-2013/14	0.74	0.22	-1.24	-0.08	-0.57	-8.06	0.00			688
Agriculture indicators											
How much land do you think is used for agricultural purposes for [INSERT WOREDA] district government according to data from the WCBS?	WCBS Round III (2009/10) & Round V (2012/13)	29,366	33,892	6.49	1.01	1.01					203
How much land do you think is used for pastoral purposes for [INSERT WOREDA] according to data from the WCBS?	WCBS Round III (2009/10) & Round V (2012/13)	27,009	75,005	5.97	1.00	1.00					195
What do you think is the agricultural income per household per year in [INSERT WOREDA] according to data from the WCBS?	WCBS Round III (2009/10) & Round V (2012/13)	11,738	15,868	4.58	0.32	0.32					146
What do you think is the share of households dependent on subsistence agriculture in [INSERT WOREDA] according to data from the WCBS?	WCBS Round III (2009/10) & Round V (2012/13)	24.29	23.23	1.56	0.92	0.92					132
Revenue indicators											
How many tax identification numbers do you think were issued in the last financial year according to data from the WCBS for [INSERT WOREDA]?	WCBS Round III (2009/10) & Round V (2012/13)	425	478	2.42	0.33	0.33					146
What proportion of income for this district government do you think is received from regional recurrent block grants according to data from the WCBS for [INSERT WOREDA]?	WCBS Round III (2009/10) & Round V (2012/13)	0.81	0.18	-1.97	-0.22	-0.22					190
What proportion of income for this district government do you think is received from own sources according to data from the WCBS for [INSERT WOREDA]?	WCBS Round III (2009/10) & Round V (2012/13)	0.20	0.15	2.46	-0.08	-0.08					177
Trade indicators											
How many business licenses do you think were issued in the last financial year according to data from the WCBS for [INSERT WOREDA]?	WCBS Round III (2009/10) & Round V (2012/13)	376	481	5.19	1.36	1.36					177
What do you think is the total revenue collected from issuing and renewing business licenses according to data from the WCBS for [INSERT WOREDA]?	WCBS Round III (2009/10) & Round V (2012/13)	108,473	218,841	8.69	0.32	0.32					172
Civil Service Rules											
What is the amount of regular working hours for a civil servants according to the Civil Service Proclamation? (Paragraph 32)	Civil Servants Proclamation	39									
What is the allowed amount of annual leave in the first year of service according to the Civil Service Proclamation? (Paragraph 37, Item 1)	Civil Servants Proclamation	20									
What is the maximum amount of annual leave for a civil servant who has served for more than one year according to the Civil Service Proclamation? (Paragraph 37, Item 2)	Civil Servants Proclamation	30									
What is the allowed amount of maternity leave according to the Civil Service Proclamation? (Paragraph 41, Item 2)	Civil Servants Proclamation	90									
What are the types of different disciplinary actions that a civil servant can face for the breach of discipline according to the Civil Service Proclamation? (Paragraph 67)	Civil Servants Proclamation	6									
What is the notice time required before resigning according to the Civil Service Proclamation? (Paragraph 78, Item 1)	Civil Servants Proclamation	30									

Notes: The unit of observation is the district. EMIS is the Education Management Information System of the Ministry of Education, Ethiopia. HSDP is the Health Sector Development Plan of the Ministry of Health, Ethiopia. WCBS is the District government and City Benchmarking Survey, undertaken by the Ministry of Public Service and Human Resource Development and World Bank. Skewness refers to Pearson's moment coefficient of skewness. Column 4 rep teacher ratio are not displayed due to gaps in the panel series; test statistics and autogressive parameters are not shown for the demographic indicators as these are from a single observation from the census of 2007; test statistics and autogressive parameters are not shown for the agriculture, revenue, or trade indicators as these are from a two-period panel dataset, with insufficient observations to conduct the tests. Figures are rounded to two decimal places.

Table A3: World Management Survey Indicators

Aggregate Index	Topic	Question	Score 1	Score 3	Score 5
Monitoring	Monitoring	In what kind of ways does your Directorate track how well it is delivering services? Can you give me an example?	Directorate does not track performance.	Directorate tracks a number of performance indicators. These are seen and reviewed by senior management only.	Full set of indicators are tracked formally and continuously. Reviews are conducted regularly and involve representative of all directorate staff groups. The results of the review are formally communicate to all directorate staff.
Monitoring	Monitoring	Are you involved in performance review for your Directorate? If so, how often does this occur?	Not involved in performance review;	Bi-annually	Monthly
Other	Targeting	Does your Directorate have a clear set of targets derived from the organization's goals and objectives? Are they used to determine your work schedule?	The directorate does not have defined targets.	Targets are assigned to the directorate, as well as to the manager and employee levels, and these are generally well understood by mid-level staff. However the tasks assigned to staff are not always related to those targets.	Targets are clearly defined for the directorate, manager, and employee levels, and are well understood by all staff. All tasks are directly derived from the targets, which are regularly reviewed to ensure they remain on track.
Other	Targeting	When you arrive at work each day, do you and your colleagues know what their individual roles and responsibilities are in achieving the organization's goals?	Staff do not know what their roles and responsibilities are.	Staff have a good idea of their roles and responsibilities but it is not always clear how they contribute to their organization's goals.	Staff have a very good understanding of their roles and responsibilities. Their own roles and goals are clearly interconnected to those of their organization.
Other	Targeting	How are targets and performance measures communicated to staff in your directorate?	Neither targets nor performance measures are communicated to staff.	Targets and performance measures are formally communicated to managers and team leaders.	Targets and performance measures are formally communicated and understood by all staff.
Other	Performance incentives	How would under-performance be tolerated in your Directorate? Can you give me an example of how such a case would be dealt with?	Poor performers stay in their positions (no consequences).	Poor performance is identified through evaluation and is addressed through concrete action. Although this applies to most staff, some individuals/staff groups get away with it.	Poor performers are identified through regular reviews and are put on a formal performance improvement plan immediately. This applies to all staff.
Other	Performance incentives	Given past experience, have members of [respondent's organization] been disciplined for breaking the rules of the civil service?	There are no consequences for bad behaviour/ breaking the rules.	Bad behaviour is addressed through concrete action, but the underlying issues are not addressed.	Bad behaviour/ breaking the rules is addressed through concrete action. If any employee breaks the rules, the underlying issues will be identified and rectified. This applies to all employees.
Other	Performance incentives	Does your Directorate use performance, targets, or indicators for tracking and rewarding (financially or non-financially) the performance of its employees?	Staff are rewarded equally (or not rewarded) irrespective of performance. Individual performance is not tracked formally	There is a formal staff evaluation system in place and performance is rewarded (financially or non-financially). However, there are no clear system or criteria for rewarding staff	There is a formal staff evaluation system and performance is rewarded (financially or non- financially). Rewards are given as a consequence of well-defined and monitored individual achievements. This applied to all staff.
Other	Roles	When staff in your Directorate are given tasks in their daily work, how much discretion do they have to carry out their assignments? Can you give me an example?	How officers carry out their assignments is decided by senior managers. Officers have no say.	How officers carry out their assignments is jointly decided by the officer and senior managers. Senior managers tend to drive the decisions.	Officers have complete autonomy in deciding how to carry out their tasks.
Other	Roles	Can most staff in your Directorate make substantive contributions to the policy formulation and implementation process?	Staff do not contribute to policy formulation, nor to decisions about implementation.	Staff can contribute to policy formulation and decisions about implementation, but there is no formal forum through which to do this. Contributions typically only occur when problems arise.	Management expects all staff to contribute to policy formulation and decisions about implementation (formally or informally), and considers this part of their duties.
Other	Roles	Is the workload of achieving your Directorate's targets evenly distributed across its different employees, or do some groups consistently shoulder a greater burden than others?	A small minority of staff undertake the vast majority of work within the directorate.	The burden of the directorate's work is more or less distributed equally among staff. A small minority get away with working significantly less than others.	The burden of the directorate's work is distributed equally among staff. Tasks are assigned in such a way that the amount of time required and the level of difficulty are balanced out so no member of staff finds him/herself overburdened.
Other	Roles	Thinking about all the projects that your Directorate has been involved in since your appointment here, would you say that managers and supervisors try to use the right staff for the right job?	Staff are allocated to tasks randomly.	Managers try to use the right staff for the right job but do not go to great lengths to ensure this, or are met with institutional constraints which may prevent them from doing so.	The right staff are always used for a task. Allocation of tasks is based on staffs' documented skills and competencies.
Other	Flexibility	Does your Directorate make efforts to adjust to the specific needs and requirements of communities, clients, or other stakeholders?	The directorate uses the same procedures no matter what.	The directorate tailors procedures to the specific needs of its stakeholders, but struggles when those needs are complex.	The directorate tailors all procedures to the specific needs of its stakeholders. The evolution of those needs results in adaptation to plans, project and policies.
Other	Flexibility	How flexible would you say your Directorate is in terms of responding to new and improved work practices or reforms?	New practices are not adopted/ integrated in the directorate.	New ideas or practices are adopted, but in an informal and/ or isolated manner. The directorate encourages the adoption of new practices, however it is slow to integrate them into its operations (more than a year).	The adoption of new ideas and practices is an integral part of the directorate's work. New practices are regularly reviewed and considered, and once adopted and integrated across the directorate within 6 months.
Other	Staff involvement/ contribution	How do problems in your directorate get exposed and fixed?	Ad-hoc, no set process for improvement Deal with problems as they arise without following an established procedureOnce fixed, no further action taken No suggestions from staff	Existing process to deal with problems Improvements made through meetings Focus on finding solutions, not prevention of future problems Suggestions from staff involved through meetings (formal or informal)	Exposing problems and suggesting solutions and improvements is part of all staffs' daily duty. Continuous improvement is part of the culture of the organization.
Other	Staff involvement/ contribution	What kind of feedback do you get in staff meetings?	No feedback from staff.	Staff provide feedback in meetings but in an unstructured manner. Focus on bad performance.	Staff provide the feedback on which action plans will be based. Focus on both good and bad performance. Details of the meetings are recorded and communicated to all staff.
Other	Staff involvement/ contribution	Let's say you've agreed to a follow up plan at one of your meetings, what would happen if the plan wasn't enacted?	No action taken. No changes made in the operations process.	Failure can be found in regular meetings (weekly, even monthly for long-term plans) or at standard points before the deadline. Plans can be altered in order to achieve expected results on time.	In addition to 4, tools can be checked up and reported to the manager in charge. Meetings (formal/ informal) are held to look into the root causes of problems and preventive actions are taken for future similar task.
Other	Staffing	Do you think the management of your Directorate think about attracting talented people to your Directorate and then doing their best to keep them? For example, by ensuring they are happy and engaged with their work.	Directorate does not put emphasis on talent	Senior management believes that attracting and developing talent is important, but there is no clear system for identifying, attracting or retaining such talent.	Senior management believes that attracting and developing talent is important. There is a clear system for identifying and attracting talent, developing and retaining talent.
Other	Staffing	If two senior level staff joined your Directorate five years ago and one was much better at their work than the other, would he/she be promoted through the service faster?	No promotion system (no one in the organization has been promoted for years) The promotion system is based on tenure	The promotion system is based on performance. Organization may have internal limitations (e.g. few position openings), but do everything to get around them (e.g. extra training).	Promotion system is based on performance. Organization actively identifies, develops and promotes top performers. Regular assessments, clear set of indicators and personalised career plans for individuals (regularly revised).

Table A4: Factor Analysis of Management Practices

Iterated Principal Factor Analysis

Factor Weights

	(1) Factor 1	(2) Factor 2
Monitoring	0.69	-0.37
Staffing	0.66	0.20
Incentives	0.80	-0.087
Targeting	0.79	-0.17
Staff Involvement	0.86	-0.0033
Roles	0.84	0.10
Flexibility	0.62	0.38
Number of observations	348	348

Notes: The unit of observation is the organization. Management practices are organization-level average management practices. Each topic is comprised of individual items, which are aggregated at the individual manager level and averaged across all managers in an organization. Factor weights are based on a factor analysis of the seven topics for two factors, using iterated principal factors for two factors to be retained. Figures are rounded to two significant figures.

Table A5: Management, District Characteristics, and Official Characteristics

Dependent Variable: Organization Aggregate Management Practices in Column 1; Organization Average Management Practice: Monitoring (Z-Score) in Column 2; Organization Average Management Practice: Other (Z-Score) in Column 3; Organization Average Years in Civil Service in Column 4; Organization Proportion of Civil Servants With Undergraduate Education in Column 5.

OLS Estimates

Standard Errors: Clustered at the district government-sector level in all columns

	(1) Management and District and Official Characteristics	(2) Monitoring and District and Official Characteristics	(3) Alignment and District and Official Characteristics	(4) Average Tenure in Civil Service and District and Official Characteristics	(5) Proportion of Civil Servants With Undergraduate Education and District and Official Characteristics
District Characteristic: Ethnic fractionalization (z-score)	0.074	0.0020	0.086	0.80	0.039
	[0.11]	[0.15]	[0.11]	[0.98]	[0.035]
District Characteristic: Percentage of rural inhabitants (z-score)	-0.047	-0.0031	-0.047	0.54	-0.0022
	[0.087]	[0.11]	[0.085]	[0.86]	[0.038]
District Characteristic: Population (z-score)	0.034	0.15	0.027	1.72	-0.020
	[0.11]	[0.13]	[0.11]	[0.88]	[0.043]
District Characteristic: Remoteness (z-score)	0.13	0.039	0.13	0.21	0.030
	[0.077]	[0.094]	[0.074]	[0.71]	[0.048]
District Characteristic: Poverty rate (z-score)	0.015	0.19	-0.0064	0.27	-0.062
	[0.076]	[0.10]	[0.075]	[0.89]	[0.040]
Individual Characteristic: Years in position (organization average)	0.011	0.010	0.011		
	[0.046]	[0.056]	[0.044]		
Individual Characteristic: Years in organization (organization average)	-0.00010	0.012	-0.0051		
	[0.019]	[0.022]	[0.019]		
Individual Characteristic: Years in civil service (organization average)	0.011	-0.0080	0.015		
	[0.015]	[0.018]	[0.015]		
Individual Characteristic: Undergraduate education [degree=1] (organization average)	-0.036	-0.49	0.048		
	[0.23]	[0.30]	[0.22]		
Individual Characteristic: Postgraduate education [masters=1] (organization average)	0.17	-0.024	0.19		
	[0.35]	[0.38]	[0.35]		
Individual Characteristic: Gender [female=1] (organization average)	0.30	0.30	0.29		
	[0.24]	[0.31]	[0.24]		
P-Value for Joint Test of Significance of District Characteristics	0.34	0.30	0.26	0.30	0.61
P-Value for Joint Test of Significance of Individual Characteristics	0.80	0.57	0.75		
Sector and region fixed effects	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.29	0.29	0.30	0.16	0.20
Observations	122	122	122	122	122

Notes: Robust standard errors are in parentheses. All columns report OLS estimates. The unit of analysis in all columns is the organization at the district level. The sample used for the analysis is the set of district-level organizations in the sample. The dependent variable is the organization-average z-score of Management practice: Monitoring in column 1; the organization-average z-score of Management practice: Other in column 2; the organization-average number of years in the civil service in column 3; the organization-level proportion of civil servants with an undergraduate education or higher in column 4. Management practice: Monitoring is the organization-average z-score for the monitoring practices of management as per Bloom and Van Reenen (2007); Management practice: Other is the mean of the organization average z-scores for targeting, incentives, roles, flexibility, staffing, and staff involvement. Management is equal to one if the claim is from a director. Ethnic fractionalization is the district-level probability that two randomly-selected individuals in the population of the district are from different ethnicities, the remoteness of the district (the census measure of the average travel time to the nearest large urban area), the percentage of rural inhabitants in the district, the population, and the poverty rate based on the latest census. Figures are rounded to two significant figures.

Table A6: Errors and Public Resources

Dependent Variable: Log of total budget

OLS Estimates

Standard Errors: Clustered at the district government-sector level in columns (1) to (3) and at the district level in column (4)

	(1) Population	(2) Own income and poverty rate	(3) District characteristics	(4) Cluster at district level
Log(Population estimate)	0.02 [0.0077]	0.02 [0.0073]	0.02 [0.0073]	0.02 [0.0088]
Log(Population)	0.47 [0.044]	0.45 [0.050]	0.43 [0.049]	0.45 [0.098]
Log(Own income)		0.06 [0.022]	0.06 [0.022]	0.056 [0.047]
Log(Poverty rate)		0.18 [0.036]	0.16 [0.038]	0.18 [0.077]
District characteristics	No	No	Yes	Yes
Adjusted R-squared	0.39	0.42	0.43	0.42
Observations	4074 [408]	4074 [408]	4074 [408]	4074 [85]

Notes: Clustered standard errors are in parentheses. Standard errors clustered at the district government-sector interacted level in columns (1) to (3) and at the district level in column (4). All columns report OLS estimates. The unit of analysis in all columns is the organization at the claim level. The sample used for the analysis is the full set of claims over population. The dependent variable is the log of total budget of the district in all columns. District characteristics include a measure of ethnic fractionalization in the district, the remoteness of the district (the census measure of the average travel time to the nearest large urban area), the percentage of rural inhabitants in the district, and the poverty rate based on the latest census. Income from own sources is missing for 1454 observations. The mean of the rest of the sample is used as an imputed value and an indicator for whether an imputed value is included in columns (2) to (4). Figures are rounded to two significant figures.

Table A7: Errors and Appraisal Scores

Dependent Variable: Within-organization z-score of civil servant's total appraisal score

OLS Estimates

Standard Errors: Clustered at the district government-sector level

	(1) Unconditional	(2) Tier fixed effects	(3) Sector fixed effects	(4) Indicator fixed effects	(5) Percentile fixed effects	(6) Region fixed effects	(7) Individual characteristics	(8) Capital control variables	(9) Management practices
Absolute error (z-score)	-0.021 [0.013]	-0.02 [0.013]	-0.02 [0.013]	-0.02 [0.013]	-0.02 [0.013]	-0.03 [0.013]	-0.02 [0.012]	-0.03 [0.012]	-0.02 [0.012]
Don't know	-0.085 [0.066]	-0.12 [0.080]	-0.13 [0.086]	-0.14 [0.088]	-0.14 [0.088]	-0.096 [0.10]	-0.083 [0.092]	-0.031 [0.094]	-0.040 [0.092]
Tier fixed effects	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector fixed effects	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Indicator fixed effects	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Percentile fixed effects	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Region fixed effects	No	No	No	No	No	Yes	Yes	Yes	Yes
Individual characteristics	No	No	No	No	No	No	Yes	Yes	Yes
Capital control variables and noise controls	No	No	No	No	No	No	No	Yes	Yes
Management practices	No	No	No	No	No	No	No	No	Yes
Adjusted R-squared	0.0024	0.0033	0.0035	-0.00049	-0.00099	0.018	0.082	0.14	0.14
Observations [clusters]	3963 [135]	3963 [135]	3963 [135]	3963 [135]	3963 [135]	3963 [135]	3963 [135]	3963 [135]	3963 [135]

Notes: Clustered standard errors are in parentheses. Standard errors clustered at the district government-sector interacted level. All columns report OLS estimates. The unit of analysis in all columns is the claim of the district government-level official. The sample used for the analysis is the set of respondents making claims for which there is also evaluation data for the civil servant. The dependent variable is the within-organization z-score of the civil servant's evaluation outcome as per their annual appraisal. Tier fixed effects control for whether the respondent works in a federal, regional, or district government office. Sector fixed effects are indicators for whether the claim is within an agriculture, education, health, revenue, or trade organization. Indicator fixed effects control for the specific variable being claimed over. Individual characteristics includes: years in position; years in the organization; years in the civil service; an indicator for whether the respondent has an undergraduate degree; and an indicator for whether the respondent has a masters degree; an indicator for whether the respondent is female; and an indicator for whether the respondent is a manager. Capital control variables include the proportion of managers in the organization that state that a management information system is in place; the organization average response to "In what proportion of projects would you say information flows effectively between individuals in the same organization?"; the organization average response to "In what proportion of projects would you say information flows effectively between organizations?"; the organization average response of managers to "How many personnel do you manage?"; the organization average response of employees to "How many people would you say regularly give you tasks as part of your formal work duties?"; alignment (the z-score of the row mean of the organization-average response to "To what extent do you believe that your organization's mission is aligned to your own mission?" and the organization proportion of selecting "I am contributing to fulfilling that mission on an everyday basis" to "How consistent with your mission are the various tasks and activities assigned to you on a day-to-day basis?"). Noise controls include the time of day of the survey, day-of-survey fixed effects, enumerator fixed effects, an indicator for the enumerator's subjective assessment of the quality of the interview; an indicator for the duration of the interview (decile of duration fixed effects); controls for the average tenure of the managers in the organization and the proportion of female managers in the organization. Management practices include Management practice: Monitoring and Management practice: Other. Management practice: Monitoring is the organization average of the monitoring topic, constructed from the z-scores of each individual item noted in table A3. Management practice: Other is the organization average of the average across the z-scores of targeting, incentives, roles, flexibility, staffing, and staff involvement, detailed in table A3. Figures are rounded to two significant figures.

Table A8: Determinants of Superior Information in Hierarchies

Dependent Variable: Z-score of errors across policy domains

Analysis of Variance (ANOVA) Estimates in column (1); OLS estimates in column (2)

Partial sum of squares and F-statistic p-values in parentheses in column (1); OLS coefficient and clustered standard error at the district-government-sector level, in parentheses in column (2); F-statistics and associated p-values from the joint test across all categories for factor variables preceded by "F=" and "P=" respectively in column (2)

	(1) ANOVA	(2) Regression coefficients
Hierarchical structure		
District Government [Yes=1]	8.71 [0.00]	-0.33 [0.14]
Manager	0.04 [0.83]	0.012 [0.072]
Sector	3.61 [0.41]	F=0.6 P=0.67
Capital control variables		
Feature of Claim: Indicator	27.09 [0.05]	F=1.92 P=0.02
Feature of Claim: Percentile	9.38 [0.01]	F=2.73 P=0.07
Bureaucrat Characteristic: Years in position	0.00 [0.95]	-0.00049 [0.012]
Bureaucrat Characteristic: Years in organization	0.19 [0.65]	-0.0022 [0.0047]
Bureaucrat Characteristic: Years in civil service	0.00 [0.98]	0.00013 [0.0040]
Bureaucrat Characteristic: Undergraduate education [degree=1]	0.52 [0.45]	-0.069 [0.12]
Bureaucrat Characteristic: Postgraduate education [masters=1]	0.02 [0.89]	-0.0092 [0.078]
Bureaucrat Characteristic: Gender [female=1]	0.53 [0.44]	0.046 [0.065]
Organization Characteristic: MIS exists	0.04 [0.84]	0.011 [0.18]
Organization Characteristic: Proportion of projects civil servant has adequate information on service delivery	11.74 [0.00]	0.01 [0.0052]
Organization Characteristic: Proportion of projects information flows effectively within organizations	4.84 [0.02]	-0.01 [0.0057]
Organization Characteristic: Proportion of projects information flows effectively across organizations	0.08 [0.77]	0.0011 [0.0055]
Organization Characteristic: Span of control	3.85 [0.04]	-0.01 [0.0053]
Organization Characteristic: Number of principals	3.44 [0.05]	0.075 [0.050]
Marginal benefits		
Management practices: Monitoring	4.35 [0.03]	-0.15 [0.072]
Management practices: Other	3.33 [0.05]	0.16 [0.11]
Alignment: Mission alignment index	1.10 [0.27]	0.14 [0.18]
Model	188.47 [0.00]	
Residual	1239.29	
District fixed effects		Yes
Adjusted R-squared	-	0.080
Number of observations [clusters]	1455	1455 [92]

Notes: Column (1) reports ANOVA partial sum of square estimates. Column (2) reports OLS regression coefficients. Standard errors clustered at the district-government-sector level. Column (2) reports the standard error in brackets; for factor variables with multiple categories, the F-statistic of joint test of all category indicators are reported after F= and the p-value is reported below after P=. The unit of analysis in all columns is the claim. The sample used for the analysis is the set of respondents that make claims about the same districts. Sector fixed effects are indicators for whether the claim is within an agriculture, education, health, revenue, or trade organization. Percentile fixed effects are indicators for the 75th, 50th or 25th percentile for the respective indicator for the respective region. Indicator fixed effects are indicator variables for which specific variable the claim is being made over. Manager is equal to one if the claim is from a manager. MIS exists is the organization proportion of managers that respond "Yes" to "Do you currently have a Management Information System (MIS) programme in place and functioning?". Proportion of projects civil servant has adequate information on service delivery is the organization average response to "In what proportion of projects would you say you have necessary information regarding the current state of service delivery in your jurisdiction / area of work?". Proportion of projects information flows effectively within organizations is the organization average response to "In what proportion of projects would you say information flows effectively between individuals in the same organization?". Proportion of projects information flows effectively across organizations is the organization average response to "In what proportion of projects would you say information flows effectively between organizations?". Span of control is the organization average response across managers of the response to "How many personnel do you manage?". Number of principals is the organization average response of employees to "How many people would you say regularly give you tasks as part of your formal work duties?". Management practice: Monitoring is the organization average of the monitoring topic, constructed from the z-scores of each individual item noted in table A3. Management practice: Other is the organization average of the average across the z-scores of targeting, incentives, roles, flexibility, staffing, and staff involvement, detailed in table A3. Alignment: Mission alignment index is the z-score of the row mean of the organization-average response to "To what extent do you believe that your organization's mission is aligned to your own mission?" and the organization proportion of employees selecting "I am contributing to fulfilling that mission on an everyday basis" to "How consistent with your mission are the various tasks and activities assigned to you on a day-to-day basis?". Figures are rounded to two significant figures.

Table A9: Heterogeneity in Effects: Quantile Regression By Absolute Error

Dependent Variable: Z-score of errors across policy domains in columns

Quantile regression estimates

Standard Errors: Clustered at the district government-sector level

	(1) Percentile = 0.1	(2) Percentile = 0.2	(3) Percentile = 0.3	(4) Percentile = 0.4	(5) Percentile = 0.5	(6) Percentile = 0.6	(7) Percentile = 0.7	(8) Percentile = 0.8	(9) Percentile = 0.9
Federal Government [Yes=1]	0.0095 [0.034]	0.045 [0.038]	0.057 [0.053]	0.055 [0.055]	0.071 [0.072]	0.014 [0.078]	-0.011 [0.092]	-0.021 [0.10]	0.066 [0.13]
District Government [Yes=1]	-0.08 [0.039]	-0.15 [0.053]	-0.21 [0.080]	-0.21 [0.078]	-0.20 [0.076]	-0.21 [0.088]	-0.27 [0.10]	-0.40 [0.11]	-0.26 [0.14]
Manager [Yes=1]	0.019 [0.029]	0.019 [0.031]	0.031 [0.031]	0.038 [0.032]	0.0074 [0.051]	-0.014 [0.066]	0.0035 [0.075]	0.023 [0.090]	0.083 [0.10]
Sector fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District government fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Indicator fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Percentile fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1455 [92]								

Notes: Standard errors are in parentheses. Standard errors clustered at the district-government-sector interacted level. Columns report quantile regression estimates for different quantiles of the distribution of the z-score of absolute errors. The unit of analysis in all columns is the claim. The sample used for the analysis is the set of respondents that make claims about the same districts. The dependent variable is the z-score of the absolute error. Sector fixed effects are indicators for whether the claim is within an agriculture, education, health, revenue, or trade organization. Percentile fixed effects are indicators for the 75th, 50th or 25th percentile for the respective indicator for the respective region. Indicator fixed effects control for the service delivery indicator within which the claim is made. Manager is equal to one if the claim is from a manager. Individual characteristics included in the specification: tenure in the civil service of the respondent making the claim; tenure in the organization; an indicator for whether the respondent is female; an indicator for whether the respondent holds an undergraduate degree; an indicator for whether the respondent holds a masters degree. Figures are rounded to two significant figures.

Table A10: Decentralization and Information Robustness

Dependent Variable: Z-score of errors across policy domains in columns (1) to (5) and column (8); Rank of errors, in terms of the proportion within distribution, across policy domains in column (6); Skewness adjusted z-scores of errors across policy domains in column (7)

OLS Estimates

Standard Errors: Clustered at the district government-sector level in column (1) and columns (6) to (8); Clustered at the claimed district government level in column (2); Clustered at the respondent level in column (3); Clustered at the organization-of-respondent level in column (4); Clustered at the region-sector level in column (5).

	(1) Disaggregating central dummy	(2) Clustering at claimed district government level	(3) Clustering at respondent level	(4) Clustering at the organization-of-respondent level	(5) Clustering at region-sector level	(6) Rank of Error (proportion)	(7) Skewness adjusted z-score	(8) Management controls
District Government [Yes=1]	-0.31 [0.098]	-0.33 [0.096]	-0.33 [0.064]	-0.33 [0.081]	-0.33 [0.094]	-0.10 [0.027]	-0.39 [0.098]	-0.28 [0.16]
Federal Government [Yes=1]	0.042 [0.091]							
Manager	0.039 [0.067]	0.034 [0.070]	0.034 [0.065]	0.034 [0.068]	0.034 [0.073]	-0.00034 [0.018]	0.029 [0.059]	0.060 [0.11]
Sector fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District government fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Indicator fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Percentile fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region fixed effects	No	No	No	No	No	No	No	Yes
Capital control variables and noise controls	No	No	No	No	No	No	No	Yes
Adjusted R-squared	0.068	0.069	0.069	0.069	0.069	0.11	0.11	0.057
Observations [clusters]	1455 [92]	1455 [42]	1455 [687]	1455 [122]	1455 [38]	1455 [92]	1455 [92]	1455 [92]
Federal respondents [percentage]	195 [28%]	195 [28%]	195 [28%]	195 [28%]	195 [28%]	195 [28%]	195 [28%]	195 [28%]
Regional respondents [percentage]	234 [34%]	234 [34%]	234 [34%]	234 [34%]	234 [34%]	234 [34%]	234 [34%]	234 [34%]
District government respondents [percentage]	260 [38%]	260 [38%]	260 [38%]	260 [38%]	260 [38%]	260 [38%]	260 [38%]	260 [38%]

Notes: Clustered standard errors are in parentheses. Standard errors clustered at the district government-sector interacted level in column (1) and columns (6) to (8); clustered at the claimed district level in column (2); clustered at the respondent level in column (3); clustered at the organization-of-respondent level in column (4); clustered at the region-sector interacted level in column (5). All columns report OLS estimates. The unit of analysis in all columns is the claim. The sample used for the analysis is the set of respondents that make claims about the same districts. The dependent variable is the z-score of absolute error across all policy domain items in columns (1) to (5) and column (6), where the z-score is calculated within the distribution of errors within an indicator; the rank of the error in proportion terms (proportion of the maximum rank within the indicator) within the distribution of errors within an indicator in column (6); the skewness-adjusted z-score is the dependent variable in column (7) - the skewness adjusted z-score is the z-score of the absolute error if the variable is a proportion variable (between 0 and 1), if the variable is not a proportion and is statistically significantly skewed relative to the normal distribution (Royston,1992), then the log of the variable is taken and then the z-score is created from the logged values. Sector fixed effects are indicators for whether the claim is within an agriculture, education, health, revenue, or trade organization. Percentile fixed effects are indicators for the 75th, 50th or 25th percentile for the respective indicator for the respective region. District government fixed effects identify the district that the respondent is making claims about. Indicator fixed effects control for the service delivery indicator within which the claim is made. Management is equal to one if the claim is from a director. Individual characteristics included in the specification: tenure in the civil service of the respondent making the claim; tenure in the position; tenure in the organization; an indicator for whether the respondent is female; an indicator for whether the respondent holds an undergraduate degree; an indicator for whether the respondent holds a masters degree. Capital control variables include the proportion of managers in the organization that state that a management information system is in place; the organization average response to "In what proportion of projects would you say information flows effectively between individuals in the same organization?"; the organization average response to "In what proportion of projects would you say information flows effectively between organizations?"; the organization average response of managers to "How many personnel do you manage?"; the organization average response of employees to "How many people would you say regularly give you tasks as part of your formal work duties?"; alignment (the z-score of the row mean of the organization-average response to "To what extent do you believe that your organization's mission is aligned to your own mission?" and the organization proportion of selecting "I am contributing to fulfilling that mission on an everyday basis" to "How consistent with your mission are the various tasks and activities assigned to you on a day-to-day basis?"). Noise controls include the time of day of the survey, day-of-survey fixed effects, enumerator fixed effects, an indicator for the enumerator's subjective assessment of the quality of the interview; an indicator for the duration of the interview (decile of duration fixed effects); controls for the average tenure of the managers in the organization and the proportion of female managers in the organization. Figures are rounded to two significant figures.

Table A11: Decentralization and Management

Dependent Variable: Management practice: Monitoring in columns (1) to (4); Management practice: Other in columns (5) to (8)

OLS Estimates

Robust standard errors

	Management practice: Monitoring				Management practice: Other			
	(1) Unconditional	(2) Sector FE	(3) Individual characteristics (organization average)	(4) Capital control variables	(5) Unconditional	(6) Sector FE	(7) Individual characteristics (organization average)	(8) Capital control variables
District Government [Yes=1]	0.13 [0.15]	0.16 [0.15]	0.41 [0.21]	0.54 [0.21]	-0.12 [0.12]	-0.10 [0.12]	0.14 [0.14]	0.25 [0.14]
Sector fixed effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Individual characteristics (organization average)	No	No	Yes	Yes	No	No	Yes	Yes
Capital control variables	No	No	No	Yes	No	No	No	Yes
Adjusted R-squared	-0.0023	0.034	0.078	0.25	0.00018	0.049	0.10	0.24
Observations	122	122	122	122	122	122	122	122

Notes: Standard errors are in parentheses. Robust standard errors. All columns report OLS estimates. The unit of analysis in all columns is the organization. The tiers of government are grouped into District Government and Centralized Government which incorporates both Federal and Regional tiers, hence the omitted category for tier in the regressions is Centralized Government. The sample used for the analysis is the set of respondents that make claims about the same districts. The dependent variable is Management practice: Monitoring in columns (1) to (4) and Management practice: Other in columns (5) to (8). Management practice: Monitoring is the organization average of the monitoring topic, constructed from the z-scores of each individual item noted in table A3. Management practice: Other is the organization average of the average across the z-scores of targeting, incentives, roles, flexibility, staffing, and staff involvement, detailed in table A3. Management practices are in z-scores. Individual characteristics (organization average) include the organization average of: the number of years in current position; the number of years in current organization; the number of years in the service; an indicator for whether the employee has an undergraduate degree; an indicator for whether the employee has a masters degree; and an indicator for whether the employee is female. Capital control variables include the proportion of managers in the organization that state that a management information system is in place; the organization average response to "In what proportion of projects would you say information flows effectively between individuals in the same organization?"; the organization average response to "In what proportion of projects would you say information flows effectively between organizations?"; the organization average response of managers to "How many personnel do you manage?"; the organization average response of employees to "How many people would you say regularly give you tasks as part of your formal work duties?"; alignment (the z-score of the row mean of the organization-average response to "To what extent do you believe that your organization's mission is aligned to your own mission?" and the organization proportion of selecting "I am contributing to fulfilling that mission on an everyday basis" to "How consistent with your mission are the various tasks and activities assigned to you on a day-to-day basis?"). Figures are rounded to two significant figures.

Table A14: Robustness: Decentralization and Management

Dependent Variable: Z-score of errors across policy domains in all columns
OLS Estimates
Standard Errors: Clustered at the district government-sector level in all columns

	Monitoring and Incentives		Exclude Capital Control Variables		District Characteristics	
	(1) District Government Baseline	(2) Interaction Between District Government and Monitoring & Incentives	(3) District Government Baseline	(4) Interaction Between District Government and Monitoring & Incentives	(5) District Government Baseline	(6) Interaction Between District Government and Monitoring & Incentives
District Government [Yes=1]	-0.26 [0.16]	-0.33 [0.16]	-0.26 [0.094]	-0.26 [0.092]	-0.27 [0.16]	-0.35 [0.17]
Manager	-0.026 [0.076]	-0.016 [0.076]	-0.0022 [0.074]	0.0034 [0.075]	-0.027 [0.076]	-0.019 [0.077]
Management Practice: Monitoring & Incentives	-0.18 [0.11]	-0.066 [0.12]				
Management Practice: Other	0.29 [0.13]	0.28 [0.12]	0.32 [0.099]	0.32 [0.098]	0.31 [0.11]	0.32 [0.10]
District Government [Yes=1] x Management Practice: Monitoring & Incentives		-0.28 [0.15]				
Management Practice: Monitoring			-0.16 [0.071]	-0.14 [0.086]	-0.21 [0.067]	-0.14 [0.080]
District Government [Yes=1] x Management Practice: Monitoring				-0.072 -0.072		-0.20 [0.14]
Sector, region, percentile, indicator fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
Noise controls	Yes	Yes	Yes	Yes	Yes	Yes
Capital control variables	Yes	Yes	No	No	Yes	Yes
District characteristics	No	No	No	No	Yes	Yes
Adjusted R-squared	0.061	0.063	0.056	0.055	0.064	0.066
Observations [clusters]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]

Notes: denotes significance at 1%, at 5%, and at 10% level. Clustered standard errors are in parentheses. Standard errors clustered at the district government-sector interacted level in all columns. All columns report OLS estimates. The unit of analysis in all columns is the claim. The sample used for the analysis is the set of respondents that make claims about the same districts. The dependent variable is the z-score of absolute error across all policy domain items in all columns. Sector fixed effects are indicators for whether the claim is within an agriculture, education, health, revenue, or trade organization. Percentile fixed effects are indicators for the 75th, 50th or 25th percentile for the respective indicator for the respective region. District government fixed effects identify the district that the respondent is making claims about. Indicator fixed effects control for the service delivery indicator within which the claim is made. Management practice: Monitoring & Incentives is the organization-average z-score for the monitoring and incentives practices of management as per Bloom and Van Reenen (2007); Management practice: Other is the mean of the organization average z-scores for targeting, roles, flexibility, staffing, and staff involvement. Management is equal to one if the claim is from a director. Management practice: Monitoring is the organization-average z-score for the monitoring practices of management as per Bloom and Van Reenen (2007); Management practice: Other is the mean of the organization average z-scores for targeting, incentives, roles, flexibility, staffing, and staff involvement. Management is equal to one if the claim is from a director. Individual characteristics included in the specification: tenure in the civil service of the respondent making the claim; tenure in the position; tenure in the organization; an indicator for whether the respondent is female; an indicator for whether the respondent holds an undergraduate degree; an indicator for whether the respondent holds a masters degree. Capital control variables include the proportion of managers in the organization that state that a management information system is in place; the organization average response to "In what proportion of projects would you say information flows effectively between individuals in the same organization?"; the organization average response to "In what proportion of projects would you say Information flows effectively between organizations?"; the organization average response of managers to "How many personnel do you manage?"; the organization average response of employees to "How many people would you say regularly give you tasks as part of your formal work duties?"; alignment (the z-score of the row mean of the organization-average response to "To what extent do you believe that your organization's mission is aligned to your own mission?" and the organization proportion of selecting "I am contributing to fulfilling that mission on an everyday basis" to "How consistent with your mission are the various tasks and activities assigned to you on a day-to-day basis?"). Noise controls include the time of day of the survey, day-of-survey fixed effects, enumerator fixed effects, an indicator for the enumerator's subjective assessment of the quality of the interview; an indicator for the duration of the interview (decile of duration fixed effects); controls for the average tenure of the managers in the organization and the proportion of female managers in the organization. District characteristics include a measure of ethnic fractionalization in the district, the remoteness of the district (the census measure of the average travel time to the nearest large urban area), the percentage of rural inhabitants in the district, the population, and the poverty rate based on the latest census. Figures are rounded to two significant figures.

Table A13: Balance Table**Means and standard deviations**

	(1) Control	(2) Treatment	(3) T-test [1]=[2]
Official's Characteristics			
Age	34.12 [8.34]	34.04 [8.73]	0.08 [0.52]
Years in position	2.92 [3.03]	2.54 [2.34]	0.38 [0.16]
Years in organization	7.69 [7.10]	7.34 [7.24]	0.34 [0.44]
Years in civil service	12.04 [8.25]	12.46 [8.90]	-0.42 [0.53]
Number of different organizations in service	1.32 [1.52]	1.46 [1.97]	-0.14 [0.11]
Grade	7.26 [12.59]	6.66 [10.22]	0.60 [0.69]
Education [undergraduate degree=1]	0.76 [0.43]	0.76 [0.43]	0.00 [0.03]
Education [masters degree=1]	0.01 [0.12]	0.01 [0.11]	0.00 [0.01]
Gender [female=1]	0.19 [0.39]	0.17 [0.38]	0.02 [0.02]
Manager	0.27 [0.44]	0.26 [0.44]	0.01 [0.03]
Number of bureaucrats	441	679	1120
Organizational Characteristics			
Number of bureaucrats per organization	4.43 [0.79]	4.44 [0.91]	-0.01 [0.10]
Number of heads per organization	0.95 [0.21]	0.97 [0.16]	-0.02 [0.02]
Number of managers per organization	0.93 [0.26]	0.89 [0.32]	0.04 [0.03]
Number of employees per organization	2.54 [0.73]	2.58 [0.78]	-0.03 [0.09]
Ratio of employees to heads	2.56 [0.71]	2.58 [0.78]	-0.02 [0.09]
Ratio of employees to managers	2.52 [0.74]	2.61 [0.71]	-0.09 [0.09]
Span of control (employees per manager)	5.64 [6.06]	6.20 [8.43]	-0.56 [0.90]
Number of organizations	127	196	323

Notes: Standard deviations are in parentheses. The unit of observation for the first panel is the civil servant and for the second panel the organization.

Only employees and managers responded to the information module and hence only these individuals are used for the statistics in the first panel. These statistics are based on the survey sample and not on administrative data. The span of control is taken from the survey question "How many personnel do you manage?"; the total is taken for each organization if there is more than one manager; and the organization-level statistics are reported (each organization has an equal weight in the 'By Organization' statistics). Heads refers to heads of organization or deputy heads of organization. 'Manager' refers to managers within the organization. Grade is the official civil service grade of the professional civil servant, ranging from 1 to 17 in the sample. The 2013-14 National Civil Service Human Resource Statistics Abstract (Ministry of Civil Service) records the overall female employee percentage to be 35%; this statistic includes frontline staff. Figures are rounded to two decimal places.

Table A14: Information Experiment Unconditional to Conditional Results

Dependent Variable: Z-score of errors across policy domains in all columns

OLS estimates in all columns

Standard Errors: Clustered at the district government-sector level in all columns

	(1) Treatment only	(2) Sector FEs	(3) Region FEs	(4) Percentile FEs	(5) Indicator FEs	(6) Individual Controls	(7) Capital Control variables	(8) Noise Controls
Treated District [Yes=1]	0.053	0.072	0.097	0.097	0.10	0.099	-0.042	-0.15
	[0.087]	[0.097]	[0.090]	[0.088]	[0.086]	[0.088]	[0.083]	[0.085]
<i>P-value of coefficient clustered at district government-sector level</i>	{0.54}	{0.46}	{0.28}	{0.27}	{0.23}	{0.27}	{0.62}	{0.078}
<i>P-value of coefficient clustered at district level</i>	(0.51)	(0.43)	(0.27)	(0.25)	(0.20)	(0.23)	(0.65)	(0.057)
Sector fixed effects	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region fixed effects	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Percentile fixed effects	No	No	No	Yes	Yes	Yes	Yes	Yes
Indicator fixed effects	No	No	No	No	Yes	Yes	Yes	Yes
Individual controls	No	No	No	No	No	Yes	Yes	Yes
Capital control variables	No	No	No	No	No	No	Yes	Yes
Noise controls	No	No	No	No	No	No	No	Yes
Adjusted R-squared	-0.00018	-0.0018	0.021	0.034	0.030	0.027	0.043	0.057
Number of observations [clusters]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]

Notes: All columns report OLS regression coefficients and standard errors in brackets. Standard errors in brackets. Standard errors clustered at the district government-sector interacted level in all columns. The unit of analysis in all columns is the claim. The sample used for the analysis is the set of respondents that make claims about the same districts. The dependent variable is the z-score error for any claim made within any policy domain in all columns. The tiers of government are grouped into District Government and Centralized Government which incorporates both Federal and Regional tiers, hence the omitted category for tier in the regressions is Centralized Government. Treated district is an indicator equal to one if the district received an information package. Sector fixed effects are indicators for whether the claim is within an agriculture, education, health, revenue, or trade organization. Percentile fixed effects are indicators for the 75th, 50th or 25th percentile for the respective indicator for the respective region. Indicator fixed effects are indicator variables for which specific variable the claim is being made over. Management practice: Monitoring is the organization-average z-score for the monitoring practice of management as per Bloom and Van Reenen (2007); Management practice: Other is the mean of the organization average z-scores for targeting, incentives, roles, flexibility, staffing, and staff involvement. Manager is equal to one if the claim is from a manager. Individual characteristics included in the specification: tenure in the civil service of the respondent making the claim; tenure in the position; tenure in the organization; an indicator for whether the respondent is female; an indicator for whether the respondent holds an undergraduate degree; an indicator for whether the respondent holds a masters degree. Capital control variables include the proportion of managers in the organization that state that a management information system is in place; the organization average response to "In what proportion of projects would you say information flows effectively between individuals in the same organization?"; the organization average response to "In what proportion of projects would you say Information flows effectively between organizations?"; the organization average response of managers to "How many personnel do you manage?"; the organization average response of employees to "How many people would you say regularly give you tasks as part of your formal work duties?"; alignment (the z-score of the row mean of the organization-average response to "To what extent do you believe that your organization's mission is aligned to your own mission?" and the organization proportion of selecting "I am contributing to fulfilling that mission on an everyday basis" to "How consistent with your mission are the various tasks and activities assigned to you on a day-to-day basis?"). Noise controls include the time of day of the survey, day-of-survey fixed effects, enumerator fixed effects, an indicator for the enumerator's subjective assessment of the quality of the interview; an indicator for the duration of the interview (decile of duration fixed effects); controls for the average tenure of the managers in the organization and the proportion of female managers in the organization. Figures are rounded to two significant figures.

Table A15: Robustness: Information Experiment

Dependent Variable: Z-score of errors across policy domains in all columns

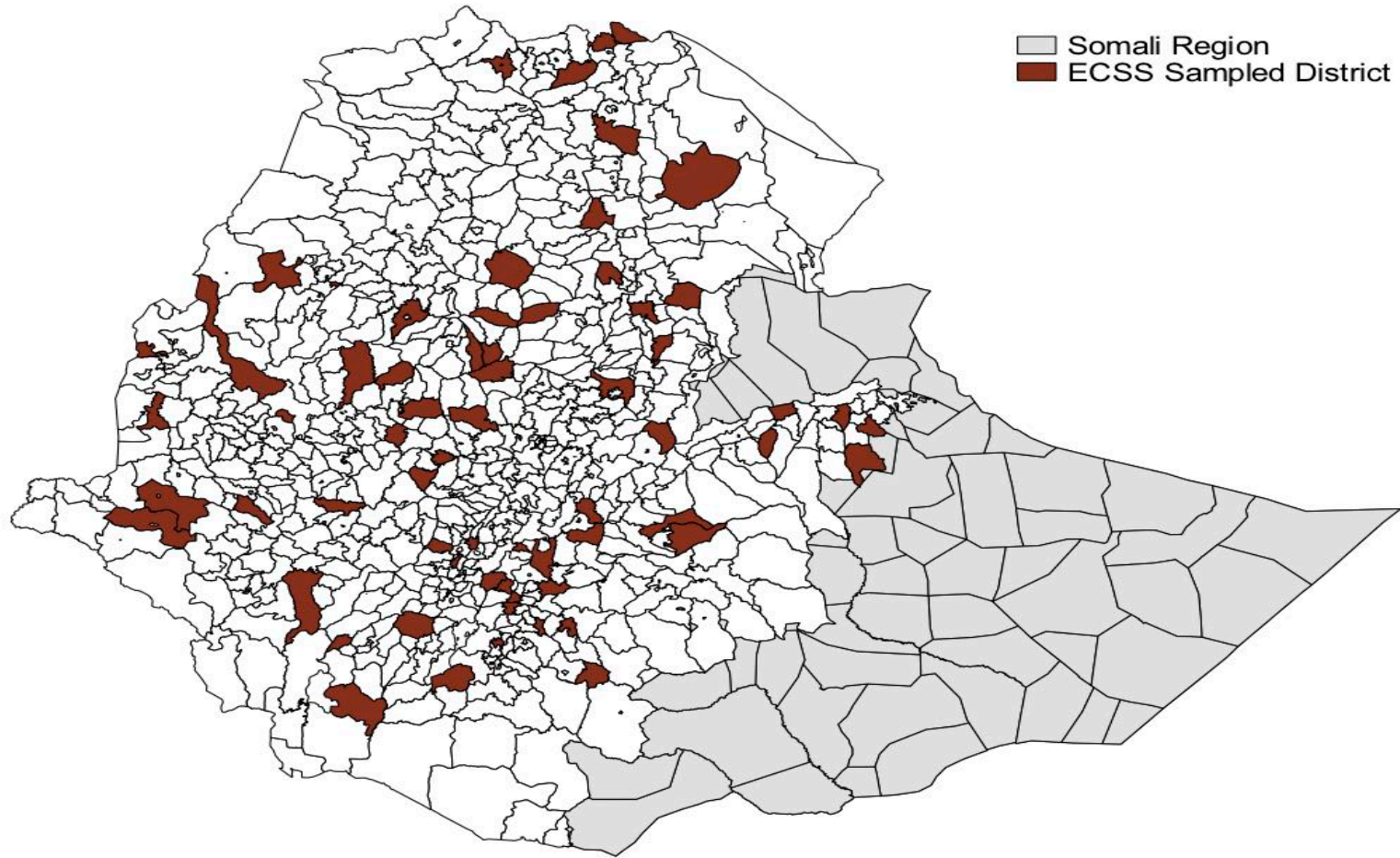
OLS Estimates

Standard Errors: Clustered at the district government-sector level in all columns

	Monitoring and Incentives			Exclude Capital Control Variables			District Characteristics		
	(1) Treated District Baseline	(2) Interaction Between Treated District and Monitoring & Incentives	(3) Interaction Between Treated District and Other	(4) Treated District Baseline	(5) Interaction Between Treated District and Monitoring & Incentives	(6) Interaction Between Treated District and Other	(7) Treated District Baseline	(8) Interaction Between Treated District and Monitoring & Incentives	(9) Interaction Between Treated District and Other
District Government [Yes=1]	-0.41 [0.17]	-0.47 [0.16]	-0.43 [0.16]	-0.28 [0.093]	-0.33 [0.090]	-0.32 [0.092]	-0.40 [0.17]	-0.45 [0.17]	-0.42 [0.17]
Manager	-0.025 [0.075]	-0.012 [0.076]	-0.015 [0.076]	0.000059 [0.073]	0.0066 [0.074]	0.0078 [0.074]	-0.024 [0.076]	-0.016 [0.077]	-0.014 [0.077]
Management Practice: Monitoring & Incentives	-0.18 [0.11]	-0.49 [0.14]	-0.17 [0.11]						
Management Practice: Other	0.27 [0.13]	0.29 [0.13]	0.025 [0.17]	0.31 [0.11]	0.32 [0.10]	0.061 [0.15]	0.27 [0.11]	0.30 [0.11]	0.028 [0.14]
Treated District [Yes=1]	-0.26 [0.098]	-0.36 [0.098]	-0.36 [0.11]						
Treated District x Management Practice: Monitoring & Incentives		0.41 [0.13]		-0.039 [0.099]	-0.069 [0.094]	-0.15 [0.10]	-0.21 [0.11]	-0.24 [0.10]	-0.31 [0.11]
Treated District x Management Practice: Other			0.35 [0.14]			0.36 [0.14]			0.38 [0.14]
Management Practice: Monitoring				-0.16 [0.076]	-0.38 [0.11]	-0.16 [0.071]	-0.19 [0.069]	-0.43 [0.11]	-0.19 [0.066]
Treated District x Management Practice: Monitoring					0.30 [0.13]			0.33 [0.12]	
Sector, region, percentile, indicator fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Noise controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Capital control variables	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes
District characteristics	No	No	No	No	No	No	Yes	Yes	Yes
Adjusted R-squared	0.064	0.071	0.068	0.055	0.060	0.060	0.066	0.072	0.071
Observations [clusters]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]	1455 [92]

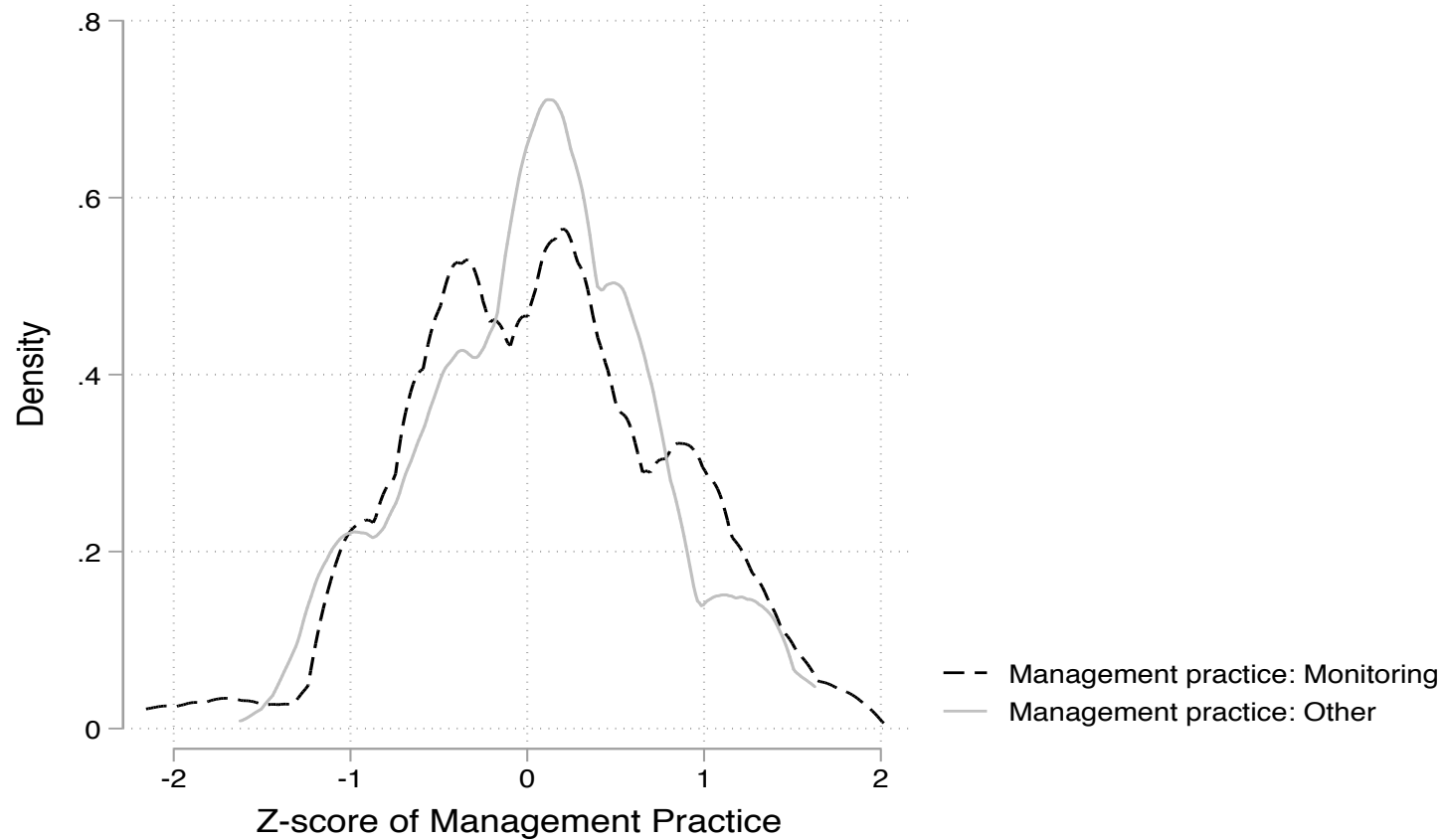
Notes: denotes significance at 1%, at 5%, and at 10% level. Clustered standard errors are in parentheses. Standard errors clustered at the district government-sector interacted level in all columns. All columns report OLS estimates. The unit of analysis in all columns is the claim. The sample used for the analysis is the set of respondents that make claims about the same districts. The dependent variable is the z-score of absolute error across all policy domain items in all columns. Sector fixed effects are indicators for whether the claim is within an agriculture, education, health, revenue, or trade organization. Percentile fixed effects are indicators for the 75th, 50th or 25th percentile for the respective indicator for the respective region. District government fixed effects identify the district that the respondent is making claims about. Indicator fixed effects control for the service delivery indicator within which the claim is made. Management practice: Monitoring & Incentives is the organization-average z-score for the monitoring and incentives practices of management as per Bloom and Van Reenen (2007); Management practice: Other is the mean of the organization average z-scores for targeting, roles, flexibility, staffing, and staff involvement. Management is equal to one if the claim is from a director. Management practice: Monitoring is the organization-average z-score for the monitoring practices of management as per Bloom and Van Reenen (2007); Management practice: Other is the mean of the organization average z-scores for targeting, incentives, roles, flexibility, staffing, and staff involvement. Management is equal to one if the claim is from a director. Individual characteristics included in the specification: tenure in the civil service of the respondent making the claim; tenure in the position; tenure in the organization; an indicator for whether the respondent is female; an indicator for whether the respondent holds an undergraduate degree; an indicator for whether the respondent holds a masters degree. Capital control variables include the proportion of managers in the organization that state that a management information system is in place; the organization average response to "In what proportion of projects would you say information flows effectively between individuals in the same organization?"; the organization average response to "In what proportion of projects would you say information flows effectively between organizations?"; the organization average response of managers to "How many personnel do you manage?"; the organization average response of employees to "How many people would you say regularly give you tasks as part of your formal work duties?"; alignment (the z-score of the row mean of the organization-average response to "To what extent do you believe that your organization's mission is aligned to your own mission?" and the organization proportion of selecting "I am contributing to fulfilling that mission on an everyday basis" to "How consistent with your mission are the various tasks and activities assigned to you on a day-to-day basis?"). Noise controls include the time of day of the survey, day-of-survey fixed effects, enumerator fixed effects, an indicator for the enumerator's subjective assessment of the quality of the interview; an indicator for the duration of the interview (decile of duration fixed effects); controls for the average tenure of the managers in the organization and the proportion of female managers in the organization. District characteristics include a measure of ethnic fractionalization in the district, the remoteness of the district (the census measure of the average travel time to the nearest large urban area), the percentage of rural inhabitants in the district, the population, and the poverty rate based on the latest census. Figures are rounded to two significant figures.

Figure A1: Map of Sampled Districts for Ethiopian Civil Servants Survey



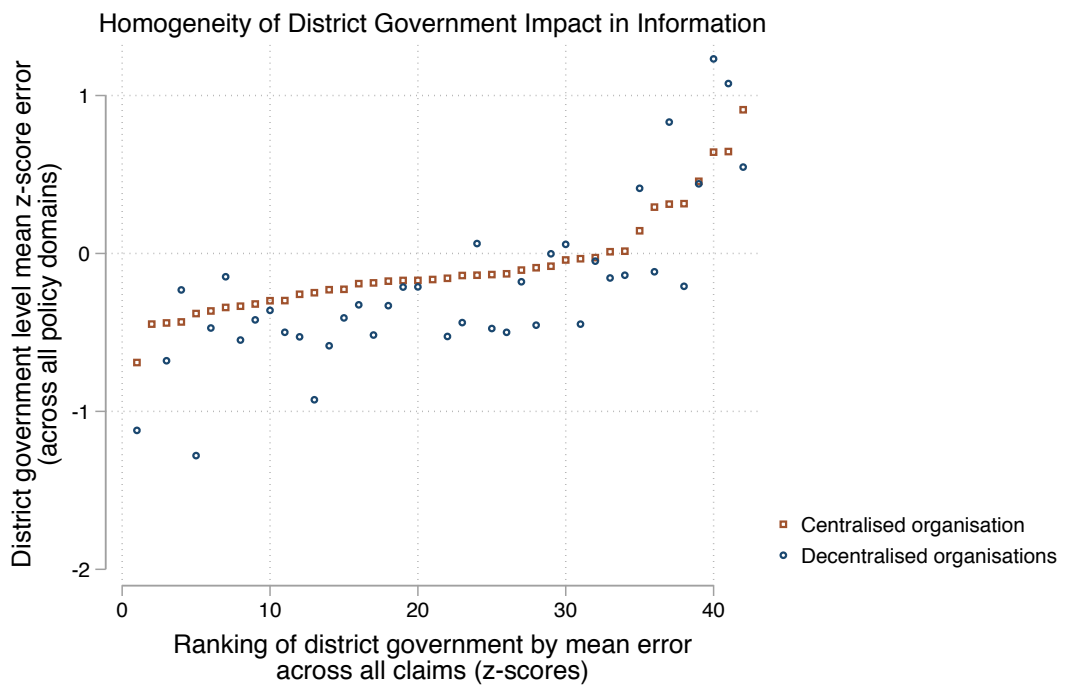
Notes: The figure shows a heat map of districts across Ethiopia based on the Census 2007. The districts shaded in dark red are those sampled in the Ethiopian Civil Servants Survey. The gray areas are districts within the Somali region of Ethiopia, which was mostly excluded from the sampling frame of the survey due to security considerations.

Figure A2: Distributions of Management Indices



Notes: The figure shows the distributions of the Management practice: Monitoring (black dashed line) and Management practice: Other (gray solid line). Management practice: Monitoring is the organization average of the monitoring topic, constructed from the z-scores of each individual item noted in table A3. Management practice: Other is the organization average of the average across the z-scores of targeting, incentives, roles, flexibility, staffing, and staff involvement, detailed in table A3.

Figure A3: Assessing Homogeneity of District Impact



Notes: The figure shows the average claims for centralized (federal and regional) organizations and decentralized (district government) organizations, averaged by the district government that they are claiming over. The red squares refer to centralized organizations and the blue circles to decentralized organizations. The sample used for the analysis is the set of respondents that make claims about the same districts. The error is in terms of z-score within the indicator.