MANAGEMENT AND BUREAUCRATIC EFFECTIVENESS: EVIDENCE FROM THE GHANAIAN CIVIL SERVICE^{*}

IMRAN RASUL

DANIEL ROGGER

MARTIN J.WILLIAMS[†]

JULY 2019

Abstract

We study the relationship between management practices and bureaucratic output, using a survey of the universe of Ghanaian civil servants across 45 Ministries and Departments, combined with hand-coded administrative data on over 3600 projects these organizations undertake. We first demonstrate that there is considerable variation across government organizations, both in management styles/quality, and project delivery. We then show that project delivery exhibits a positive partial correlation with autonomy/discretion-related management practices, but a negative partial correlation with incentives/monitoring-related management practices. While these results contrast with the frequent policy emphasis on introducing top-down monitoring and incentives as a means to elicit effort from bureaucrats, we show the findings are consistent with theories of bureaucratic coordination, intrinsic motivation, and influence activities, that all imply the provision of incentives/monitoring can reduce output in bureaucracies. Finally, we show the external validity of these relationship between management practices and bureaucratic output in a separate sample of bureaucrats and outputs from the Nigerian Civil Service. JEL Classification: J33, O20.

^{*}We gratefully acknowledge financial support from the International Growth Centre [1-VCS-VGHA-VXXXX-33301] and the World Bank's i2i trust fund financed by UKAID. We thank Julien Labonne, Anandi Mani, Don Moynihan, Arup Nath, Simon Quinn, Raffaella Sadun, Itay Saporta-Eksten, Chris Woodruff and numerous seminar participants for valuable comments. Jane Adjabeng, Mohammed Abubakari, Julius Adu-Ntim, Temilola Akinrinade, Sandra Boatemaa, Eugene Ekyem, Paula Fiorini, Margherita Fornasari, Jacob Hagan-Mensah, Allan Kasapa, Kpadam Opuni, Owura Simprii-Duncan, and Liah Yecalo-Tecle provided excellent research assistance, and we are grateful to Ghana's Head of Civil Service, Nana Agyekum-Dwamena, members of the project steering committee, and the dozens of Civil Servants who dedicated their time and energy to the project. This study was approved by UCL's Research Ethics Committee. All errors remain our own.

[†]Rasul: University College London and the Institute for Fiscal Studies [i.rasul@ucl.ac.uk]; Rogger: World Bank Research Department [drogger@worldbank.org]; Williams: Blavatnik School of Government, University of Oxford [martin.williams@bsg.ox.ac.uk].

1 Introduction

The relationship between the management practices under which public servants operate and bureaucratic output is central to research in public administration [Wilson 1989, Lynn *et al.* 2000, Ingraham *et al.* 2003], and in understanding how state capabilities shape economic development [Besley and Persson 2011, Acemoglu and Robinson 2012, Pepinsky *et al.* 2017]. Two broad schools of thought stand out. If bureaucrats are viewed as agents who minimize effort or whose preferences otherwise diverge from their principals, then they should be managed through top-down tools of control such as incentive and monitoring schemes that elicit effort and minimize moral hazard [Duflo *et al.* 2012]. An alternative view is that public bureaucracies ought to delegate significant autonomy and discretion to bureaucrats, relying on their professionalism and expertise to deliver public services [Simon 1983, Rose-Ackerman 1986, Miller and Whitford 2016].

The existing evidence has some important limitations. Many empirical studies of management practices – such as incentive programs – focus on one instance of the practice (often implemented under controlled conditions), on one type of worker (often frontline bureaucrats), or on one type of bureaucratic output [Meier and O'Toole 2002, Hasnain *et al.* 2012, Khan *et al.* 2015]. At the same time, investigations of performance across a fuller range of public sector activities has often relied on subjective measures of organizational performance [Meier and O'Toole 2012]. Meanwhile, quantitative studies on bureaucratic discretion have mainly focused on its potential downsides in terms of discrimination [Einstein and Glick 2017] or corruption [Olken and Pande 2012], rather than on its broader impacts on performance (Andersen and Moynihan [2016] is a recent exception).

We contribute to this debate by studying the relationships between a spectrum of management practices and the full range of bureaucratic output, across 45 ministries and departments in the central government of Ghana. To measure management practices, we survey the universe of 3000 professional-grade civil servants in Ghana's central government, and construct measures of management quality, adapting the methodological innovations of Bloom and Van Reenen [2007] and Bloom *et al.* [2012] (henceforth BSVR) from organizational economics to the public sector. This enables us to construct indices of management style and quality related to monitoring/incentives, autonomy/discretion, and other practices, based not on subjective self-reported perceptions or specific *de jure* rules, but on probing and objectively benchmarked interviews that seek to capture the *de facto* management practices used in practice.

To measure output, we develop a novel approach exploiting the fact that each organization is required to provide quarterly and annual progress reports of their planned activities against their actual achievements. We collect, digitize, and hand-code these reports, yielding the characteristics and completion of 3620 outputs covering the entire range of bureaucratic activity, from procurement and infrastructure, to policy development, advocacy, human resource management, budgeting, and regulation.

We first document the high variation in management practices and task/project completion

across organizations. This is despite the fact they all operate under the same civil service laws, regulations and pay structure, are overseen by the same authorities, draw from the same pool of potential hires, and are located proximately to each other in the capital – sometimes even in the same building. This adds to a nascent body of research documenting within-country variation in bureaucratic quality, which has largely been undertaken through qualitative case studies of exceptional organizations [Leonard 2010], using input-based measures of bureaucratic capacity [Gingerich 2013, Bersch *et al.* 2016], or using subjective perception surveys [Owusu 2006].

We then estimate the relationships between management indices for monitoring/incentives and autonomy/discretion, and output completion. We exploit the fact that multiple organizations conduct each output type, so we estimate the partial correlation of each management practice with output completion conditional on output-type fixed effects. This accounts for heterogeneity in bureaucracies arising from the composition of outputs they are tasked to implement, as well as an extensive array of additional controls.

We find that both dimensions of management practices robustly correlate with output completion. However, they have opposing signs: a one standard deviation increase in management practices related to providing autonomy/discretion to bureaucrats is associated with a 25 percentage point (pp) increase in the likelihood work on an output is initiated and a 28pp increase in the likelihood it is fully completed; in contrast, a one standard deviation increase in management practices related to the provision of incentives/monitoring to bureaucrats is associated with a 28pp decrease in the likelihood an output is initiated, and a decrease of 18pp in the likelihood it is fully completed. These results are striking in the context of a lower-middle income country such as Ghana, where concerns about bureaucratic effectiveness and moral hazard among public officials are especially salient among academics and citizens alike.

Methodologically, this implies that, because the underlying autonomy/discretion and monitoring/incentives indices are positively correlated, estimating the relationships between a single set of management practices and output without accounting for other practices – as has often been done – leads to significant omitted variable bias. This justifies our approach of seeking to understand these relationships simultaneously across a spectrum of management practices, rather than focusing on the effects of a single dimension of management practice.

While these results are correlational, we probe them further by investigating their consistency with mechanisms from incentive theory through which incentives could have a negative effect on bureaucratic output: coordination, intrinsic motivation, and influence activities. We find evidence consistent with each mechanism. Although our context does not allow for clean causal identification of these effects, the empirical support we find for these mechanisms makes it unlikely that our findings are driven entirely by reverse causality. Our findings are thus an important complement to (quasi-)experimental studies in expanding our understanding of the relationship between management and performance in state bureaucracies.

Finally, to interrogate the external validity of our findings, we modify our core specification

to replicate the analysis of Rasul and Rogger [2018], who study the relationship between output and these two dimensions of management using comparable measurement approaches in Nigeria's Federal Civil Service. The findings are qualitatively and quantitatively similar across contexts, suggesting common mechanisms at play across bureaucracies in sub-Saharan African.

Section 2 describes our context and data. Section 3 describes the variation in management and productivity within Ghana's Civil Service. Section 4 presents our empirical method and main results. Sections 5 and 6 investigate mechanisms and external validity. Section 7 concludes.

2 Context and Data

Ghana is a lower-middle income state home to 28 million individuals, with a central government bureaucracy structured along lines reflecting its British colonial origins and various postindependence reforms. We study the universe of 45 Ministries and Departments in the Civil Service. The headquarters of these organizations are all in Accra, but have responsibility for public projects and activities implemented nationwide.¹ Ministries and Departments are overseen by the Office of the Head of Civil Service (OHCS). OHCS decides on all hiring, promotion, transfer, and (in rare circumstances) firing of bureaucrats. While OHCS develops and promulgates official management regulations and processes, Ministries' and Agencies' compliance with these is imperfect, with *de facto* management practices varying across organizations. All Ministries and Departments have the same statutory levels of autonomy and political oversight structures.

Our analysis of bureaucrats focuses on the professional grades of technical and administrative officers within these Ministries and Departments. We exclude grades covering cleaners, drivers, most secretaries, etc. On average, each organization employs 64 professional grade bureaucrats. Senior bureaucrats are those classified as a 'Director (Head of Division) or Acting Director' or as a 'Deputy Director or Unit Head (Acting or Substantive)'. By this definition, the span of control of senior bureaucrats over non-seniors is 4.5, with considerable variation across Ministries.²

45% of bureaucrats are women, 70% have a university degree, and 31% have a postgraduate degree. Bureaucrats enjoy stable employment once in service: the average bureaucrat has 14 years in service, with their average tenure in their current organization being 9 years, with infrequent transitions between organizations.

Our analysis is based on two data sources. First, we hand-coded quarterly and annual progress reports from Ministries and Departments, covering outputs ongoing between January and Decem-

¹Ghana distinguishes between the Civil Service and the broader Public Service, which includes autonomous agencies under the supervision of sector ministries, and frontline implementers such as the Police Service. Our sample is restricted to headquarter offices of Civil Service organizations.

 $^{^{2}}$ In Ghana, grades of technical and administrative bureaucrats are officially 'senior' officers while grades covering cleaners, drivers etc. are 'junior' officers, regardless of tenure/seniority. We restrict our sample to 'senior' officers in the formal terminology. However, we use the terms senior and non-senior in their colloquial sense to refer to hierarchical relationships within professional grades.

ber 2015. As detailed below, these reports enable us to code the individual outputs under the remit of each organization, and the extent to which they are initiated or successfully completed. Second, we surveyed 2971 bureaucrats from all 45 civil service organizations over the period August to October 2015. As detailed below, civil servants were questioned on topics including their background characteristics and work history in service, job characteristics and responsibilities, engagement with stakeholders outside the civil service, perceptions of corruption in the service, and their views on multiple dimensions of management practices.

2.1 Coding Output Completion

Worldwide, civil service bureaucracies differ greatly in whether and how they collect data on their performance, and few international standards exist to aid cross country comparisons. To therefore quantify the delivery of public sector outputs in our context, we exploit the fact that each Ghanaian civil service organization is required by OHCS to provide quarterly and annual progress reports. Organizations differ in their reporting formats and coverage, and some either did not produced reports for this time period or produced them in a format that was infeasible to code. We are able to use the progress reports of 30 Ministries and Departments (and our civil servant survey covers 2247 bureaucrats in these organizations). Figure A1 provides a snapshot of a typical progress report and indicates the information coded from it.

Progress reports cover the *entire* range of bureaucratic activity. While some of these outputs are public-facing, others are internal functions or intermediate outputs. We use these progress reports to identify 3620 tasks or projects ('outputs' henceforth) underway during 2015. The outputs undertaken by each organization in a given year are determined through an annual planning and budgeting process jointly determined between: the core executive, mainly the Ministry of Finance and the sector minister representing government priorities; the organization's management, based in large part on consultatively developed medium-term plans; and ongoing donor programs. This schedule of outputs is formalized in the organization's annual budget (approved by Parliament) and annual workplan. The quarterly and annual reports we use to code output completion thus detail the output that the organization's workplan committed the organization to work on during out study period.

The Appendix describes how we hand-coded and harmonized the information to measure output completion across the Ghanaian civil service. Three key points are of note in relation to this process. First, each quarterly progress report was codified into output line items using a team of trained research assistants and a team of civil servant officers seconded from the Management Services Department (MSD), an organization under OHCS tasked with analyzing and improving management in the civil service. MSD officers are trained in management and productivity analysis and frequently review organizational reports, making them ideally suited to judging output characteristics and completion. Second, coders were tasked to record output completion on a 1-5 scoring grid, where a score of one corresponds to, "No action was taken towards achieving the target", three corresponds to, "Some substantive progress was made towards achieving the target. The output is partially complete and/or important intermediate steps have been completed", and a score of five corresponds to, "The target for the output has been reached or surpassed." Outputs can be long-term or repeated (e.g. annual, quarterly) outputs. There were at least two coders per output.³

Third, as progress reports are self-compiled by bureaucracies, an obvious concern is that low performing bureaucracies might intentionally manipulate their reports to hide the fact. To check the validity of progress reports, we matched a sub-sample of 14% of outputs from progress reports to output audits conducted by external auditors through a separate exercise undertaken by OHCS. Auditors are mostly retired civil servants, overseen by OHCS, and they obtain documentary proof of output completion. For matched outputs, 94% of the completion levels we code are corroborated based on the qualitative descriptions of completion in audits.

The types of outputs included in the data are revealing of the full scope of activity of bureaucracies. The most common output type in Ghanaian central government bureaucracies relates to human resource management ('monitoring, training and personnel management', 29%), while 23% of outputs relate to policy advocacy and development. Infrastructure and procurement together comprise around a third of outputs, with other categories of bureaucratic activity comprising a relatively smaller share. Another feature that is important to our research design is that each output type is implemented by many different organizations and each organization implements multiple output types, allowing us to disentangle the performance of bureaucracies from the types of outputs they undertake.

2.2 Measuring Management

We follow BSVR's approach to measuring management practices, but adapt their procedures and survey to the Ghanaian public sector setting. Survey team leaders were recruited from the private sector, with an emphasis on previous experience of survey work in Ghana. We worked closely with team leaders to give them an appreciation and understanding of the practices and protocols of the civil service. We then collaborated with OHCS to recruit junior public officials to act as our enumerators. The Head of Service ensured their commitment to the survey process by stating the research team would monitor enumerator performance and that these assessments would influence future posting opportunities. We trained the team leaders and public officials jointly, including intensive practice interview sessions, before undertaking the first few interviews together.

Over the period from August to November 2015, our enumerators interviewed 2971 bureaucrats employed at 45 organizations. This constitutes 98% of all eligible staff in these organizations,

 $^{^{3}}$ Given the tendency for averaging scores across coders to reduce variation, we use the maximum and minimum scores to code whether outputs are fully complete/never initiated. We show robustness to alternative approaches to aggregating scores.

with the remainder mostly having been out of the office during the survey period. Interviews were conducted in person, but were double-blind in that interviewers had never worked in the organizations in which they were interviewing and did not know their interviewees, and likewise interviewees did not know their interviewers. As we elicited information on management practices in *individual* surveys administered to bureaucrats, both those in managerial roles and those being managed, we can construct management practices based on alternative sets of respondents. The measure we use for our core analysis averages management scores over the most senior divisional-bureaucrat reports. The median (mean) number of senior managers per organization is 13 (20).

We adapted BSVR's methodology to cover fourteen practices across six dimensions of management practice: roles, flexibility, incentives, monitoring, staffing and targets. Table A1 details each management related question and the scoring grid used by our enumerators for each question. Following BSVR, for each question enumerators would first ask what practices were used in an open-ended way, then probe respondents' responses and ask for examples to ascertain what practices are actually in use (much as a qualitative interview would), as opposed to simply asking for respondents' perceptions of management quality. Interviewers would then use this information to score each practice on a continuous 1-5 scale, where 1 represents non-use or inconsistent/incoherent use of that practice, and 5 represents strong, consistent, and coherent use of that practice. To further anchor the scores and provide comparability across organizations, the scoring grid for each practice is benchmarked to actual descriptions of the practices in use. This improves on more commonly used Likert-style measures of perceptions of management practices, which are vulnerable to differential anchoring across respondents and organizations.

The answers to questions on roles and flexibility are then combined to produce a measure of management practices related to autonomy/discretion. This index captures the extent to which organizations delegate decision-making, participation, innovation, and flexibility downwards, capturing both the absence of controls and the presence of structures to empower agents and support their use of their discretion. The answers to questions on incentives and monitoring scores are combined to produce an incentives/monitoring measure of management practices, which captures the extent to which organizations measure and review performance indicators and use performance to reward good performance and punish poor performance. The answers to questions on staffing and targeting topics are combined into a measure of 'other' management practices. The scores on each practice are converted into normalized z-scores by taking unweighted means of the underlying z-scores (so are continuous variables with mean zero and variance one by construction).

For the autonomy/discretion index, greater autonomy/discretion thus corresponds to higher scores, and for the incentives/monitoring measure the provision of stronger incentives/monitoring corresponds to higher scores. We do not presume that higher scores correspond to "good management" in the sense that they necessarily improve performance; rather, these relationships are what we try to estimate empirically.

3 Variation in Management and Output

We first provide novel descriptive evidence on within-country, cross-organization variation in output completion and in the management practices with which organizations are run. Figure 1A shows substantial variation in output completion across civil service organizations, whether measured in the proportion of outputs started/finished or in the average score on our 1-5 scale. To quantify this variation, we note that the 75th percentile organization has an average completion rate 22% higher than the 25th percentile organization. This output-based measure provides powerful evidence of the variation in actual bureaucratic performance across bureaucracies within a government, thus building on and extending the existing input-, survey-, and perception-based measures of this variation [Ingraham *et al.* 2003, Gingerich 2013, Bersch *et al.* 2016].

Figure 1B shows that there is also immense variation in the management practices bureaucrats are subject to across organizations along both dimensions of management practice. While organizations' scores on the two practices are positively correlated ($\rho = .67$), there is considerable variation, with some organizations scoring relatively higher on one measure than the other. Although quantitative comparison of these differences is complicated since our scaling is by necessity somewhat artificial, these quantitative scores are based on meaningful underlying qualitative variation. The differences across organizations are vast: the 75th percentile organization has a raw autonomy/discretion score that is 145% higher than 25th percentile organization, and the corresponding gap is 97% for the incentives/monitoring index.

Management thus differs across organizations in style, quality and the relative use of different dimensions of practices. While a mainly qualitative literature has previously documented variation in state capacity and performance within states, our findings represent large-scale evidence that within-government variation in management and performance is systematic and does not consist merely of a handful of problem organizations or "pockets of effectiveness" [Tendler 1997, Leonard 2010].

4 Empirical Method and Main Results

To examine how management practices relate to output completion, we estimate the following OLS specification where the unit of analysis is output i of type j in organization n:

$$y_{ijn} = \gamma_1 M P - autonomy_n + \gamma_2 M P - incentives_n + \gamma_3 M P - other_n + \beta_1 P C_{ijn} + \beta_2 O C_n + \lambda_j + \epsilon_{ijn}.$$
(1)

 y_{ijn} is an indicator of whether the output is initiated, or fully completed, or a continuous measure of the output completion rate. Management practices (*MP*-) are measured using the autonomy/discretion, incentives/monitoring and other indices, and PC_{ijn} and OC_n are output and

organizational controls.⁴ Many organizations implement the same output type j, so we can control for output-type fixed effects λ_j in (1), as well as fixed effects for the broad sector the implementing organization operates in.⁵ We cluster standard errors by organization n, the same level of variation as management practices.

The partial correlations of interest are γ_1 and γ_2 , the effect size of a one standard deviation change in management practices along the respective margins of autonomy and incentives/monitoring. These estimates are based on our measurement of *de facto* management practices in an organization. While management in an organization can be described qualitatively both by the style of management (what type of practices the organization is trying to implement) and the quality of implementation of these practices, our measure of management collapses both of these into a single dimension for each practice. We thus estimate the relationship between output and the management practices that organizations are *actually* using, rather than the management practices they are trying to use or an idealized version of them.

Table 1 presents our main results. To illustrate the omitted variable bias that occurs when analyzing one set of management practices in isolation, Column 1 first estimates the relationship between output and autonomy/discretion without controlling for organizations' other management practice, and Column 2 does the same for incentives/monitoring. Column 3 then presents our core specification using binary output completion as the dependent variable, while Columns 4 and 5 present the same specification with binary completion and a continuous completion index as outcomes respectively. A consistent set of findings emerges: (i) management practices providing bureaucrats more autonomy/discretion are robustly *positively* correlated with the likelihood of output completion ($\hat{\gamma}_1 > 0$); (ii) management practices related to the provision of incentives/monitoring to bureaucrats are robustly *negatively* correlated with the likelihood of output completion ($\hat{\gamma}_2 < 0$).

A one standard deviation increase in the autonomy/discretion index is associated with an increase in the likelihood an output is fully completed by 28pp; a one standard deviation increase in incentives/monitoring is associated with a decrease in the likelihood it is fully completed by 18pp. These magnitudes are of economic and statistical significance: the backdrop is that 21% of

⁴Output controls comprise controls for whether the output is regularly implemented by the organization or a one off, whether it is a bundle of interconnected outputs, and whether the division has to coordinate with actors external to government to implement the output. Organizational controls comprise a count of the number of interviews undertaken and organization-level controls are the share of the workforce with degrees, the share of the workforce with postgraduate qualifications, and the span of control. Following BVSR, we condition on 'noise' controls related to the management surveys: these are averages of indicators of the seniority, gender, and tenure of all respondents, the average time of day the interview was conducted and of the reliability of the information as coded by the interviewer.

⁵Output type fixed effects relate to whether the primary classification is listed as Advocacy and Policy Development, Financial and Budget Management, ICT Management and Research, Monitoring/Training/Personnel Management, Physical Infrastructure, Permits and Regulation, or Procurement. Sector fixed effects relate to whether the output is in the administration, environment, finance, infrastructure, security/diplomacy/justice or social sector.

outputs are never initiated, and only 34% are fully completed. The dramatically different point estimates between Columns 1-2 and Columns 3-5 illustrate that examining autonomy/discretion (incentives/monitoring) in isolation would lead to a downward (upward) bias on the point estimates, due to the underlying positive correlation between these measures.

The coefficient on each index represents a partial correlation of that set of practices with output. Thus, while we find that management practices related to incentives and monitoring are negatively related to output conditional on the level of autonomy and discretion being used in the same organization, this does not imply that all incentives and monitoring are bad for output. Rather, it implies that organizations seem to be overbalancing what we describe as their portfolio of management practices inefficiently towards incentives and monitoring at the expense of autonomy and discretion. While this finding is more nuanced than a simple "incentives and monitoring are bad", this reflects the multi-faceted nature of management.

The Appendix shows the results to be robust to alternative codings of completion rates, samples, estimation methods, fixed effects specifications, alternative clusterings of the standard errors, and introducing an extensive set of controls for the intrinsic motivation, personality traits, and cognitive skills of each organization's officials.

5 Mechanisms

Our core finding confirms the two dimensions of management practice emphasized by the public administration and economics literatures do indeed robustly correlate to effective public service delivery in the Ghanaian context. To further understand the factors driving these observed relationships, we unpack the negative relationship we find between incentives/monitoring and output. Incentive theory provides a rich set of predictions about the impacts of incentives in public sector contracting environments [Dixit 2002, Besley and Ghatak 2005, Finan *et al.* 2017], so we focus our analysis on this aspect of our results. Our data allows us to study three mechanisms in detail.

5.1 Coordination

Many bureaucratic outputs require coordination with stakeholders external to the organization such as the private sector, politicians, and community groups. We therefore examine whether the negative impact of incentives/monitoring is exacerbated in organizations where a greater share of their outputs require engagement with stakeholders outside of their organization, which likely increases coordination costs. Our survey asks about engagement with members of civil society, Ministers, Members of parliament, Member(s) of the local government units, the private sector, traditional authorities, community or religious group(s), and the media.

Column 1 of Table 2 interacts this variable measuring the frequency of interaction with external stakeholders with the incentives/monitoring index. We find the marginal effect of incentives/monitoring on output completion rates is even more negative in organizations where there is a greater need to coordinate with such external stakeholders. This is as expected if the baseline negative coefficient on incentives/monitoring were reflecting that top-down, rigid control mechanisms can backfire in public sector contexts characterized by the need for coordination with external actors.

5.2 Public Service Motivation

Incentives and monitoring might crowd out the intrinsic motivation of bureaucrats [Perry and Wise 1990, Benabou and Tirole 2006, Besley and Ghatak 2018]. A long established literature suggests those that self-select into public service might be relatively more intrinsically motivated than those in the private sector. Performance incentives or monitoring might then be detrimental if such practices crowd out intrinsic motivation. As part of our civil servants survey, we obtained individual measures of the public service motivation (PSM) of bureaucrats, using an abbreviated version of the standard Perry scale [Perry 1996]. This includes four sub-indices on motivation related to attraction to policy making, compassion for underprivileged people, commitment to the public interest, and self-sacrifice. These indices are often found to be positively associated with various measures of individual commitment, pro-social behavior, and performance [Belle 2013, Dal Bo *et al.* 2013]. We then examine how management practices related to incentives/monitoring interplay with these dimensions of intrinsic motivation.

The results are shown in Columns 2-5 of Table 2. Column 2 shows the negative relationship of incentives/monitoring with output is reinforced in organizations whose bureaucrats on average score more highly on the 'policy making' dimension of PSM (this sub-index measures intrinsic interest in the structures and procedures of policymaking). Columns 3 and 4 show that this negative relationship is partly ameliorated when bureaucrats in the organization score higher on the PSM dimensions of 'compassion' and 'public interest'. This runs counter to the notion that incentive provision crowds out efforts of intrinsically motivated individuals: if anything, as in Ashraf *et al.* [2014], our evidence suggests intrinsically motivated bureaucrats work harder in the face of poorly designed incentives/monitoring practices (so as to ameliorate the negative effects of these practices). However, in neither case does this interaction fully offset the negative overall coefficient on incentives/monitoring except for a small handful of bureaucrats at the extreme top end of the PSM distribution. Finally, Column 5 shows that the interaction of the 'self-sacrifice' dimension of PSM with management practices is not significant.

5.3 Influence Activities

The use of high-powered incentives and associated monitoring structures may distort bureaucratic agents' effort by encouraging them to engage in 'influence activities' to curry favor with their

managers [Milgrom 1988, Milgrom and Roberts 1988]. For example, agents may react to the provision of incentives not by improving performance towards the desired objective, but by engaging in lobbying, manipulating or distorting information revealed to managers, or other dysfunctional responses. There is scope for these distortions in public sector contexts where bureaucratic actions and outputs are difficult for managers to observe, leading to more reliance on subjective performance evaluation and thus a vulnerability to influence activities and bias.

We investigate whether these issues related to influence activities and subjective performance evaluation may be a factor in the observed relationship between incentives/monitoring and output by examining situations where the social connectedness between managers and their subordinates may exacerbate such distortions. We measure social connectedness in two ways: (i) the proportion of non-senior bureaucrats in an organization that overlapped in time at university as an undergraduate, with their most senior civil servant; (ii) the proportion of non-senior bureaucrats in an organization in the same ethnic group as their most senior civil servant. As Columns 6 and 7 of Table 2 show, we find a negative interaction between incentives/monitoring and both measures of social connectedness. This suggests these management practices might be capturing schemes in place that effectively allow for subjective performance evaluation or influence activities to take place between socially tied senior and non-senior bureaucrats.

5.4 Corruption

A final question relates to the interplay between management practices and corruption, a salient issue in public administration in much of the world. In particular, one concern about providing civil servants autonomy and discretion is that this might increase their engagement in corrupt behaviors. This concern perhaps motivates the policy emphasis on instituting top-down control and monitoring mechanisms. In 2015 Ghana scored at the 53rd percentile on the World Governance Indicators's Control of Corruption measure [World Bank 2018], making it a representative context in which to investigate these questions.

Our survey includes two simple approaches to eliciting perceptions of corruption from bureaucrats. First, we focus on corruption by senior bureaucrats by asking what proportion of unofficial payments are shared with the superior in the hypothetical scenario that, 'Imagine that a corrupt bureaucrat extracts unofficial payments. Typically in your organization, what proportion of the unofficial payments does s/he share with the following types or groups of people?' The second measure used a similar question structure, but asked officials the proportion of outputs on which they 'observed others breaking service rules for their own benefit'.

Table 3 shows that for neither measure of corruption do we find evidence that the interaction of autonomy/discretion with corruption is associated with a significantly lower likelihood of output completion. For the second measure, the negative association of incentives/monitoring with output is ameliorated if there is a greater proportion of recent outputs on which officials report observing others engaging in corrupt practices. This provides suggestive evidence that the negative consequences of corruption may be partially offset by the provision of incentives/monitoring to bureaucrats.

6 External Validity

Our findings in Ghana provide novel evidence about the relationships between management practices and bureaucratic output, and the mechanisms driving these relationships. A natural question is whether the same patterns exist in other countries. To investigate this we compare our findings from Ghana to work linking management practices for middle-tier bureaucrats and public sector output in the Federal Civil Service of Nigeria [Rasul and Rogger 2018]. In the Nigerian context, Rasul and Rogger [2018] coded independent engineering assessments of output completion rates for 4700 public projects and held focus group discussions of senior management in 63 organizations to generate BSVR-style management scores. While there are some differences in the scope of outputs included in the analysis and the approach to measuring management (as the Appendix details), there are nonetheless enough commonalities in measurement to allow us to examine the external validity of our findings from Ghana.

Table 4 repeats our core results for Ghana, alongside their equivalent measures in Nigeria. Columns 1 to 3 refer to Ghana and are the same specification as Table 1 Columns 3-5, except without sector fixed effects as the Nigeria data lacks this variable; Columns 4 to 6 refer to Nigeria. A common set of results emerge across contexts: (i) autonomy/discretion is robustly positively correlated with output initiation, full completion and completion rates; (ii) incentives/monitoring is robustly negatively correlated with output initiation, full completion and completion rates. Moreover, the estimates show similar effect sizes of both dimensions of management practice on the initiation and full completion margins, in which the two settings are most comparable. While our data in Ghana enables us to go significantly beyond Rasul and Rogger [2018] in investigating the mechanisms underlying these relationships, the similar core finding in each setting is suggestive of a pattern that is not limited only to Ghana. Establishing robust findings across similar contexts moves the knowledge frontier closer to establishing stylized facts.

An alternative approach to external validity is to use theory to predict in which type of contexts we would expect to find similar results. In this sense, lower-middle income countries like Ghana and Nigeria might be thought to be 'least likely' contexts for bureaucratic autonomy and discretion to be a positive force, given the salience of concerns about corruption and low overall government effectiveness. Similarly, if bureaucratic rigidity leads to incentives and monitoring backfiring, then contexts with high degrees of informality might also be thought to be least likely contexts to find negative associations between incentives/monitoring and output. Our examination of mechanisms also provides empirical guidance as to the type of contexts where incentives/monitoring are likely to be relatively more negative than autonomy/discretion as an approach to management: where coordination costs are high, where intrinsic motivation is driven by policymaking interest rather than compassion or public interest, and where there are high degrees of social connection between bureaucrats and their superiors. These predictions suggest promising avenues for further investigation in other contexts.

7 Conclusion

We investigate how two prominent approaches to public sector management – top-down control through monitoring and incentives, versus relying on bureaucratic autonomy and discretion – are related to bureaucratic output. We find robust positive partial correlations between output and organizational practices related to autonomy and discretion, but robust negative associations with management practices related to incentives and monitoring. We then provide evidence on three mechanisms that underpin these associations, show that bureaucratic autonomy and discretion does not seem to aggravate concerns over corruption, and provide evidence that our core findings may also pertain in other empirical contexts.

While we have demonstrated the robustness of our findings against an extensive range of organizational, individual, and output characteristics to rule out many alternative explanations, and shown that many of the mechanism driving this result are consistent with theoretical predictions, it is nonetheless possible that there exist additional unobserved factors that (partially) explain the observed associations. In this sense, we view this study as an important complement to a nascent body of experimental studies focused on high level bureaucrats as well as more nuanced qualitative research in advancing our knowledge over how management practices are related to output in public bureaucracies [Banerjee *et al.* 2014, Bandiera *et al.* 2019, Hjort *et al.* 2019, Rasul *et al.* 2019].

A Appendix

A.1 Measuring Bureaucratic Output

In Ghana each civil service organization is required to provide quarterly and annual progress reports. These detail targets and achievements for individual outputs. The process of measuring output for each organization then comprised two steps. First, extracting the data from organizations' reports (which differed slightly in their formats) into a standardized template. Second, coding variables based on the standardized data.

Figure A1 shows a snapshot of a typical quarterly progress report. The unit of observation is the task or output, defined as the most disaggregated output reported. For each quarterly progress report, we codified output line items using a team of trained research assistants and a team of civil servant officers seconded from the Management Services Department in the Civil Service. Each output was thus assigned to an organization (ministry or department) and to a division within that organization. For organizations in which reporting formats did not specify which division was responsible for a particular output, coders were supplied with information about the divisions in the organization and assigned each output to the division that was most likely responsible for it, in consultation with civil servants and/or research assistants who were familiar with the organization. In cases where the two coders assigned an output to different divisions, a manager made a judgment about which division to assign the output to.

A.2 Extracting and Standardizing

Although organizations' reports differed in their format and variable coverage, we extracted the following standard variables for each organization (leaving them blank where the variable was missing).

Output Level 1 The name or short description of the output specifying the action to be taken during the time period, at the most disaggregated or fine-grained level available. For instance, in Figure A1, this is 'Develop draft competition policy'. This variable defines the unit of observation, and by definition, cannot be missing.

Output Level 2 The name or short description of the output, aggregated to one level higher than in *Output level 1*. Many organizations reported outputs that were nested into broader outputs, or whose completion required multiple sequential or simultaneous smaller outputs to be completed. For example, in Figure A1 the *Output level 2* for 'Develop draft competition policy' is 'Competition Policy Developed and Approved.' Multiple outputs can thus share the same *Output level 2*.

Output Level 3 The same as Output level 2, but one level of aggregation higher. As in Figure A1, this level of aggregation was frequently unreported, but was extracted where relevant.

Budget Allocation/Cost The budgeted cost of the output. This was reported infrequently. **Baseline Completion Level** Where reported, the level of attainment on the output at the start of the time period.

Actual Output The actual attainment or work done during the time period. Together with the target level of achievement for the time period (from Output level 1) and (where relevant) the baseline level of completion, this is used to code output completion (as described in more detail below).

Remarks Where reported, the organization's comments about the output. These often explain why the target level of attainment was not achieved during the time period.

A.3 Coding

After extracting the data, our team of civil servants and output research assistants coded a fixed list of variables for each output (at the most disaggregated level, *Output level 1*). As the variables to be coded required coders to interpret and judge the information being reported by each organization, coding was undertaken by two independent coders, with reconciliation led by managers where necessary. Below is a list of all variables coded for each output.

Output Type (primary) Which category best describes this output? Coders had to select one of the following: (i) Advocacy, outreach and stakeholder engagement/relations; (ii) Financial & budget management; (iii) ICT management and/or development; (iv) Monitoring, review, & audit; (v) Permits and regulation; (vi) Personnel management; (vii) Physical infrastructure – office & facilities; (viii) Physical infrastructure – public infrastructure and projects; (ix) Policy development; (x) Procurement; (xi) Research; (xii) Training.

Output Type (secondary) If output covers more than one category, select the secondary category here. Coders had to select one of the same twelve categories as above.

Period/Regular vs. One-off *Is the output repeated (e.g. weekly, quarterly, annually) or one-off (no planned repetition)? Coders had to* select one of: (i) Periodic/ regular (e.g. weekly, quarterly, annually); (ii) One-off (no planned repetition).

Output Scope How narrowly is the output defined? Does it include multiple tasks, or even multiple outputs? Coders had to select one of: (i) Single activity (one step in a larger activity, has no value on its own; e.g. hold a meeting about writing a policy); (ii) Single output (multiple steps, has value on its own; e.g. write a policy); (iii) Bundle of outputs (multiple outputs that each have their own value; e.g. write four policies)].

Technical Complexity *Does the output require specific technical or scientific knowledge, beyond the level most civil servants would have?* Coders had to select one of: (i) No technical knowledge required (any senior civil servant could do this); (ii) Technical knowledge is required (special education or training needed).

Coordination Required *Does the division have to coordinate or interact with other actors in order to achieve the output?* Coders could select any of the following that applied: (i) Requires action from other divisions in the organization; (ii) Requires action from other government organizations; (iii) Requires action from stakeholders outside government.

Target Clarity *How precise, specific, and measurable is the target?* Coders had to answer on a 1-5 scale (where integers and half values were both permitted) using the following scoring guidelines. Score 1: Target is undefined or so vague it is impossible to assess what completion would mean; Score 3: Target is defined, but with some ambiguity; Score 5: There is no ambiguity over the target – it is precisely quantified or described.

Output Clarity *How precise, specific, and measurable is what the division actually achieved?* Coders had to answer on a 1-5 scale (where integers and half values were permitted) using the

following scoring guidelines. Score 1: Output information is absent or so vague it is impossible to assess completion; Score 3: Output information is given but there is some ambiguity over whether the target was met; Score 5: Output information is clear and unambiguous.

Completion Status *How did actual achievement compare to the target?* Coders had to answer on a 1-5 scale (where integers and half values were permitted) using the following scoring guidelines. Score 1: No action was taken towards achieving the target; Score 3: Some substantive progress was made towards achieving the target. The output is partially complete and/or important intermediate steps have been completed; Score 5: The target for the output has been reached or surpassed.

Completion Remarks Were any challenges/ obstacles mentioned? Coders could select all that applied from the following: (i) awaiting action from another division, organization or stakeholder; (ii) 2 = Procurement/sourcing delay or problem; (iii) Sequencing issue (can't start until another output has been completed); (iv) Lack of technical knowledge to complete activity; (v) Delayed/ non-release of funds; (vi) Unexpected event; (vii) Activity not due.

There are at least two coders per output. Given the tendency for averaging scores to reduce the measured variation, we use the maximum and minimum scores to code whether outputs are fully complete/never initiated respectively. We show robustness of our main result to alternative methods by which to combine codings.

A.3.1 Nigeria: Measuring Output Completion

The output data for Rasul and Rogger [2018], RR henceforth, is based on the Federal Civil Service in Nigeria, and exploited the following historical event. In 2006/7 the Nigerian Government undertook the Overview of Public Expenditure in NEEDS (the 'OPEN initiative'), in which it traced, by output, the use and impact of a representative sample of 10% of federal social sector expenditures approved in 2006/7. Under the OPEN initiative, expert teams visited projects to record their implementation. This monitoring process was independent of civil servants: projects were evaluated by teams of independent engineers and civil society.⁶ Monitoring teams visited project sites 18 months after projects were approved. The projects studied had 12 month completion schedules. The material from all projects recorded in OPEN initiative reports was hand-coded for 63 federal civil service organizations, covering 4721 outputs. 11 project types are covered (road, borehole, training etc.) with boreholes being the modal project type, and 75% of projects relating to small-scale infrastructure.

⁶A system of checks and balances were further put in place to underpin the credibility of the initiative. First, a centralized team of technocrats monitored the evaluation teams, providing them with training and opportunities for standardization of their methods. Second, evaluators were asked to provide material/photographic/video evidence to support their reports. Third, random checks were performed on evaluated sites.

A.3.2 Nigeria: Measuring Management Practices

In the Nigerian context, RR also followed BSVR's approach to measuring management practices, adapting their survey tool to the Nigerian public sector setting. Management practices were elicited from senior management in each organization. While each manager filled in their own questionnaire, enumerators looked for a consensus and recorded that in their own questionnaire. This is the information used to construct management practice indices for each organization. Enumerators held double blind interviews in the 63 organizations with output completion data. The management survey covered nine topics: roles, flexibility, incentives, monitoring, culture, targeting, facilities, skills and staffing. RR combined answers to the roles and flexibility questions to construct an index of management practices capturing bureaucrats' autonomy. The answers on the incentives and monitoring topic areas were combined to construct an index capturing the incentives/monitoring management practices bureaucrats operate under. All remaining topics were combined to construct an 'other' management practices index. The responses to each practice were converted into normalized z-scores by taking unweighted means of the underlying z-scores (so are continuous variables with mean zero and variance one by construction), where, as in Ghana, both are increasing in the commonly understood notion of 'better management'.

A.4 Robustness

Appendix Table A2 provides a battery of checks on our core estimates. These show the results to be robust to alternative codings of completion rates, samples, estimation methods, and fixed effects specifications. More precisely, we first redefine output completion rates to be the average of the codings of the two enumerators designated to each progress report. The result in Column 2 shows the baseline results continue to hold even when we reduce the variation in completion rates this way. Column 3 excludes outputs implemented by the largest organization in terms of number of outputs; Column 4 excludes the five smallest organizations by number of outputs. Columns 5 and 6 exclude organizations at the top and bottom of the autonomy/discretion and incentives/monitoring management scales respectively. Column 7 uses only the 30 organizations for whom we have coded output completion data for, to define the management z-scores (and so does not define the z-scores based on all 45 organizations for which these scores are available based on our civil servant survey). Column 8 reports the result of estimation a specification analogous to (1) but using a fractional regression to account for the fact that output completion rates lie between zero and one. Finally, in Column 9 we control for output-sector level fixed effects (so allowing for sector specific impacts of output types).

Appendix Table A3 shows the results to be robust to alternative clusterings of the standard errors, including robust standard errors, allowing them to be clustered by output type within organization (so at the jn level), by output type within sector, and by sector. In all cases, the coefficients of interest, $\hat{\gamma}_1$ and $\hat{\gamma}_2$, remain precisely estimated and statistically different from zero

at conventional significance levels (p < 0.05 throughout).

One concern in interpreting these results is that management practices may be endogenous to other variables that also affect management practices, so that these associations cannot be interpreted as causal. While the our extensive array of control variables address some such concerns, a remaining possibility is that organizations endogenously vary management practices according to the quality of their personnel. For example, organizations with highly capable or intrinsically motivated staff might rationally provide them with more autonomy and discretion, whereas organizations with low-quality or non-intrinsically motivated staff might rely more on the topdown control mechanisms of incentives and monitoring, and these characteristics of organizations' personnel may (partially) drive output.

To address this, in Appendix Table A4 we control for the cognitive skills, intrinsic motivation, and personality of personnel in each implementing organization, collected in our survey. These variables are averaged across each individual in the organization and include: scores on the Raven's test of cognitive ability; four indices of public service motivation (policymaking interest, compassion, public interest, and self-sacrifice) from an abbreviated version of the Perry [1996] scales; and scores on each component of the big-5 personality scales. Columns 1-3 introduce these three sets of variables in different combinations, and Column 4 introduces all three sets together. We estimate these four specifications separately for our core output measure (binary completion) in Panel A, and subsequently for our two alternative measures (binary initiation and continuous completion status) in Panels B and C.

The coefficient $\hat{\gamma}_1$ on autonomy/discretion is positive and significant in ten of the twelve specifications, positive and insignificant in one, and negative and significant in the other. The coefficient $\hat{\gamma}_2$ on incentives/monitoring remains negative and highly significant in all specifications. A practical challenge in estimating these extremely demanding specifications is multicollinearity among these characteristics, and between these organizational averages and other organization-level controls. This leads one to two noise controls to drop out in a handful of specifications. Despite this, the overall pattern is consistent with our core specifications and provides suggestive evidence that our results are not primarily driven by endogenous determination of management practices based on bureaucrat characteristics.

References

- [1] ACEMOGLU.D AND J.A.ROBINSON (2012) Why Nations Fail, Crown Business, New York.
- [2] ANDERSON.S.C AND D.P.MOYNIHAN (2016) "Bureaucratic Investments in Expertise: Evidence from a Randomized Controlled Field Trial," *Journal of Politics* 78: 1032-44.
- [3] ASHRAF.N, O.BANDIERA AND K.JACK (2014) "No Margin, No Mission? A Field Experiment on Incentives for Pro-social Tasks," *Journal of Public Economics* 120: 1-17.

- [4] BANDIERA.O, M.C.BEST, A.Q.KHAN AND A.PRAT (2019) Incentives and the Allocation of Authority in Organizations: A Field Experiment with Bureaucrats, mimeo LSE.
- [5] BANERJEE.A, R.CHATTOPADHYAY, E.DUFLO, D.KENISTON AND N.SINGH (2014) Improving Police Performance in Rajasthan, India: Experimental Evidence on Incentives, Managerial Autonomy and Training, NBER WP17912.
- [6] BELLE.N (2013) "Experimental Evidence on the Relationship Between Public Service Motivation and Job Performance," *Public Administration Review* 73: 143-53.
- [7] BENABOU.R AND J.TIROLE (2006) "Incentives and Prosocial Behavior," *American Economic Review* 96: 1652-78.
- [8] BERSCH, KATHERINE.S, 'E.PRAÇA AND M.TAYLOR (2016) "State Capacity, Bureaucratic Politicization, and Corruption in the Brazilian State," *Governance* 30: 105-24.
- [9] BESLEY.T AND M.GHATAK (2005) "Competition and Incentives with Motivated Agents," American Economic Review 95: 616-36.
- [10] BESLEY.T AND M.GHATAK (2018) "Prosocial Motivation and Incentives," Annual Review of Economics 10: 411-38.
- [11] BESLEY.T AND T.PERSSON (2011) Pillars of Prosperity, Princeton University Press, Princeton, New Jersey.
- [12] BLOOM.N AND J.VAN REENEN (2007) "Measuring and Explaining Management Practices Across Firms and Countries," *Quarterly Journal of Economics* 122: 1351-408.
- [13] BLOOM.N, R.SADUN AND J.VAN REENEN (2012) "The Organization of Firms Across Countries," *Quarterly Journal of Economics* 127: 1663-705.
- [14] DAL BO.E, F.FINAN AND M.ROSSI (2013) "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service," *Quarterly Journal of Economics* 128: 1169-218.
- [15] DIXIT.A (2002) "Incentives and Organizations in the Public Sector: An Interpretive Review," Journal of Human Resources 37: 696-727.
- [16] DUFLO.E, R.HANNA AND S.P.RYAN (2012) "Incentives Work: Getting Teachers to Come to School," American Economic Review 102: 1241-78.
- [17] EINSTEIN.K.L AND D.M.GLICK (2017) "Does Race Affect Access to Government Services? An Experiment Exploring Street-Level Bureaucrats and Access to Public Housing," American Journal of Political Science 61: 100-16.

- [18] FINAN.F, B.A.OLKEN, AND R.PANDE (2017) "The Personnel Economics of the State," in A.V.Banerjee and E.(eds.) *Handbook of Field Experiments*, Elsevier.
- [19] GINGERICH.D (2013) "Governance Indicators and the Level of Analysis Problem: Empirical Findings from South America," *British Journal of Political Science* 43: 505-40.
- [20] HASNAIN.Z, N.MANNING AND J.H.PIERSKALLA (2012) Performance-related Pay in the Public Sector, World Bank Policy Research WP6043.
- [21] HJORT.J, D.OREIRA, G.RAO AND J.F.SANTINI (2019) How Research Affects Policy: Experimental Evidence from 2,150 Brazilian Municipalities, NBER WP25941.
- [22] INGRAHAM.P.W, P.G.JOYCE AND A.K.DONAHUE (2003) Government Performance: Why Management Matters, London: Johns Hopkins University Press.
- [23] KHAN.A.Q, A.I.KHWAJA AND B.OLKEN (2015) "Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors," *Quarterly Journal of Economics* 131: 219-71.
- [24] LEONARD.D.K (2010) "Pockets' of Effective Agencies in Weak Governance States: Where are they Likely and Why does it Matter?," *Public Administration and Development* 30: 91-101.
- [25] LYNN.L, C.HEINRICH AND C.HILL (2000) "Studying Governance and Public Management: Challenges and Prospects," *Journal of Public Administration Research and Theory* 10: 233-61.
- [26] MEIER.K.J, AND L.J. O'TOOLE JR (2002) "Public Management and Organizational Performance: The Effect of Managerial Quality," *Journal of Policy Analysis and Management* 21: 629-43.
- [27] MEIER.K.J, AND L.J. O'TOOLE JR (2012) "Subjective Organizational Performance and Measurement Error: Common Source Bias and Spurious Relationships," *Journal of Public Administration Research and Theory* 23: 429-56.
- [28] MILGROM.P.R (1988) "Employment Contracts, Influence Activities, and Efficient Organization Design", Journal of Political Economy 96: 42-60.
- [29] MILGROM.P.R AND J.ROBERTS (1988) "An Economic Approach to Influence in Organizations", American Journal of Sociology 94: S154-79.
- [30] MILLER.G AND A.B.WHITFORD (2016) Above Politics: Bureaucratic Discretion and Credible Commitment, Cambridge: Cambridge University Press.
- [31] OLKEN.B AND R.PANDE (2012) "Corruption in Developing Countries," Annual Review of Economics 4: 479-509.

- [32] OWUSU.F (2006) "Differences in the Performance of Public Organisations in Ghana: Implications for Public-Sector Reform Policy," *Development Policy Review* 24: 693-705.
- [33] PEPINSKY.T.B, J.H.PIERSKALLA AND A.SACKS (2017) "Bureaucracy and Service Delivery," Annual Review of Political Science 20: 13.1-13.20.
- [34] PERRY.J.L (1996) "Measuring Public Service Motivation: An Assessment of Construct Reliability and Validity," Journal of Public Administration Research and Theory 6: 5-22.
- [35] PERRY.J.L AND L.R.WISE (1990) "The Motivational Basis of Public Service," *Public Admin*istration Review 50: 367-73.
- [36] RASULI AND D.ROGGER (2018) "Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service," *Economic Journal* 128: 413-46.
- [37] RASULI, M.J.WILLIAMS AND D.ROGGER (2019) Training for Productivity? Experimental Evidence from Ghana's Civil Service, mimeo UCL.
- [38] ROSE-ACKERMAN.S (1986) "Reforming Public Bureaucracy Through Economic Incentives?," Journal of Law, Economics and Organization 2: 131-61.
- [39] SIMON.W (1983) "Legality, Bureaucracy, and Class in the Welfare System," Yale Law Journal 92: 1198-269.
- [40] TENDLER.J (1997) Good Government in the Tropics, Baltimore: Johns Hopkins University Press.
- [41] WILSON.J (1989) Bureaucracy: What Government Agencies Do and Why They Do It, New York: Basic Books.
- [42] WORLD BANK (2018) World Governance Indicators, available at: http://info.worldbank.org/governance/WGI/#reports.



Figure 1A: Output Completion by Organization

Notes: Multiple coders assessed an output such that here we take the minimum assessment of initiation and the maximum assessment of completion, so that it is possible for proportion started to be lower than proportion completed (as it is for one organization). Completion status is a continuous 1-5 score for each output, here rescaled to 0-1.



Figure 1B: Management Practices by Organization

Notes: Organization z-scores presented for organizations with output data available.

Table 1: Management of Bureaucrats and Public Service Delivery

Standard Errors: Clustered by Organization OLS Estimates

Dependent Variable:	(1) Output completion [binary]	(2) Output completion [binary]	(3) Output completion [binary]	(4) Output initiation [binary]	(5) Completion rate [0-1 continuous]
Management-Autonomy/Discretion	0.10***		0.28***	0.25***	0.18***
	(0.02)		(0.04)	(0.04)	(0.03)
Management-Incentives/Monitoring		0.04	-0.18***	-0.28***	-0.16***
		(0.03)	(0.05)	(0.04)	(0.03)
Management-Other			-0.07***	-0.09***	-0.04***
			(0.02)	(0.02)	(0.01)
Noise Controls	Yes	Yes	Yes	Yes	Yes
Organizational Controls	Yes	Yes	Yes	Yes	Yes
Output Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Output Type, Sector				
Observations (clusters)	3620 (30)	3620 (30)	3620 (30)	3620 (30)	3620 (30)

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. Standard errors are in parentheses, and are clustered by organization throughout. All Columns report OLS estimates. The dependent variable in Columns 1 to 3 is a dummy variable that takes the value 1 if the project is completed and 0 otherwise. The dependent variable in Column 4 is a dummy variable that takes the value 1 if the project completion (that is a continuous measure between zero and one). Output controls comprise controls for whether the output is regularly implemented by the organization or a one off, whether it is a bundle of interconnected outputs, and whether the division has to coordinate with actors external to government to implement the output. Organizational controls comprise a count of the number of interviews undertaken and organization-level controls are the share of the workforce with degrees, the share of the workforce with postgraduate qualifications, and the span of control. Following BVSR, we condition on `noise' controls related to the management surveys: these are averages of indicators of the seniority, gender, and tenure of all respondents, the average time of day the interview was conducted and of the reliability of the information as coded by the interviewer. Output type fixed effects relate to whether the primary classification is listed as Advocacy and Policy Development, Financial and Budget Management, ICT Management and Research, Monitoring/Training/Personnel Management, Physical Infrastructure, Permits and Regulation, or Procurement. Sector fixed effects relate to whether the output is in the administration, environment, finance, infrastructure, security/diplomacy/justice or social sector. Figures are rounded to two decimal places.

Table 2: Mechanisms

Dependent Variable: Project Fully Completed Standard Errors: Clustered by Organization

OLS Estimates

	Coordination		Influence Activities				
Interaction Variable:	(1) Stakeholder engagement	(2) Policymaking	(3) Compassion	(4) Public Interest	(5) Self-sacrifice	(6) Shared University Time	(7) Same Ethnicity
Management-Autonomy/Discretion	0.28***	0.28***	0.34***	0.34***	0.28***	0.28***	0.28***
	(0.04)	(0.04)	(0.04)	(0.06)	(0.06)	(0.03)	(0.03)
Management-Incentives/Monitoring	-0.18***	-0.14***	-0.19***	-0.20***	-0.16***	-0.22***	-0.10***
	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.04)	(0.04)
Management-Other	-0.07***	-0.09***	-0.11***	-0.11***	-0.08***	-0.06***	-0.09***
	(0.02)	(0.03)	(0.01)	(0.03)	(0.03)	(0.01)	(0.03)
Interaction Variable							
Main effect	0.05**	0.02	-0.08***	-0.09***	-0.04***	0.10***	0.10***
	(0.02)	(0.02)	(0.01)	(0.03)	(0.01)	(0.02)	(0.02)
Interaction with Management-	-0.06**	-0.05**	0.08***	0.12***	0.04	-0.11***	-0.10***
Incentives/Monitoring	(0.02)	(0.02)	(0.02)	(0.04)	(0.04)	(0.02)	(0.03)
Noise and Organizational Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Output Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Output Type, Sector	Output Type, Sector	Output Type, Sector	Output Type, Sector	Output Type, Sector	Output Type, Sector	Output Type, Sector
Observations (clusters)	3620 (30)	3620 (30)	3620 (30)	3620 (30)	3620 (30)	3620 (30)	3620 (30)

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. Standard errors are in parentheses, and are clustered by organization throughout. All columns report OLS estimates. The dependent variable in all columns is a dummy variable that takes the value 1 if the output is completed and 0 otherwise. The variable with which management scores are interacted is a z-score defined across the 30 organizations we study. In Column 1, 'Stakeholder Engagement' is the proportion of outputs on which officials at an organization report that they interact with stakeholders outside of their organization. In Columns 2-5, the variables represent z-scores of organization averages of aggregate Perry Public Service Motivation scores on the sub-categories of attraction to policy making, compassion for underprivileged people, commitment to the public interest, and self-sacrifice respectively. In Column 6, 'Shared University' refers to the proportion of officials in an organization who share an ethnicity with the senior civil servant in the unit. In Column 7, 'Same Ethnicity' refers to the proportion of officials in an organization-level controls are the division has to coordinate with actors external to government to implement the output. Organizational controls comprise a count of the number of interviews undertaken and organization-level controls are the share of the workforce with degrees, the share of the workforce with postgraduate qualifications, and the span of control. Following BVSR, we condition on `noise' controls related to the management surveys: these are averages of indicators of the seniority, gender, and tenure of all respondents, the average time of day the interview was conducted and of the reliability of the information as coded by the interviewe. Output type fixed effects relate to whether the primary classification is listed as Advocacy and Policy Development, Financial and Budget Management, ICT Management and Research, Monitoring/Training/Personnel Management, Physical Infrastructure, Permits and Reg

Table 3: Management Practices and Corruption

Dependent Variable: Project Fully Completed Standard Errors: Clustered by Organization OLS Estimates

	(1) Rents Received by Superior	(2) Observe Others' Corruption
Management-Autonomy/Discretion	0.28***	0.33***
	(0.06)	(0.04)
Management-Incentives/Monitoring	-0.13**	-0.20***
	(0.07)	(0.07)
Management-Other	-0.07***	-0.14***
	(0.02)	(0.03)
Interaction Variable		
Main effect	0.05**	0.04
	(0.02)	(0.03)
Interaction with Management-	-0.01	0.00
Autonomy/Discretion	(0.03)	(0.02)
Interaction with Management-	0.02	0.11**
Incentives/Monitoring	(0.05)	(0.04)
Noise and Organizational Controls	Yes	Yes
Output Controls	Yes	Yes
Fixed Effects	Output Type, Sector	Output Type, Sector
Observations (clusters)	3620 (30)	3620 (30)

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. Standard errors are in parentheses, and are clustered by organization throughout. All columns report OLS estimates. The dependent variable in all columns is a dummy variable that takes the value 1 if the output is completed and 0 otherwise. The variable with which management scores are interacted is z-score defined across the 30 organizations we study. In Column 1, 'Rents Received by Superior' represents the average proportion of corruption rents received by a unit superior within an organization. In Column 2, 'Other Breaking' refers to the proportion of recent outputs on which officials report observing others engaging in corrupt practices. Output controls comprise controls for whether the output is regularly implemented by the organization or a one off, whether it is a bundle of interconnected outputs, and whether the division has to coordinate with actors external to government to implement the output. Organizational controls comprise a count of the number of interviews undertaken and organization-level controls are the share of the workforce with degrees, the share of the workforce with postgraduate qualifications, and the span of control. Following BVSR, we condition on `noise' controls related to the management surveys: these are averages of indicators of the seniority, gender, and tenure of all respondents, the average time of day the interview was conducted and of the reliability of the information as coded by the interviewer. Output type fixed effects relate to whether the primary classification is listed as Advocacy and Policy Development, Financial and Budget Management, ICT Management and Research, Monitoring/Training/Personnel Management, Physical Infrastructure, Permits and Regulation, or Procurement. Sector fixed effects relate to whether the output is in the administration, environment, finance, infrastructure, security/diplomacy/justice or social sector. Figures are rounded to two decimal places.

Table 4: External Validity

Standard Errors: Clustered by Organization

OLS Estimates

	Ghana			Nigeria		
	(1) Output completion [binary]	(2) Output initiation [binary]	(3) Completion rate [0-1 continuous]	(4) Output completion [binary]	(5) Output initiation [binary]	(6) Completion rate [0-1 continuous]
Management-Autonomy/Discretion	0.23***	0.13***	0.12***	0.16***	0.15***	0.18***
	(0.03)	(0.05)	(0.02)	(0.02)	(0.03)	(0.03)
Management-Incentives/Monitoring	-0.11**	-0.16***	-0.10***	-0.10***	-0.16***	-0.14***
	(0.06)	(0.06)	(0.04)	(0.02)	(0.02)	(0.02)
Management-Other	-0.09***	-0.07***	-0.04**	0.06**	0.06**	0.08***
	(0.02)	(0.02)	(0.01)	(0.03)	(0.03)	(0.02)
Noise and Organizational Controls	Yes	Yes	Yes	Yes	Yes	Yes
Output Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Output Type	Output Type	Output Type	Output Type	Output Type	Output Type
Observations (clusters)	3620 (30)	3620 (30)	3620 (30)	4721 (201)	4721 (201)	4721 (201)

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. Standard errors are in parentheses, and are clustered by organization throughout. All Columns report OLS estimates. The dependent variable in Columns 1 and 4 is a dummy variable that takes the value 1 if the output is initiated and 0 otherwise. The dependent variable in Columns 2 and 5 is a dummy variable that takes the value 1 if the output is completed and 0 otherwise. The dependent variable in Columns 3 and 6 is an index of output completion that is a continuous measure between zero and one. For details of controls and fixed effects, see text (for Ghana) and appendix (for Nigeria). Figures are rounded to two decimal places.

Management Practice	nagement Practice Topic Indicative Question		Score 1	Score 3	Score 5	
Autonomy/Discretion	Roles	Can most senior staff in your division make substantive contributions to the policy formulation and implementation process?	Senior staff do not have channels to make substantive contributions to organizational policies, nor to the management of their implementation.	Substantive contributions can be made in staff meetings by all senior staff but there are no individual channels for ideas to flow up the organization.	It is integral to the organization's culture that any member of senior staff can substantively contribute to the policies of the organization or their implementation.	
		When senior staff in your division are given tasks in their daily work, how much discretion do they have to carry out their assignments? Can you give me an example?	Officers in this division have no real independence to make decisions over how they carry out their daily assignments. Their activities are defined in detail by senior colleagues or organizational guidelines.	Officers in this division have some independence as to how they work, but strong guidance from senior colleagues, or from rules and regulations.	Officers in this division have a lot of independence as to how they go about their daily duties.	
		Is the burden of achieving your division's targets evenly distributed across its different officers, or do some individuals consistently shoulder a greater burden than others?	A small minority of staff undertake the vast majority of substantive work within the division.	A majority of staff make valuable inputs, but it is by no means everyone who pulls their weight.	Each member of the division provides an equally valuable contribution, working where they can provide their highest value.	
		Would you say that senior staff try to use the right staff for the right job?	Often tasks are not staffed by the appropriate staff. Staff are allocated to tasks either randomly, or for reasons that are not associated with productivity.	Most jobs have the right staff on them, but there are organizational constraints that limit the extent to which effective matching happens.	The right staff are always used for a task.	
	Flexibility	Does your division make efforts to adjust to the specific needs and peculiarities of communities, clients, or other stakeholders?	The division uses the same procedures no matter what. In the face of specific needs or community/ client peculiarities, it does not try to develop a 'better fit' but automatically uses the default procedures.	The division makes steps towards responding to specific needs and peculiarities, but stumbles if the specific needs are complex. Often, tailoring of services is often unsuccessful.	The division always redefines its procedures to respond to the needs of communities/ clients. It does its best to serve each individual need as best as it can.	
		How flexible would you say your division is in terms of responding to new and improved work practices?	There is no effort to incorporate new ideas or practices. When practice improvements do happen, there is no effort to disseminate them through the division.	New ideas or practices are sometimes adopted but in an ad hoc way. These are sometimes shared informally or in a limited way, but the division does not actively encourage this or monitor their adoption.	Seeking out and adopting improved work practices is an integral part of the division's work. Improvements are systematically disseminated throughout the division and their adoption is monitored.	

Table A1: Defining Management Practices

Management Practice	ctice Topic Indicative Question		Score 1	Score 3	Score 5	
Given p ncentives/Monitoring Performance would u Incentives tolerate		Given past experience, how would under-performance be tolerated in your division? Poor performance is not addressed or is inconsister addressed. Poor performer rarely suffer consequences are removed from their positions.		Poor performance is addressed, but on an ad hoc basis. Use of intermediate interventions, such as training, is inconsistent. Poor performers are sometimes removed from their positions under conditions of repeated poor performance.	Repeated poor performance is systematically addressed, beginning with targeted intermediate interventions. Persistently poor performers are moved to less critical roles or out of the organization.	
		Given past experience, are members of [respondent's organization] disciplined for breaking the rules of the civil service?	Breaking the rules of the civil service does not carry any consequences in this division. Guilty parties do not receive the stipulated punishment.	An officer may break the rules infrequently and not be punished. An officer who regularly breaks the rules may be disciplined, but there would be no other specific actions beyond this. The underlying drivers of the behavior can persist indefinitely.	Any officer who breaks the rules of the civil service is punished; the underlying driver is identified and rectified. On- going efforts are made to ensure the issue does not arise again.	
		Does your division use performance, targets, or indicators for tracking and rewarding (financially or non- financially) the performance of its officers?	Officers in the division are rewarded (or not rewarded) in the same way irrespective of their performance.	The evaluation system awards good performance in principle (financially or non-financially), but awards are not based on clear criteria/processes.	The evaluation system rewards individuals (financially or non-financially) based on performance. Rewards are given as a consequence of well-defined and monitored individual achievements.	
	Monitoring	In what kind of ways does your division track how well it is delivering services? Can you give me an example?	Measures tracked are not appropriate or do not indicate directly if overall objectives are being met. Tracking is an ad hoc process and most processes aren't tracked at all. Tracking is dominated by the head of the division.	Performance indicators have been specified but may not be relevant to the division's objectives. The division has inclusive staff meetings where staff discuss how they are doing as division.	Performance is continuously tracked, both formally with key performance indicators and informally, using appropriate indicators and including many of the divisional staff.	

Table A1 Continued: Defining Management Practices

Management Practice	agement Practice Topic Indicative Question		Score 1	Score 3	Score 5
Other	Staffing	Do you think about attracting talented people to your division and then doing your best to keep them? For example, by ensuring they are happy and engaged with their work.	Attracting, retaining and developing talent throughout the division is not a priority or is not possible given service rules.	Having top talent throughout the division is seen to be a key way to effectively deliver on the organizations mandate but there is no strategy to identify, attract or train such talent.	The division actively identifies and acts to attract talented people who will enrich the division. They then develop those individuals for the benefit of the division and try to retain their services.
		If two senior level staff joined your division five years ago and one was much better at their work than the other, would he/she be promoted through the service faster?	The division promotes people by tenure only, and thus performance does not play a role in promotion.	There is some scope for high performers to move up through the service faster than non- performers in this division, but the process is gradual and vulnerable to inefficiencies.	The division would certainly promote the high-performer faster, and would rapidly move them to a senior position to capitalize on their skills.
	Targeting	Does your division have a clear set of targets derived from the organization's goals and objectives? Are they used to determine your work schedule?	The division's targets are very loosely defined or not defined at all; if they exist, they are rarely used to determine our work schedule and our activities are based on ad hoc directives from senior management.	Targets are defined for the division and its individual officers (managers and staff). However, their use is relatively ad hoc and many of the division's activities do not relate to those targets.	Targets are defined for the division and individuals (managers and staff) and they provide a clear guide to the division and its staff as to what the division should do. They are frequently discussed and used to benchmark performance.
		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?	No. There is a general level of confusion as to what the organization is trying to achieve on a daily basis and what individual's roles are towards those goals.	To some extent, or at least on some days. The organization's main goals and individual's roles to achieve them are relatively clear, but it is sometimes difficult to see how current activities are moving us towards those.	Yes. It is always clear to the body of staff what the organization is aiming to achieve with the days activities and what individual's roles and responsibilities are towards that.

Table A2: Robustness

Dependent Variable: Project Fully Completed Standard Errors: Clustered by Organization OLS Estimates

	(1) Baseline	(2) Average Completion Rate	(3) Excl. Org. With Most Outputs	(4) Excl. Five Orgs. With Smallest No. of Outputs	(5) Excl. Autonomy/ Discretion Outliers	(6) Excl. Incentives/ Monitoring Outliers	(7) Defining Management Scores Using Orgs with Output Data	(8) Fractional regression	(9) Alternative Fixed Effects
Management-Autonomy/Discretion	0.28***	0.12***	0.24***	0.30***	0.35***	0.30***	0.26***	1.25***	0.26***
	(0.04)	(0.03)	(0.04)	(0.04)	(0.02)	(0.04)	(0.04)	(0.19)	(0.04)
Management-Incentives/Monitoring	-0.18***	-0.11***	-0.19***	-0.23***	-0.21***	-0.21***	-0.14***	-0.80***	-0.16***
	(0.05)	(0.03)	(0.05)	(0.05)	(0.03)	(0.05)	(0.04)	(0.22)	(0.04)
Management-Other	-0.07***	-0.03**	-0.08***	-0.06***	-0.06***	-0.07***	-0.06***	-0.35***	-0.08***
	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.09)	(0.02)
Noise Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Organizational Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Output Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Output Type, Sector	Output Type, Sector	Output Type, Sector	Output Type, Sector	Output Type, Sector	Output Type, Sector	Output Type, Sector	Output Type, Sector	Output Type Within Sector
Observations (clusters)	3620 (30)	3620 (30)	3125 (29)	3593 (26)	3379 (27)	3585 (29)	3620 (30)	3620 (30)	3620 (30)

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. Standard errors are in parentheses, and are clustered by organization throughout. Columns 1 to 7 and 9 report OLS estimates. Column 8 reports estimates from a fractional regression model. In Columns 1 and 3 through 9, the dependent variable is a dummy variable that takes the value 1 if the output is fully completed and 0 otherwise. In Column 2, we redefine output completion rates to be the average of the codings of the two enumerators designated to each progress report. Column 3 excludes outputs implemented by the largest organization in terms of number of outputs. Column 4 removes the 5 smallest organizations by number of outputs. Columns 5 and 6 exclude organizations at the top and bottom of the Autonomy/Discretion and Incentives/Monitoring management scales respectively. Column 7 uses the 30 organizations with non-missing output data to define the management z-scores. In Column 9 we control for output-sector level fixed effects (so allowing for sector specific impacts of output types). Output controls comprise controls for whether the output is regularly implemented by the organization or a one off, whether it is a bundle of interconnected outputs, and whether the division has to coordinate with actors external to government to implement the output. Organizational controls comprise a count of the number of interviews undertaken and organization-level controls are the share of the workforce with postgraduate qualifications, and the span of control. Following BVSR, we condition on `noise' controls related to the management, environe. Output type fixed effects under the management, provide the interview was conducted and of the reliability of the information as coded by the interviewer. Output type fixed effects relate to whether the primary classification is listed as Advocacy and Policy Development, Financial and Budget Management, ICT Management and Research, Monitoring/Training/Personnel Management, Physical Infrastructure, Permits and Reg

Table A3: Alternative Clustering

Dependent Variable: Project Fully Completed OLS Estimates

Clustering Level:	(1) Baseline: Organization	(2) Robust	(3) Output Type Within Organization	(4) Output Type Within Sector	(5) Sector
Management-Autonomy/Discretion	0.28***	0.28***	0.28***	0.28***	0.28***
	(0.04)	(0.05)	(0.05)	(0.06)	(0.05)
Management-Incentives/Monitoring	-0.18***	-0.18***	-0.18***	-0.18**	-0.18***
	(0.05)	(0.05)	(0.07)	(0.07)	(0.06)
Management-Other	-0.07***	-0.07***	-0.07**	-0.07**	-0.07***
	(0.02)	(0.02)	(0.04)	(0.04)	(0.02)
Noise Controls	Yes	Yes	Yes	Yes	Yes
Organizational Controls	Yes	Yes	Yes	Yes	Yes
Output Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Output Type, Sector	Output Type, Sector	Output Type, Sector	Output Type, Sector	Output Type, Sector
Observations (clusters)	3620 (30)	3620	3620 (167)	3620 (41)	3620 (6)

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. Standard errors are in parentheses, and are clustered by organization in Column 1, by output type within organization in Column 3, by output type within sector in Column 4, and by sector in Column 5. In Column 2, robust standard errors are reported. All columns report OLS estimates. The dependent variable is a dummy variable that takes the value 1 if the output is fully completed and 0 otherwise. Output controls comprise controls for whether the output is regularly implemented by the organization or a one off, whether it is a bundle of interconnected outputs, and whether the division has to coordinate with actors external to government to implement the output. Organizational controls comprise a count of the number of interviews undertaken and organization-level controls are the share of the workforce with degrees, the share of the workforce with postgraduate qualifications, and the span of control. Following BVSR, we condition on `noise' controls related to the management surveys: these are averages of indicators of the seniority, gender, and tenure of all respondents, the average time of day the interview was conducted and of the reliability of the information as coded by the interviewer. Output type fixed effects relate to whether the primary classification is listed as Advocacy and Policy Development, Financial and Budget Management, ICT Management and Research, Monitoring/Training/Personnel Management, Physical Infrastructure, security/diplomacy/justice or social sector. Figures are rounded to two decimal places.

Table A4: Bureaucrat Characteristics

Standard Errors: Clustered by Organization OLS Estimates

Bureaucrat controls	(1) PSM + Personality	(2) PSM + Cognitive	(3) Personality + Cognitive	(4) PSM + Personality + Cognitive
Par	el A: Output C	ompletion [bi	nary]	
Management-Autonomy/Discretion	0.52***	0.31***	0.64***	0.26***
	(0.03)	(0.05)	(0.03)	(0.05)
Management-Incentives/Monitoring	-0.39***	-0.17***	-0.56***	-0.18***
	(0.04)	(0.04)	(0.03)	(0.03)
Management-Other	-0.10***	-0.08**	-0.04**	-0.09***
	(0.02)	(0.04)	(0.02)	(0.02)
<u>Pa</u>	nel B: Output	Initiation [bin	ary]	
Management-Autonomy/Discretion	0.43***	0.14*	0.52***	-0.10***
	(0.02)	(0.07)	(0.05)	(0.03)
Management-Incentives/Monitoring	-0.58***	-0.27***	-0.58***	-0.13***
	(0.03)	(0.04)	(0.04)	(0.02)
Management-Other	-0.03**	0.06	-0.06***	-0.01
	(0.01)	(0.09)	(0.02)	(0.02)
Panel C	: Completion S	itatus [0-1 coi	ntinuous]	
Management-Autonomy/Discretion	0.30***	0.18***	0.40***	0.03
	(0.02)	(0.04)	(0.02)	(0.02)
Management-Incentives/Monitoring	-0.30***	-0.16***	-0.39***	-0.06***
	(0.02)	(0.02)	(0.02)	(0.01)
Management-Other	-0.06***	-0.02	-0.03**	-0.05***
	(0.01)	(0.04)	(0.01)	(0.01)
Public Service Motivation Controls	Yes	Yes	No	Yes
Personality Traits Controls	Yes	No	Yes	Yes
Cognitive Skills Controls	No	Yes	Yes	Yes
Noise, Organizational, Output Controls	Yes	Yes	Yes	Yes
Fixed Effects	Output Type, Sector	Output Type, Sector	Output Type, Sector	Output Type, Sector
Observations (clusters)	3620 (30)	3620 (30)	3620 (30)	3620 (30)

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. Standard errors are in parentheses, and are clustered by organization throughout. All columns report OLS estimates. Dependent variables are output completion (binary) for Panel A, output initiation (binary) for Panel B, and output completion (continuous) for Panel C. Public service motivation controls include z-scores of organization averages on the four PSM sub-indices (policymaking, self-sacrifice, compassion, and public interest); personality trait controls include z-scores of organization averages on the components of the Big 5 scale (conscientiousness, agreeableness, openness, extraversion, and neuroticism); cognitive skills control is the z-score of the organization average score on a Raven's matrix quiz. Output controls comprise controls for whether the output is regularly implemented by the organization or a one off, whether it is a bundle of interconnected outputs, and whether the division has to coordinate with actors external to government to implement the output. Organizational controls comprise a count of the number of interviews undertaken and organization-level controls are the share of the workforce with degrees, the share of the workforce with postgraduate qualifications, and the span of control. Following BVSR, we condition on `noise' controls related to the management surveys: these are averages of indicators of the seniority, gender, and tenure of all respondents, the average time of day the interview was conducted and of the reliability of the information as coded by the interviewer. Output type fixed effects relate to whether the primary classification is listed as Advocacy and Policy Development, Financial and Budget Management, ICT Management and Research, Monitoring/Training/Personnel Management, Physical Infrastructure, Permits and Regulation, or Procurement. Sector fixed effects relate to whether the output is in the administration, environment, finance, infrastructure, security/diplomacy/justice or social sector.

Figure A1: Quarterly Report, an Example

MINISTRY OF TRADE AND INDUSTRY - DRAFT 2015, SECOND QUARTER PROGRESS REPORT

•					
	PROGRAMMED	BUDGET	EXPECTED	ACTUAL OUTPUT	REMARKS
	ACTIVITIES	PERFORMANCE	OUTPUT/IMPACT		
		MAI	NAGEMENT AND ADMINIST	RATION	
D ¹ 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1			ADMINISTRATION		
Division name					
			LEGAL DIVISION		
	Output 1.Competition Police	Developed and App	reved		
	1 Develop Draft Competition	y Developed and Appl	Draft Competition Policy	Evaluation of technical	Procurement process is
	Policy		developed	proposals completed. Two	still on-going and expected
	Policy		developed	firms have been	to be completed by end of
				shortlisted.	Quarter 3.
Expected output					
Expected output				The two firms have been	
				invited to present financial	
				proposals.	
	2. Organize stakeholder		Stakeholder meeting held	Yet to be done	
	meeting to validate draft				
	Competition Policy				
	3. Prepare Cabinet Memo		Competition Policy	Work on the bill is still	1
	and submit Competition		submitted to Cabinet	frozen due to on-going	
	Policy to Cabinet for			work on the policy	
	approval				
	Output 2: Competition Bill pres	pared and approved	I	1	I
	1. Review the Draft	parca and approved	The draft bill reviewed	Yet to be done	Work on the hill will begin
	Competition Bill taking into		ine arere on reviewed		after the Policy has been
	competition bill taking into				

Notes: Key information used in coding is highlighted The Appendix provides details on all output data variables.