Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service[*]

Imran Rasul   Daniel Rogger[†]

October 2013

Abstract

We study how the management practices that bureaucrats operate under, correlate to the quantity and quality of public services delivered. We do so in a developing country context, exploiting data from the Nigerian Civil Service linking public sector organizations to the projects they are responsible for. For each of 4700 projects, we have hand coded independent engineering assessments of each project's completion rate and delivered quality. We supplement this information with a survey to elicit management practices for bureaucrats in the 63 civil service organizations responsible for these projects, following the approach of Bloom and Van Reenen [2007]. Management practices matter: a one standard deviation increase in autonomy for bureaucrats corresponds to significantly *higher* project completion rates of 18%; a one standard deviation increase in practices related to incentives and monitoring corresponds to significantly *lower* project completion rates of 14%. We provide evidence that the negative impacts of practices related to incentive provision/monitoring arise because bureaucrats multi-task and incentives are poorly targeted, and because these management practices capture elements of subjective performance evaluation that further leave scope for dysfunctional responses from bureaucrats. The backdrop to these results, where 38% of projects are never started, implies there are potentially large gains to marginally changing management practices for bureaucrats.

Keywords: autonomy, bureaucracy, multi-tasking, performance evaluation.
JEL Classification: H00, H11, J33, J38, M1, O20.

1   Introduction

We study the correlates of effective public service delivery in an important developing country context: Nigeria. To do so, we combine data sources linking the outputs of government bureaucracies with details of how bureaucrats are managed. On outputs, we use project level data measuring the completion, quality and complexity of over 4700 public sector projects implemented by organizations in the Nigerian civil service, including government ministries and other federal agencies. On management, we adapt the methodology set out in Bloom and Van Reenen [2007] to measure management practices for bureaucrats in each organization, along multiple dimensions. Our study sheds light on the relationship between the management practices for bureaucrats in each organization, and the quantity and quality of projects delivered by the organization.

From a macroeconomic perspective, the effective functioning of the bureaucracy is an important determinant of poverty, inequality, and economic growth as stressed by the state capabilities literature [Besley and Persson 2010]. Effective public service delivery also matters from a microeconomic perspective: program evaluations of micro-scale interventions are often partly motivated by the assumption that successful interventions can be faithfully scaled-up by governments.

However, despite the importance of government effectiveness for citizen welfare, the public administration literature is almost devoid of concrete evidence linking practices in civil service organizations to public goods outcomes [Goldfinch *et al.* 2012]. At the same time, economic analyses of incentives in the public sector have largely focused either on the selection and motivation of *politicians* [Besley 2004, Gagliarducci and Nannicini 2013], or on the response to incentives of *frontline* staff such as teachers and health workers [Glewwe *et al.* 2010, Muralidharan and Sundararaman 2011, Duflo *et al.* 2012, Ashraf *et al.* 2012, Miller and Babiarz 2013]. In both rich and poor country contexts, there is little evidence linking the managerial practices that the vital middle-tier of *bureaucrats* operate under, to public service delivery. It is this gap we start to fill.

Banerjee *et al.* [2007] highlight two constraints restricting research on public good provision in developing countries: (i) the process of project implementation is rarely quantifiable; (ii) public good quality is difficult to measure. The data we utilize allows us to make progress on both issues. More precisely, we exploit a unique period of history in the Nigerian civil service, during which the activities of public organizations were subject to detailed and independent scrutiny. As part of this process, quantitative information was collected to measure the actual *implementation* success and *quality* of public sector projects in various social sectors. The scrutineers were independent teams of engineers and members of civil society.

We have hand coded this information to obtain potentially unbiased assessments of individual project completion rates and their quality, for over 4700 public sector projects that began in 2006/7. The bulk of project types we study are construction projects, such as boreholes, buildings, dams and roads. We have also used the technical documents available for each project, to work with engineers to construct measures of each project's *complexity*, following engineering best-practice [Remington and Pollack 2007]. The aggregate budgetary cost of the projects we study is US$800 million or 8% of all social spending in Nigeria over this period.

To measure the management practices that bureaucrats operate under, we follow the methodological approach set out in Nick Bloom and John Van Reenen's pioneering work [Bloom and Van Reenen 2007, 2010]. We adapt their management surveys to the Nigerian public sector setting, taking account of insights from the public administration literature [Rose-Ackerman 1986, Wilson 1989]. We collected data on management practices for 63 organizations of the Federal Civil Service in Nigeria, including central ministries and regional development authorities. For each organization, we derive three measures of management practices: one broadly related to the *autonomy* provided to bureaucrats; one related to the provision of *incentives* and *monitoring* of bureaucrats; and one capturing multiple other aspects of management practices.

The autonomy index we construct captures the extent to which: (i) bureaucrats can input into policy formulation and implementation processes; (ii) the flexibility with which an agency can re-organize its bureaucrats to respond to best practice and project peculiarities. There are long-standing views in the public administration literature on the importance of autonomy. As Rose-Ackerman [1986] describes, at one extreme lies the view that public agencies ought to delegate as much decision making to bureaucrats as possible, relying on their professionalism and resolve to deliver public services [Simon 1983]. At the other extreme lies the Weberian view that, because the objectives of bureaucracies and society diverge, only an entirely rules-based system of public administration, that leaves little to the individual judgement of bureaucrats, can ensure consistent and acceptable levels of public service. The second reason to focus on autonomy stems from the economics of organizations literature. Despite the early prominence of autonomy in this literature [Simon 1951], and the recent contribution of Aghion and Tirole [1997] that has revitalized interest in discretion in organizations, little evidence exists on the causes and consequences of providing autonomy in organizations. Our analysis also starts to fill this gap.

The incentives/monitoring based management index we construct captures the extent to which an organization collects indicators of project performance, how these indicators are reviewed, and whether bureaucrats are rewarded for achievements reflected in these indicators. This is an important dimension of management practices given the literature on incentive theory stressing the positive impacts performance incentives and monitoring have on organizational performance. However, *a priori* the impacts of such incentives in *public sector* settings is uncertain because: (i) bureaucrats might need to engage multiple effort types, not all of which are measurable; (ii) the process by which inputs are converted to outputs is uncertain; (iii) there can be competing views on the right way to implement bureaucratic outputs; (iv) bureaucratic objectives are not clear cut; (v) performance incentives might crowd out the intrinsic motivation of those self-selected into the public sector [Perry and Wise 1990, Benabou and Tirole 2006, Francois and Vlassopoulos 2008].[1]

---

[1]Positive impacts of performance pay for teachers have been documented using RCTs in developing countries by Glewwe *et al.* [2010], Muralidharan and Sundararaman [2011], and Duflo *et al.* [2012], although Fryer [2013] finds zero or even negative impacts in the US. In health, a nascent literature documents positive impacts of performance pay in developing countries when provided to frontline workers [Miller and Babiarz 2013]. In line with our findings, Ashraf *et al.* [2012] document how non-monetary incentives elicit more effort than monetary incentives for such tasks. Perry *et al.* [2009] and Hasnain *et al.* [2012] present literature reviews suggesting a lack of evidence that incentives positively impact bureaucrat behaviors. Muralidharan [2012] provides extensive discussions of why performance related pay might be specifically suboptimal in the public sector.

We probe these issues by exploiting a third data source: a survey we fielded to a representative sample to over 4100 civil servants, corresponding to 13% of the total workforce of the 63 organizations we study. This asked bureaucrats questions related to their tenure and intrinsic motivation, their perceptions of corruption in the organization, and their employment histories as bureaucrats. We use this evidence both to shed light on how the impacts of management practice vary with bureaucrat characteristics, and also to address econometric concerns related to non-random sorting of bureaucrats into organizations as a function of the management practices in place, or the endogenous determination of management practices as a function of bureaucrat characteristics.

Central to our empirical method is that, for any given project type, *multiple* organizations are observed conducting similar project activities. For example, small-scale dams are constructed by the federal ministries of water, agriculture, and environment. We therefore assess how the delivery of the *same* project type varies depending on the management practices in place for bureaucrats in the specific organization responsible, holding constant all other project characteristics, such as their complexity and scale, as well as other characteristics of the organization and bureaucrats.

We present three core findings linking civil service management practices and public service delivery in Nigeria. First, the management practices bureaucrats operate under matter. Despite the measures of management practice related to autonomy and performance incentives being positively correlated to each other, they have *opposing* correlations with the quantity of public service delivery: a one standard deviation increase in autonomy for bureaucrats corresponds to significantly *higher* project completion rates of 18%, and a one standard deviation increase in practices related to incentives/monitoring corresponds to significantly *lower* project completion rates of 14%. Moreover, we find management practices correlate to *quality-adjusted* project completion rates in similar ways as documented for project completion rates. The backdrop to these findings in Nigeria, where 38% of public projects are never even started, implies these magnitudes are of economic as well as statistical significance. If these findings reflect causal impacts, they suggest the potential gains to marginally changing management practices are huge.

At face value, the interpretation of our findings on autonomy for bureaucrats provide support to the notion that public agencies ought to delegate some decision making to bureaucrats, relying on their professionalism and resolve to deliver public services [Simon 1983]. The evidence is less supportive of the notion that when bureaucrats have more agency or organizations are more flexibly structured, then they are more likely to pursue their own objectives that diverge from societal interests, resulting in fewer public services being delivered.

The robust *negative* correlation we document between project completion rates and management practices related to the provision of incentives and monitoring of bureaucrats, is far more surprising and counter to a large body of evidence from private sector settings. The interpretation of this result is nuanced because we measure management *practices* related to the provision of incentives and monitoring: we do not measure the use of explicit compensation schemes based on output or project completion. As such, the detrimental impacts of such practices for bureaucrats can be interpreted in at least two ways.

First, bureaucrats might operate in a classic multi-tasking environment, in which they exert

multiple types of effort, only some of which generate productive outcomes such as public projects being completed [Holmstrom and Milgrom 1991]. Other efforts that bureaucrats engage in might be termed 'processing', or 'ticking boxes'. In essence, our management practice measure might then capture an incentive system that places excessive regulatory burden or 'red tape' on bureaucrats, that has long been argued to lead bureaucrats to mis-allocate effort towards such processing activities [Kelman 1990]. As management practices in bureaucracies become more overtly concerned with incentivizing and monitoring such processing efforts (perhaps because they are much less costly to specify or monitor than more productive types of bureaucrat effort), then bureaucrats will optimally reallocate effort toward such activities, *reducing* project completion rates overall.

A second interpretation of the negative impact management practices related to performance incentives and monitoring have in this bureaucratic setting is that the practices we measure pick up *subjective* performance evaluation (SPE), rather than explicit incentives based on signals of productive effort/output. While SPE has the benefit of being based on a more rounded set of assessments, it is well known that such subjective assessments also give rise to other biases and dysfunctional responses, especially the desire of agents to engage in influencing activities to curry favor with supervisors [Milgrom 1988, Milgrom and Roberts 1988]. If so, the increased use of such mis-targeted incentives and key performance indicators can also lead to bureaucrats reallocating effort towards non-productive tasks, thus *reducing* project completion rates overall.

To understand whether the incentive/monitoring related management practices we measure capture elements of multi-tasking or SPE, our second set of results explores how the impacts of management practices vary with project and organizational characteristics. We document the negative impact of incentive based practices are even *more negative* for: (i) more complex projects (that might reflect multiple effort types needing to be exerted); (ii) for projects that organizations implement less frequently and might be less specialized in; (iii) project types that our independent engineering assessments code as being of inherently greater ambiguity or uncertainty in design and implementation. Also in line with the interpretation that incentives are mis-targeted or subjectively determined in this context we find that the negative impacts of performance-incentive practices are *ameliorated* in organizations: (i) that give more discretion to bureaucrats as measured by our index of management practices along this dimension; (ii) with better IT facilities, that might capture their greater ability to measure and target incentives towards the most productive efforts bureaucrats engage in; (iii) staffed by less experienced bureaucrats, that might reflect that bureaucrats learn how to engage in influence activities with experience.

Our third class of results shed new light on the interplay between management practices and *bureaucrat* characteristics. We first address the literature suggesting performance-based incentives might crowd out the intrinsic motivation of public sector workers. We find the negative impact of performance-based incentives is *offset* by the share of bureaucrats in the civil service organization that are intrinsically motivated. Hence, if anything, our evidence suggests crowding-in of bureaucrat effort in the presence of practices related to incentives/monitoring.

While the recent economics literature has emphasized the role of intrinsic motivation, a long-standing literature in public administration emphasizes that civil servants pursue their self-interest

5

[Tullock 1965, Wilson 1989]. This more negative view of bureaucrats spurs our final set of results, that explore how the impacts of management practices are mediated through *perceptions of corruption* among civil service organizations, as elicited in our civil servant survey. Clearly, in the context of Nigerian bureaucracies the issue of corruption cannot be ignored, and it permeates throughout our analysis. We find a large negative levels impact on completion rates of perceptions of corruption in organizations, but the negative impacts of performance-based incentives do *not vary* with perceptions of corruption. This suggests that corruption is certainly prevalent in this environment, but that there are few bureaucrats that are actually *prevented* from behaving corruptly because of marginal changes in management practice.

To provide support for a causal interpretation of our results linking management practices to project outcomes, we later tackle three econometric challenges: (i) projects being *assigned* to organizations based on their management practices; (ii) *unobserved* bureaucrat or organizational characteristics that are correlated to management practices and also drive project completion rates; (iii) management practices being *endogenously* determined.

The central contribution of the paper is to provide novel evidence on how management practices for bureaucrats in civil service organizations correlate to effective public sector service delivery, and how the impacts of each dimension of management practice vary across the characteristics of projects, organizations, and bureaucrats. Our results point to new directions for theoretical research to better understand the contracting environment in public bureaucracies, as well as highlighting specific areas in which the better measurement of inputs and outputs of public sector organizations can aid our understanding of public service delivery in the developing world.

The paper is organized as follows. Section 2 overviews relevant aspects of the Nigerian civil service. Section 3 details our data sources and empirical method. Section 4 presents our core results linking public service delivery and management practices for bureaucrats. Section 5 provides evidence to help underpin a causal interpretation of our findings. Section 6 concludes by discussing implications of our results for optimal management practices, and links our findings to the literature on improving public services. The Appendix presents further data description and robustness checks.

2   Institutional Background

Nigeria is Africa's most populous country, home to 160 million individuals, double the size of any other African country and representing 20% of the population of sub-Saharan Africa. It is a leading setting in which to understand the determinants of public service delivery in the developing world, and also shares important characteristics with other developing countries: government expenditures represent 26% of GDP, it has generally weak institutions holding government to account, and corrupt practices in public sector organizations are commonplace. The British colonial government fashioned its Nigerian administration after the British Parliamentary Civil Service System, and this is essentially what passed to the independent government of Nigeria in 1960.

Although there have been a number of periodic reforms of the civil service, and despite the fact that Nigeria has moved to a Presidential system of government, Nigeria's civil service structure still largely replicates its British colonial origins.[2]

2.1 Civil Service Organizations

The Nigerian civil service is organized into federal, state, and local government civil services. Our analysis relates exclusively to federal civil service organizations. Various *organizations*, including ministries, are established by statute to render specified public services.[3] Table A1 lists the 63 federal civil service organizations we study. These include ministries of health, education, environment, and water resources, and organizations that have regional bases (such as federal polytechnics, federal medical centres, development authorities etc.). Table A1 highlights how these organizations vary in their budget sizes, staffing levels, and decentralization. The federal ministries are typically the largest in terms of budget, with regional organizations typically having fewer staff and being deconcentrated from central government.[4]

Each civil service organization is tasked to provide various types of *project*. These include construction projects: boreholes, buildings, roads and canals; as well as non-construction projects such as procurement, training, and advocacy. Crucially for our analysis, for any given project type, *multiple* organizations are observed conducting similar project activities. For example, small-scale dams are constructed by the federal ministries of water, agriculture, and environment, and by all of the river basin development authorities. We therefore assess how the delivery of the *same* project type varies depending on the incentive structures in place for bureaucrats in the specific organization responsible, holding constant other project and organizational characteristics.

Underlying our analysis is the notion that the objectives of civil service organizations do not run entirely counter to raising project completion rates. Our analysis is informative on the importance of management practices as long as organizations place *some* weight on this outcome. Anecdotal evidence suggest senior management have sometimes been removed on the basis of poor completion rates. Moreover the types of capital expenditure on projects we focus on constitutes 46% of organizational budgets in total, on average, so they undoubtedly represent an important element of what organizations are responsible for overall.

---

[2]The constitution adopted since 1999 has many similarities with that of the United States Constitution. Legislation is enacted by a bicameral National Assembly composed of the Senate and the House of Representatives. Each federal state has its own government headed by a Governor and a state parliament. Although the introduction of a Presidential system of government in 1979 saw initial reforms to the civil service, for example under the 1988 Civil Service Reorganization Decree No. 43, later decrees reversed some of these changes.

[3]The Civil Service is governed by a set of Public Service Rules and Financial Regulations, *ad hoc* Circular Instructions, decrees circulated across government, and Gazette Notices (decrees published in the Government's gazette). Together these outline the laws regulating the business of government, and cover service appointments and exits, discipline, salaries, management of public funds, and other major aspects of official assignments.

[4]Budget figures are averages for 2006-10. Staff numbers are taken from administrative data for 2010. Decentralized organizations refer to those whose day-to-day running is largely independent of the central authority. They have boards of governors that make decisions over policy and operation, and they have a separate budget line to central ministries. In line with the literature, we refer to such organizations as being 'deconcentrated'. 'Concentrated' organizations have a direct line of responsibility to the President and National Assembly.

2.2 The Assignment of Civil Servants and Projects to Organizations

The Head of the Civil Service of the Federation organizes the postings and conditions of Nigeria's federal civil servants. Our representative survey of 4100 individual civil servants confirms this: 88% of civil servants said they had no influence over their initial posting; 60% report their current posting being 'at random', with a further 22% reporting being transferred across organizations on an impersonal basis. Hence it is unlikely that bureaucrats self-select into organizations on the basis of the management practices in place, as evidenced in greater detail later.

Once posted, civil servants enjoy job security. Our survey reveals mean tenure at the *current* organization to be almost 13 years. For senior managers (those above grade level 12) this rises to almost 16 years. The survey also reveals that movements across organizations are rare: 67% of bureaucrats report never having moved organization. This lack of mobility in the labor market for bureaucrats slows down the rate at which best practices for management spread through the bureaucracy, and might also dampen incentives for organizations to tailor management practices in order to attract certain types of bureaucrat.

Projects are assigned to organizations centrally by the National Assembly, that enacts a budget law specifying the projects to be implemented each fiscal year. The projects we study were all established in law by Budget Appropriation Bills passed in 2006 or 2007. The passage of these bills is as follows. Having received inputs from the executive branch of government, a draft Appropriation Bill is presented to the National Assembly. The draft bill is then split into sectors (water, health etc.) and sent to sectoral committees of the House and Senate. These committees are delegated to hold hearings with relevant parties, to scrutinize the proposals, define budgets for each project and decide which organization is responsible for its implementation. These committees are staffed by politicians with qualifications or experience in the relevant sector. These sectoral committees then recommend a budget for the sector to an Appropriation Committee which merges the recommendations into a single budget. This unified budget is then voted on by both houses to form that year's Budget Appropriation Bill. This legal document defines the responsibilities of civil service organizations in terms of projects to be delivered.

2.3 Management of Bureaucrats

It is at this point that the *management* of civil servants becomes crucial for the transformation of government plans into public service delivery. To better understand how management practices evolve in organizations, we held structured interviews at four of the federal organizations in Table A1. These revealed three common themes in how management practices are determined in organizations: (i) the Public Service Rules of the Nigerian civil service provide guidelines on how bureaucrats should be incentivized, and these are common to all federal organizations; (ii) the history of senior management staff that have worked in an organization might bring their own innovations to bear; (iii) the role of external events such as demands of trade unions. Taken together, these interviews consistently emphasized the nature of management practices in the civil service organizations to be *slowly evolving* over time, and not necessarily tailored to maximize

public service delivery of the types of project organizations are used to being assigned.[5]

An econometric concern is the potential reverse causality from bureaucrat effectiveness driving the management practices in place. While we cannot rule this out entirely given our data, we also reiterate that such processes for how management practices were determined were not mentioned in any of the structured interviews conducted. We address this issue in detail in Section 5.

3   Data and Empirical Method

3.1   OPEN Data on Project Completion and Quality

The Nigerian Government began a program of sweeping reforms in the major organs of government in 2003 [Nkonjo-Iweala and Osafo-Kwaako 2007]. As a result, the Nigerian Government received cancellation of its external debt to the tune of US$18 billion from the Paris Club. At the federal level, the annual savings from debt interest were channeled into the social sectors that we study. The Presidency also saw this as an opportunity to track the effectiveness of government expenditures, and so in 2006 and 2007 the Nigerian Government undertook the Overview of Public Expenditure in NEEDS, known as the 'OPEN initiative', in which it traced, at a project level, the use and impact of 10% of *all* federal Government social sector expenditures approved between 2006 and 2007. The projects selected to be part of the OPEN initiative were designed to be representative of existing social sector expenditures, but also to be informative for those projects that were most needed to be scaled-up nationwide.

Under the OPEN initiative, visits to public projects by expert teams identified the extent to which they had been implemented as planned in the Federal Budget, and embodied in each project's technical document. The Presidency contracted national and regional teams to undertake the monitoring process outside of the institutions of the civil service. Hence the public sector projects were *not* evaluated by potentially biased civil servants, but rather by teams of independent engineers and civil society. The engineers evaluating the projects were not those working on the project sites and the civil society groups were recognized third sector organizations.[6]

---

[5]The structured interviews took place more than two years after the other surveys were fielded. In the Nigerian public sector, management practices are said to take the Public Service Rules as their foundation. These rules provide the framework for the running of the public service, including rules around the distribution of authority, the nature of discipline, the provision of training, and so on. In each case, these were said to be central to determining management practice. However, we were repeatedly told that a secondary influence on an organization's management practices was the history of management staff who had worked at the organization. Officials are promoted into management positions based primarily on tenure. Views on management practices are said to be aggregated by committee, with the chief executive marshalling, rather than defining, the direction of reform. This situation, we were told, leads to a relatively slow changing management environment, but one that over the years can lead to substantial divergence in management practice across organizations. Finally, external events, such as the demands of trade unions, were said to have a third-tier influence and constrain management practices.

[6]The teams were recruited in a competitive tendering process that was regulated by the government's procurement agency. The decision to use monitoring teams independent of government was a function of the weakness of existing government systems as well as the need for impartiality [Federal Government of Nigeria 2007]. Prior to the OPEN initiative, the government had its own monitoring and evaluation systems in place (based on unannounced visits) but these were largely perceived to be ineffective.

9

We consider projects traced under the OPEN initiative that were approved in the 2006 or 2007 federal budgets. Monitoring teams visited the relevant project sites around 18 months after the project was centrally approved. All the projects we study had twelve month completion schedules, so that even accounting for any delay in the disbursement of funds, it is feasible for these projects to be completed by the time of the monitoring survey.

The OPEN evaluation teams coded: (i) whether the project had started; (ii) its stage of completion; (iii) the quality of the inputs and work. Our main outcome variable is a *continuous* measure, from zero to one, of project completion rates. A recorded completion rate of zero does not imply the organization never even attempted to work on the project. Rather, the project might have been formulated and prepared, with responsibility for implementation having been delegated to a department and bureaucrats within the organization. At that point however, progress on the project halted, with funds either being returned due to lack of use, or being lost. We cannot distinguish whether this lack of implementation reflects active or passive waste [Bandiera *et al.* 2009]. A completion rate of one implies the project matched its full specification.

To maximize data coverage on project quality, we are forced to utilize the most aggregate formulation of quality reporting. A project was either of insufficient quality, satisfactory, or commended for an 'above average or high' quality level. With this definition, we obtain 2235 observations of project quality, 2206 of which also have project completion data. We then define a project quality indicator equal to one if the project is of satisfactory quality or above.

To further ensure the accuracy of monitoring reports, the Presidency put in place a system of checks and balances. First, a centralized team of technocrats monitored the evaluation teams, providing them with training and opportunities for standardization of their methods at national conferences. Second, evaluators were asked to provide material, photographic, or video evidence to support their reports. Third, the national teams and Presidency performed random checks on evaluated sites, all of which were consistent with the findings of OPEN monitors. Evaluations of the OPEN process indicate it successfully achieved its aims [Eboh 2010, Dijkstra *et al.* 2011].

The reports of OPEN evaluators describe the fate of projects budgeted for execution in the 2006 and 2007 federal budgets [Federal Government of Nigeria 2008, 2009]. We hand coded the material from all projects recorded in OPEN initiative reports from the federal civil service organizations listed in Table A1.[7] Taken together, the coverage of projects in our sample traces 8% of all Federal Government social sector expenditures in 2006/7 budget years, corresponding to 4721 projects from 63 organizations, with an aggregate budget of around US$800 million.

We also hand coded other project level characteristics such as its budget, whether it was a rehabilitation project, and, as described in mode detail below, its complexity as defined by engineer-approved measures derived from the project's technical specifications.

Table 1 provides descriptive evidence on each project type studied. Boreholes are the most

---

[7]We have a relatively broad sample of federal social sector organizations. In the water and power sectors, we cover all the relevant federal organizations. In the health sector, we cover 28% of health organizations, with the excluded a subset of the medical service providers such as a number of Federal Medical Centres. Similarly, in education we cover 14% of education organizations, excluding a range of institutions of learning such as some Federal Colleges of Education.

common type, corresponding to 29% of all OPEN projects. Key to our empirical method is that most project types are implemented by a range of organizations: for example there are 18 civil service organizations that construct boreholes, as shown in Column 2. Hence we are later able to study the impact of management practices for bureaucrats, that vary across civil service organizations, on project implementation, *conditional on* project-type fixed effects.

Column 3 highlights the scale of projects by project type: most constitute the 'nuts-and-bolts' of rural infrastructure development: the median budget for dams is US$18,000, and the median budget for a building is US$120,000. It is because projects are relatively small scale that partly explains why multiple organizations are observed being tasked to implement similar project types. Columns 4 to 7 show completion rates by project type. Aggregating across all project types, 38% of projects are never started. Although this might emphasize the role that corruption or passive waste might play, it is not a complete explanation for bureaucrat behavior: 31% of projects are *fully* completed, and conditional on being started, the average project completion rate is .75. Examining the data by project type emphasizes the variation of completion rates across and within project types. Conditional on being started, most projects are more than 50% complete. The final column provides information on the percentage of projects rated to be of satisfactory quality by the team of independent engineers and civil society: here we tend to observe the majority of projects being ranked highly irrespective of project type.

Table 2 presents descriptive evidence on the public service delivery of the ten largest civil service organizations in our sample, as defined by the total number of projects implemented (although this maps closely to the size of organizations as measured by total budget). This again emphasizes that, with the exception of the Federal Ministry of Power and Steel, each organization is engaged in providing multiple project types. We observe huge variation across these large organizations in the percentage of projects that are never started (11% to 95%), and that are fully completed (3% to 89%). The final column provides information on the percentage of projects rated to be of satisfactory quality: here we observe far greater variation across civil service organizations (25% to 100%) than we previously documented in Table 1 across project types.

These statistics suggest there might be important factors at the organization level that drive this variation in the quantity and quality of public sector projects. We next detail how we measure one such factor: the management practices civil service bureaucrats operate under.[8]

3.2 Measuring Management Practices

There has been a revival of research investigating the impacts of management practices on the performance of private sector firms [Ichniowski *et al.* 1997, Black and Lynch 2001, Bloom and Van Reenen 2007, 2010; Bloom *et al.* 2012, 2013]. We follow Bloom and Van Reenen's (henceforth BVR) approach to measuring management practices in organizations. We adapt their survey tool and practices to the Nigerian public sector setting, taking into account long-standing views on

---

[8]Regressing project completion rates on organizational fixed effects yields an R^2 of .33, suggesting that organizational characteristics such as management practices can plausibly play an important role in determining such outcomes.

the importance of autonomy in public administration [Simon 1983, Rose-Ackerman 1986, Wilson 1989] as well as recent insights from the 'new performance management' and 'good governance agenda' perspectives [Francois and Vlassopoulos 2008, Goldfinch *et al.* 2012].

To obtain reliable information on management practices, we worked closely with members of the OPEN office in the Presidency, as well as members of the Office of the Head of the Civil Service of the Federation in undertaking the questionnaire development process. A number of pilots using semi-structured interviews like those in BVR were held to outline key similarities and deviations from the BVR methodology. After a number of months of collaborative questionnaire design, civil servants from each organization practised the survey with each other and identified where wording or phrasing was not suitable for the Nigerian context.

The management survey enumerators were trained together for a number of weeks including practice interview sessions before undertaking the first few interviews together. The aim was to ensure a consistent interview engagement across sessions. To obtain information on management practices, senior management staff from the key departments of the organization, but not the chief executive officer, were brought together in a private office to discuss managerial practice at the organization in confidence. While each manager filled in their own questionnaire, the enumerator looked for a consensus and recorded that in her own questionnaire. This is the underlying information we use to construct management practice indices for each organization.[9]

From September to November 2010, our survey team held interviews at the organizations listed in Table A1. Following BVR, interviews were 'double blind' in that: (i) managers were not told in advance they were being scored or shown a score grid; (ii) enumerators were given no information on the performance of the organization. However, the delay between the collection of the OPEN data set in 2006/7 and the Civil Servants Survey in 2010 raises the question whether civil service structures changed significantly in between data collection periods. For example, those organizations found to have low completion rates might have instigated reforms to improve management practices for bureaucrats. However, there is little evidence from other sources of any major civil service reforms being implemented over this period, or of significant changes in the political organization of federal agencies [Alabi and Fashagba 2010, Ogundiya 2011]. In addition, we find little evidence of a spike in turnover of bureaucrats around 2007, coinciding with the Presidential election: 80% of bureaucrats employed in 2010 were at the same organization in 2007.

The BVR evaluation tool elicits management practices through a semi-structured interview covering four topics: operations, targets, incentives and monitoring. We apply the BVR approach in the context of public bureaucracies, extending the series of practices elicited to cover those more relevant for managing bureaucrats. As a result, our management survey covers nine topics:

---

[9]In this context, conducting management surveys via telephone was judged to be far less likely to reveal true practices than conducting face-to-face group interviews. Second, given the interview format, individual manager responses on management practices are available, but we cannot link individual managers to specific projects and so do not utilize that information (each project is delivered by teams of bureaucrats across sub-departments): rather we use the consensus measure recorded by the enumerator. Managers were told their individual responses would remain confidential. Third, we checked whether recorded practices are sensitive to the number of managers present at interview (that does vary across organizations): we find no such relationship. Fourth, the mean interview time was 74 minutes, and we include this as part of our noise controls in the empirical analysis.

roles, flexibility, incentives, monitoring, culture, targeting, facilities, skills and staffing. We then replicate the BVR method eliciting information on each of these broad topic areas from our civil service organizations, although we do so using a more limited set of underlying questions related to each topic, and some of our questions permit only yes/no replies, while others are based on a full scoring grid. Table A2 details the questions that come under each of the nine topic areas, and how they are aggregated into what we refer to as 'autonomy', 'incentives' and 'other practice' indices (denoted CS-autonomy, CS-incentives and CS-other respectively).[10]

The questions on 'roles' assess the extent to which bureaucrats can input into policy formulation and implementation processes. The questions on 'flexibility' aim to measure whether a bureaucratic agency is able to re-organize its bureaucrats and adapt tasks to respond to best practice and project peculiarities. We combine answers to the roles and flexibility practices to construct an index of management practices capturing bureaucrats' 'autonomy' (CS-autonomy).

The questions on 'incentives' are designed to capture more familiar notions of incentive provision for bureaucrats, both positively in terms of whether rewards are given for some dimensions of service, and negatively in terms of punishments for poor service; the questions on 'monitoring' capture practices related to the collection and use of performance indicators. We combine the answers to the incentives and monitoring practices to construct an overall index of management practices capturing the performance related 'incentives' bureaucrats operate under (CS-incentives).

The remaining topics cover the following practices: the 'facilities' questions relate to how well-functioning the organization is, for example, by collecting information on the availability of electricity and internet facilities to bureaucrats; the 'skills' questions relate to the human capital of bureaucrats, especially their IT competencies, and the trainings offered to them; the 'staffing' questions focus staff recruitment and retention, as well as how workloads are spread across bureaucrats; the 'targeting' questions relate to the existence and clarity of targets for bureaucrats, and finally, the questions related to 'culture' try to elicit information on how colleagues are collectively treated and interact with others outside of the workplace. We combine the answers on all these topics to construct a third index of 'other' management practices (CS-other).

We clearly recognize there is no definitive way to solicit management practices along these various dimensions, nor a definitive way to collate them into more aggregate indices. Our approach is primarily designed to reflect two broad areas of management practice emphasized in the public administration and economics literatures as being first order determinants of bureaucrats' behavior: autonomy and performance incentives. However we later consider two extreme cases by documenting the correlation between project completion rates and: (i) a fully disaggregated

_____

[10]Hence there are two important deviations from how we elicit management practices using the BVR method. First, we tailored the precise wording of some questions to better fit our context. Second, we did not use the same universe of questions from the BVR survey. In most cases this was because we could not identify an analogous concept in the public sector that was relevant or not covered by other questions. For example, the majority of questions on lean manufacturing in BVR (e.g. 'What kinds of lean (modern) manufacturing processes have you introduced?') were not utilized. However, those on improving manufacturing processes ('How do you go about improving the manufacturing process itself?') were translated into the redefinition of procedures in response to new needs or challenges ('Does your organization make efforts to redefine its standard procedures in response to the specific needs and peculiarities of a community?').

specification showing the separate impact of all nine dimensions of management practice; (ii) an aggregate measure of management practice that collates all nine categories into a single index.[11] Given these indices measure *broad* practices, rather than reflecting *specific* compensation schemes or workplace policies, there might be less concern that such broad practices can rapidly endogenously respond to either past project completion rates, or to changing bureaucrat characteristics.

The responses to each practice in Table A2 are converted into normalized z-scores (so are continuous variables with mean zero and variance one by construction), where in all cases, variables are increasing in the commonly understood notion of 'better management'. For the CS-autonomy index, we assume greater autonomy corresponds to better management practices (our empirical analysis then assesses whether this is positively or negatively correlated to project completion rates). For our core analysis, we aggregate z-scores into the three CS- measures by taking the unweighted mean of the underlying z-scores. We later show the robustness of our results to other weighting schemes.

Two further points are of note. First, the CS-autonomy and CS-incentive management scores are *positively* correlated. Hence in the cross section of federal organizations, the provision of autonomy and performance incentives do not appear to be substitutes. Second, these correlations are not high: the CS-based measures have a correlation coefficient of .24. Combined with the underlying variation in each measure across organizations, this opens up the possibility to precisely identify the separate relationship of each measure to public service delivery.[12] The marginal impacts of these two measures can also be separately identified from that of the CS-other index: the CS-autonomy (CS-incentives) index has a correlation of .17 (.43) with the CS-other measure.

These indices, and the underlying management questions, provide us with our core explanatory variation. Following BVR, we also collected data on the interviewees for each survey, interview characteristics and the quality of the survey session. These 'noise controls' will also be conditioned on in our baseline empirical specifications as described later.

3.3   Project Complexity and Other Variables

When relating project outcomes to management practices for bureaucrats, it is important to condition on project complexity. To measure this we collaborated with a pair of Nigerian engineers familiar with the OPEN initiative and a group of international scholars with research interests in project complexity. The complexity indicators were based on the detailed technical specifications specified for each project, and are constructed following engineering best practice that emphasizes multiple dimensions of complexity [Remington and Pollack 2007]. The Appendix: (i) details the construction of these indices, and presents descriptive statistics for them; (ii) describes checks we put in place, using multiple engineers, to establish the validity of these complexity measures.

---

[11]We have also used principle components analysis to assess the importance of individual practices through factor analysis. For the matrix of all nine sub-indices, we find the first factor explains 28% of the variation.

[12]Such substitution might have been observed if bureaucrats have strong career concerns, and so performance incentives are not required once autonomy is given to individuals. Alternatively, if bureaucrats are intrinsically motivated they might need only to be provided autonomy, and indeed, the provision of explicit incentives might crowd out their intrinsic motivation.

These complexity indicators reflect the number of inputs and methods needed for the project, the ease with which the relevant labor and capital inputs can be obtained, ambiguities in design and project implementation, and the overall difficulty in managing the project. Some of these features relate to the hallmark differences in goods supplied between public and private sectors. Our empirical approach then aims to hold constant the complexity of the project along these dimensions, including issues related to organizations needing to sub-contract project implementation to a private sector firm, for example. This allows us to focus in on the correlation between managerial practices for bureaucrats and project completion rates, all else equal.[13]

The other project level controls include the project's budget and whether it was a rehabilitation project. Finally, organizational variables we control for include administrative data on the number of employees at each organization, the proportion of staff with graduate/postgraduate qualifications, and the organization's total budget. In some specifications we also condition on controls for the state in which a given project is located. We construct these state controls using aggregated data provided by the National Bureau of Statistics.

3.4 Empirical Method

Our baseline empirical specification has as its unit of observation project i of type j in organization n. The project types are listed in Table 1, and the federal civil service organizations from which management practices have been elicited are listed in Table A1. We estimate the following OLS specification, where y_{ijn} corresponds to the project completion rate, or the assessment of project quality, as described in Table 1, and the three main indices of management practice are the CS-autonomy, CS-incentives and CS-other indicators described above,

$$y_{ijn} = \gamma_1 CS\text{-}autonomy_n + \gamma_2 CS\text{-}incentives_n + \gamma_3 CS\text{-}other_n + \beta_1 PC_{ijn} + \beta_2 OC_n + \lambda_j + \epsilon_{ijn}, \quad (1)$$

PC_{ijn} includes project characteristics such as the project complexity, log project budget and whether the project is a rehabilitation or not. OC_n includes organization level controls such as the log number of staff, log total organization budget, log capital budget, and the proportions of officials with a college and postgraduate degree. Following BVR, within OC_n we also condition on 'noise' controls related to the management surveys. These include four interviewer dummies, indicators of the seniority, gender, and tenure of the managers who responded, the day of the week the interview was conducted, the time of day the interview was conducted, a dummy variable indicating whether the interview was conducted during Ramadan, the duration of the interview, and an indicator of the reliability of the information as coded by the interviewer, which is a simply a subjective assessment as to whether the interview went well.

As typically many organizations are observed implementing the same project type j, we control

---

[13]Our civil servant survey also helps to shed some light on the relationship between bureaucratic organizations and such third party contractors/suppliers/consultants. For example, only 6% of civil servants agreed with the statement that, the most successful contractors "are aligned with the government in some way"; only 13% of civil servants reported having been offered a 'small present', 'money', or an 'expensive present' by such contractors.

for project fixed effects λ_j in our baseline specification (1). Our parameters of interest are γ_1 and γ_2: as each CS- measure is a standardized z-score, these coefficients measure the effect size of a one standard deviation change in management practices along the respective margins of autonomy and performance incentives. As our main specification controls for project fixed effects, we allow the standard errors to be clustered by project type-organization. In the Appendix we demonstrate the robustness of our results to allowing standard errors to be clustered by organization (across all its project types). Our working sample is based on 4721 projects from 63 organizations on which we have data on management practices and project, organization and bureaucrat characteristics.

In specification (1) we are implicitly assuming that, within project type and controlling for project and organizational characteristics, the underlying production function is the same across projects. Specification (1) then corresponds to a reduced form representation of an underlying production function in which management practices convert the raw total of available bureaucratic labor into effective labor inputs in the completion of public projects.

To provide support for a causal interpretation of the estimated partial correlations linking management practices to project outcomes $(\widehat{\gamma}_1, \widehat{\gamma}_2)$, we need to tackle three econometric challenges: (i) projects are *non-randomly assigned* to organizations based on their management practices; (ii) *unobserved* bureaucrat or organizational characteristics that are correlated to management practices and also drive project completion rates; (iii) management practices being *endogenously* determined by bureaucrat characteristics or project outcomes. In the next section we present our OLS results on the partial correlations between various dimensions of management practice and project completion rates. Section 5 then provides evidence to address these econometric concerns.

4   Core Results

4.1   Project Completion Rates

Table 3 presents our main results on how civil service management practices correlate with project completion rates. Column 1 only controls for the three CS- measures. We find that higher levels of autonomy for bureaucrats correlate to significantly higher completion rates $(\widehat{\gamma}_1 > 0)$. The impact of greater use of performance incentives/monitoring for bureaucrats on project completion is however *negative* and also significantly different from zero $(\widehat{\gamma}_2 < 0)$. Columns 2 to 4 sequentially condition on noise controls and organizational characteristics, project characteristics, and project fixed effects. Throughout, we find both management practices have significant impacts on project completion rates at the 1% significance level.[14]

Our preferred specification is in Column 4 with project fixed effects. This shows that a one standard deviation increase in CS-autonomy corresponds to significantly higher project completion

---

[14]Among the noise controls, neither the age, gender, or tenure of interviewees significantly impact project completion rates. Nor do whether the interviews were undertaken in the morning, their length, or the quality rating of the interview. Noise controls that are significant include whether the interviews where undertaken on a Tuesday or Saturday (negative), or during Ramadan (positive). A few of the interviewer dummies are significant. The exclusion of the noise controls does not change the qualitative nature of the main results reported.

rates of 18%. A one standard deviation increase in CS-incentives corresponds to significantly lower project completion rates of 14%. The stability of both coefficients of interest and their standard errors across Columns helps ameliorate the concern that there likely remain other omitted controls that would cause large changes in magnitude and significance of the coefficients of interest.

The estimated coefficient on the other dimension management practice of CS-other, $\widehat{\gamma}_3$, is less stable in magnitude and significance across specifications. This might be as expected given that it is composed of the widest range of underlying components. In all specifications, better management practices on this CS-other dimension are positively correlated with project completion rates; in Column 4 the estimated coefficient is indeed significant at conventional levels, although the effect size is significantly smaller than for the two dimensions of management practice.

In Appendix Table A4 we show the robustness of these core findings to assuming the disturbance terms ϵ_{ijn} in (1) are clustered by organization (across all project types), correcting p-values for potential biases due to a small number of clusters [Cameron $et\ al.$ 2008]. In Figure A2 and Appendix Table A5 we probe the robustness of our baseline result along eight margins: (i) defining threshold completion rates that deem the project usable and seeing how management practices relate to reaching these thresholds (especially those related to completion rates of zero and one); (ii) restricting the sample to the largest/smallest organizations; (iii) removing organizations at the tails of the CS-autonomy and CS-incentives measures; (iv) considering the impacts of managerial practices on construction and non-construction projects separately; (v) considering the impacts of managerial practices on projects implemented by centralized and decentralized organizations separately; (vi) controlling for characteristics of the state in which the project are located, and exploring how the results vary depending on whether projects are located in the North or South of the country that defines the principal cultural divide in Nigeria; (vii) alternative constructions of the CS- management practice indices (rather than the equal weighting procedure); (viii) using a fractional regression model rather than OLS.

One concern is that project completion rates may not be a useful outcome indicator: projects may be almost completed but to a low quality standard. We therefore construct a 'quality-adjusted' completion rate where the proportion completed is multiplied by a binary quality indicator. Where quality is unsatisfactory, whatever the level of completion, this variable is set to zero. As described in Section 3, the drawback of using this outcome measure is that information on project quality is only available for around half the projects for which we have project completion data, originating in 51 civil service organizations. To ease comparison of the samples used, we note that if we re-estimate our main specification with the project completion rate as the dependent variable, but for the sample in which quality data is available, the partial correlations of each CS- measure with project completion rates are not much different to our baseline specification in Column 4.

Column 5 then shows the relationship between management practices and quality-adjusted project completion rates. We find the CS- measures to quantitatively impact quality-adjusted project completion rates in a similar way to project completion rates: higher levels of CS-autonomy are associated with significantly *higher* quality projects, and higher levels of CS-incentive are associated with significantly *lower* quality projects.

4.2 Exploring the Negative Impact of Performance Incentives

Our results confirm that the two dimensions of management practice we focus on: autonomy and performance incentives/monitoring, do indeed correlate to the quantity and quality of public services delivered, as emphasized by the public administration and economics literatures respectively. The overall positive correlation of CS-autonomy with project completion rates provides support to the notion that public agencies ought to delegate some decision making to bureaucrats, relying on their professionalism and resolve to deliver public services [Simon 1983]. On average, the evidence is less supportive of the notion that when bureaucrats have more agency or organizations are more flexibly structured, then they are more likely to pursue their own, potentially corrupt, objectives that diverge from societal interests, resulting in fewer public services being delivered.

The robust *negative* correlation we document between project completion rates and management practices related to the provision of incentives and monitoring of bureaucrats, is far more surprising and counter to a large body of evidence from private sector settings. As described in the introduction, the evidence on the impacts of performance-related incentives in public sector settings is mixed (often focusing on the impacts of specific compensation schemes to frontline workers).[15] Ours is among the first evidence to suggest the possibility that such management practices, *broadly* defined, have negative impacts on the vital tier of civil service *bureaucrats*.

The interpretation of the negative partial correlation is nuanced because we measure management *practices* related to the provision of incentives and monitoring: we do not necessarily measure the use of explicit compensation schemes based on project completion. As such, the detrimental impacts of such practices for bureaucrats can be interpreted in at least two ways. First, bureaucrats might operate in a classic multi-tasking environment, exerting some types of effort that can be labelled as 'processing' efforts, that do not lead to project completion rates, and also exerting other more productive types of effort that raise completion rates. Our management practice measure might then capture an incentive system that places excessive regulatory burden or 'red tape' on bureaucrats (say through the use of mis-targeted key performance indicators), that has long been argued to lead bureaucrats to mis-allocate effort towards processing activities [Kelman 1990]. Alternatively, our management practices related to performance incentives and monitoring might well pick up *subjective* performance evaluation (SPE), rather than explicit incentives based on signals of effort/output. While SPE has the benefit of being based on a more rounded set of assessments, such subjective assessments also give rise to other biases and dysfunctional responses, especially the desire of agents to engage in influencing activities to curry favor with supervisors [Milgrom 1988, Milgrom and Roberts 1988]. If so, the increased use of such mis-targeted incentives and key performance indicators can also lead to bureaucrats reallocating

---

[15]In health, two recent Cochrane reviews have come to different conclusions on the efficacy of pay for performance [Flodgren *et al.* 2011, Scott *et al.* 2011]. Perry *et al.* [2009] review 57 studies on pay for performance in the public sector and conclude 'pay-for-performance continues to be adopted but persistently fails to deliver'. Hasnain *et al.* [2012] review over 60 public sector studies, and find the vast majority are for tasks where outputs are more easily measurable such as teachers, health workers, and revenue inspectors. They argue there is simply insufficient evidence of the impact of incentives on bureaucrats.

effort towards non-productive tasks, thus *reducing* project completion rates overall.[16]

To begin to shed light on these possible interpretations, Column 6 of Table 3 interacts the two CS- measures of management practices for autonomy and performance incentives. We find evidence of an interplay between the two: the negative impacts of performance related pay on project completion rates are offset when bureaucrats are provided more autonomy and organizations are more flexible. This might be indicative of the fact that with such autonomy and flexibility, bureaucrats can, on the margin, more easily work around any poorly targeted elements of the incentive/monitoring environment they face that might otherwise lead to dysfunctional responses such as them exerting processing efforts or engaging in influence activities.[17]

To probe this further we now investigate how characteristics of projects, organizations and bureaucrats interplay with management practices. By doing so we aim to shed light on the extent to which our measured management practices might capture that multi-tasking concerns are first order for bureaucrats, or that they are subject to SPE and so might engage in influence activities.

4.2.1 Project and Organizational Characteristics

We first establish whether the impact of management practices related to performance incentives vary with the complexity of projects, assuming that more complex projects require more varied effort types to be exerted. If so, performance incentives/monitoring might be especially hard to design for such projects. Column 1 in Table 4 takes this prediction to the data by interacting the CS-incentives measure with the continuous measure of project complexity. For expositional ease, this interaction term is defined in terms of its deviation from mean, so the coefficients on CS-autonomy and CS-incentives are interpreted as the marginal effect of these practices, evaluated at the mean of project complexity. We see the negative impacts of CS-incentives related practices for bureaucrats are *exacerbated* in more complex projects, in line with a multi-tasking interpretation.

To probe further the idea that it is difficult to target incentives/monitoring for bureaucrats towards those efforts that aid project completion, we examine how the impacts of this management practice vary across projects conducted by the organization. More precisely, for each organization we define the *modal* project type assigned to it. We then create a dummy equal to zero if project i is of this modal type and equal to one if it corresponds to a more atypical project that the organization is tasked to complete. Column 2 shows the impact of performance incentives to indeed be even more negative among non-modal project types (the magnitude of the impact being around 40% of that on modal project-types). This suggests management practices related to performance are better tailored to the modal project type each organization is engaged in.

---

[16]Baker [2002] develops a multi-tasking model to characterize how distortion and risk affect the value and use of performance measures. The model highlights how in public bureaucracies, that cannot use stock incentives and have nebulous objective functions, this leads to a fundamental difficulty in defining 'good' performance measures. Hence the potential for dysfunctional responses when high powered incentives are utilized in such settings.

[17]Hence there are gains to organizations from coupling together good management practices along these two dimensions. Indeed as discussed in Section 3.2, we do observe the two CS- measures in the cross section of organizations having a correlation coefficient of .24, so better management practices along both dimensions are weakly bundled together in most cases.

Our third approach to shed light on whether incentives/monitoring for bureaucrats might be poorly targeted exploits information on the inherent riskiness/ambiguity of projects as encompassed in their technical specifications. Again we might think of project types with more inherent design ambiguity to be those for which incentives and monitoring are hardest to accurately target. To explore this, we consider projects of different types j to be of systematically different levels of ambiguity, and then proceed in two steps. First, we measure this design ambiguity using subcomponents of the complexity indictor described in Table A3. In particular we construct a z-score based on the design uncertainty, implementation uncertainty, design ambiguity and implementation ambiguity components of the project complexity metric. We then take the average of this over all projects of type j, denoting the average ambiguity of projects of type j by $\overline{\sigma}_j$.

Second, we estimate a specification analogous to (1) for a given project type j, obviously excluding project fixed effects. In our sample, there are sufficient numbers of projects implemented by enough organizations to estimate this for five project types: boreholes, buildings, dams, procurement and training. For each project type j we then obtain an estimate of the partial correlation between CS-incentives and project completion rates, $\widehat{\gamma}_{2j}$, conditional on CS-autonomy. Figure 1 then plots the five $(\widehat{\gamma}_{2j}, \overline{\sigma}_j)$ pairs, as well as a cubic best fit.

The evidence suggests a negative relationship across project types between inherent design ambiguities, as assessed by engineering best practice, and the marginal impact of management practices related to incentives and monitoring on project completion rates.[18] In short, projects with the greatest ambiguities in design and implementation, might be those for which incentives/monitoring schemes are hardest to design well, and so illicit the most dysfunctional responses from bureaucrats. The heterogenous impacts documented across project types in Figure 1 might also shed light on why the existing literature documents such mixed findings for the impacts of incentive provision in public sector settings. Our findings highlight such impacts might well vary with the nature of tasks/projects, such as their complexity and design ambiguity. Finally, we note that the average ambiguity of project types $\overline{\sigma}_j$ correlates with the proportions of those projects that have completion rates of zero. This mitigates against the idea that corrupt practices are entirely the reason behind low project completion rates. Rather, bureaucrats might simply shy away from implementing the most ambiguously designed projects.

Our fourth approach to understanding whether incentive/monitoring schemes might be poorly designed in this setting uses the intuition that if the negative impacts of performance incentives reflect the inability of organizations to correctly target such incentives to the relevant types of bureaucrat effort, this problem might be ameliorated in organizations with better IT facilities, who presumably have better capability to objectively measure effort types. To explore this we interact our CS-incentives measure with the practices captured in CS-facilities: as Table A2 shows, this relates to the availability of computing facilities at the organization. In this specification we redefine CS-other to exclude the CS-facilities component. The result in Column 3 of Table 4 indeed

---

[18]Given the CS-performance and CS-autonomy are positively correlated and have opposite signed impacts on project completion rates, this negative relationship between risk and incentives would be attenuated if we do not control for CS-autonomy when estimating $\widehat{\gamma}_{2j}$ [Prendergast 2002].

shows the impacts of CS-incentives to be significantly less negative in organizations with better IT facilities. Moreover, there is no significant *levels* impact of CS-facilities – the provision/accessibility of IT facilities for bureaucrats only seem to matter in combination with management practices related to performance incentives/monitoring.

4.2.2 Bureaucrat Characteristics

The remaining specifications consider how the impact of management practices related to performance incentives vary with three bureaucrat characteristics: their tenure, intrinsic motivation and perceptions of organizational corruption. To measure civil servant characteristics along each dimension we use the survey we administered to a representative sample of officials at each organization. We interviewed 4148 civil servants from the 63 federal organizations studied here, corresponding to around 13% of their total workforce.

As described in Section 2, Nigerian bureaucrats enjoy long tenure. Tenure can interplay with the effectiveness of management practices. On the one hand, longer serving bureaucrats might learn over time how best to respond to incentives by exploiting other flexibilities. On the other hand, if bureaucrats are subject to SPE they might learn how best to engage in influence activities. To check for this, Column 4 in Table 4 controls for an interaction between the average tenure of bureaucrats in the organization (in deviation from mean) with the CS-incentives measure, as well as the direct impact of tenure. We find the negative impacts of performance-based incentives are even worse in organizations staffed by more experienced bureaucrats, as might be consistent with bureaucrats learning to engage in influencing activities. We also note there is no direct levels effect of tenure on project completion rates, so that it is not the case that bureaucrats naturally reduce effort over time as they become embedded within long-standing norms of poor standards in civil service organizations. Rather, there appears to be a specific interplay between bureaucrat tenure and the use of performance incentives/monitoring for bureaucrats.

A burgeoning literature suggests those attracted to public service might be relatively more *intrinsically motivated* than those working in the private sector. Performance incentives might then be detrimental if they crowd out such intrinsic motivation [Rose-Ackerman 1986, Perry and Wise 1990, Benabou and Tirole 2006, Francois and Vlassopoulos 2008]. To measure civil servant's intrinsic motivation, we asked bureaucrats which factor that had most influenced them to *originally* enter the civil service from the following options: 'I was interested in the type of work', 'income prospects', 'the prestige associated with such a job', 'the stable career path that a job in the service affords', 'the chance to serve Nigeria', 'it was the only employment I could get', 'educational opportunities', 'other'. We define those that answered, 'the chance to serve Nigeria' as being intrinsically motivated. Roughly a third of officials state that they entered the civil service to serve Nigeria (with the percentage being slightly higher among less senior bureaucrats).[19] For

---

[19]In the public administration literature, public service motivation is usually measured using the scale developed in Perry [1996], based on statements related to politics, public service and pro-social activities. This is the approach also followed in Dal Bo *et al.* [2013]. Alternative approaches employed in the economics literature include: (i) dictator games to examine how many resources an individual transfers to a pro-social task [Ashraf *et al.* 2012];

each organization, we then construct the fraction of intrinsically motivated bureaucrats.

Column 5 then shows how the impacts of CS-incentives varies by the intrinsic motivation of bureaucrats in the organization. The previously documented negative impact of providing performance incentives to bureaucrats, is significantly offset when a greater share of bureaucrats are themselves intrinsically motivated. This suggests that although on average providing performance incentives to bureaucrats has negative impacts on public service delivery, this is *less* the case when such incentives are provided to intrinsically motivated bureaucrats. Indeed, our evidence suggests incentives/monitoring crowd-in the effort of intrinsically motivated bureaucrats. If our CS-incentives measure captures SPE being in place for bureaucrats, the result highlights that intrinsically motivated bureaucrats exert more productive efforts in organizations where such SPE are utilized to a greater extent, rather than engaging in influence activities, say.[20]

While the recent economics literature has emphasized the importance of the intrinsic motivation of bureaucrats, a long-standing literature in public administration emphasizes that civil servants might pursue their own self-interest or be disinclined to exert effort [Wilson 1989]. This more negative view of bureaucrats spurs our final set of results, that explore how the impacts of incentives/monitoring are mediated through perceptions of corruption among civil service organizations. Corruption in public bureaucracies is a first order issue in Nigeria, and in many countries at similar stages of development (although the fact that 31% of projects are completed fully also suggests corruption is not all pervasive).

To elicit information on perceptions of corruption, we began by discussing vignettes with bureaucrats, then made those scenarios closer to the bureaucrat's actual situation, and finally asked individuals about their own observations and experiences of corruption. We asked on what proportion of recent projects the official had worked on, did they observe 'others breaking service rules for their own benefit'. On average, officials stated that on 38% of projects such observations of corrupt practice had been made, that, by chance, coincides with the proportion of projects with a zero completion rate. We aggregate this to the organization level to construct the proportion of projects bureaucrats report having observed corrupt practices on.[21]

Column 6 shows how the impacts of our CS-incentives measure varies by perceptions of corruption among bureaucrats. To begin with we note the robustly negative *levels* impacts of our corruption measure on project completion rates. This affirms that our measure is indeed capturing some element of civil servant behavior that is deleterious for public service delivery. However, we see that the negative impacts of performance-based incentives do not vary with perceptions of

---

(ii) charitable contributions [Buurman *et al.* 2012]. In our civil servant survey, 'the chance to serve Nigeria' was the modal answer given. The other two most frequent reasons were 'I was interested in the type of work' and 'the stable career path that a job in the service affords', that were each given by around 20% of individuals

[20] Ashraf *et al.* [2012] present evidence from a field experiment in Zambia for workers hired to engage in a pro-social task that also documents how performance incentives crowd-in the effort of intrinsically motivated workers.

[21] We also asked whether officials had themselves been put under pressure to: (i) change the project location; (ii) change project specifications; (iii) help select particular contractors/suppliers/consultants; (iv) divert some of the funds. Aggregating responses into an organizational average, officials stated that they had experienced such pressures on 19% of projects. We prefer to use the measure related to *observed* corrupt practices over this measure because officials are obviously cautious when potentially incriminating themselves.

corruption among bureaucrats. In short, the evidence suggests that corruption has a direct and quantitatively large negative impact on project completion rates, but that marginal changes in management practices have no impact on such behaviors.[22]

Finally, we note that we have also explored the heterogeneous impacts of management practices related to autonomy. However, we place less attention on these findings because economic theory is less developed to guide such an empirical analysis. For completeness we note the impacts of management practices related to autonomy appear to be generally more homogeneous than the impacts of incentives/monitoring: they do not significantly vary with project complexity, non-modal project types, bureaucratic tenure, or the prevalence of reported corrupt behavior by bureaucrats. This last result again emphasizes that corrupt practices are unlikely to be affected by marginal changes in management practice. We do however find that the positive impacts on project completion rates of practices related to autonomy are significantly stronger when a greater proportion of bureaucrats report being intrinsically motivated: providing more autonomy and flexibility to bureaucrats might be especially beneficial when bureaucrats are socially minded. We leave for future research the exploration of such heterogeneous impacts.

5   Econometric Concerns

To view the estimated partial correlations on our CS- measures as being informative of causal impacts, we need to tackle three econometric challenges broadly related to the following issues: (i) projects being *non-randomly assigned* to organizations based on their management practices; (ii) *unobserved* bureaucrat or organizational characteristics that are correlated to management practices and also drive project completion rates; (iii) management practices being *endogenously* determined by bureaucrat characteristics or project outcomes.

5.1   Project Assignment

The first class of concerns relate to the assignment of projects to organizations on the basis of management practices in place. For example, if better managed organizations are more likely to be assigned harder-to-implement projects, this creates a spurious *negative* correlation between our CS- measures and project completion rates. This suggests $(\widehat{\gamma}_1, \widehat{\gamma}_2)$ are both biased in the *same* direction, and hence we note that a more complicated explanation is required to imply that $\widehat{\gamma}_1$ is both upwards biased and $\widehat{\gamma}_2$ is downwards biased. This issue notwithstanding, we address concerns about the non-random assignment of projects in two ways.

We first use a conditional logit model to estimate the factors determining the assignment of

---

[22]We have explored whether there exists within-sample values of the interacted variables in Table 4 at which the marginal impact of CS-performance becomes positive. Generally, this is not the case: even for the least complex projects or the most IT advanced organizations the marginal impact of management practices related to incentives/monitoring is negative. The one exception is bureaucratic tenure: in 8% of organizations average bureaucratic tenure is less than 7 years, and the coefficients from Column 4 then imply that the total marginal impact of CS-performance on project completion rates is positive.

project i to organization n, including the management practices in place in organization n. As described in the Appendix and Table A6, we find no robust evidence of a correlation between the assignment of projects to an organization and any of the three CS- measures. Our second approach conducts further analysis at the organization level, estimating the following specification,

$$w_n = \theta_1 CS\text{-}autonomy_n + \theta_2 CS\text{-}incentives_n + \theta_3 CS\text{-}other_n + \pi X_n + \epsilon_n, \tag{2}$$

where observations are for the 63 organizations listed in Table A1, w_n are measures related to the set of projects assigned to organization n by Parliament, the CS- measures are as previously described, and X_n includes the same organization level controls and noise controls as previously described. Robust standard errors are reported.

The results are presented in Table A7 and show that: (i) $\widehat{\theta}_1$ and $\widehat{\theta}_2$ are both positive, so that better managed organizations in dimensions related to autonomy and performance incentives are assigned more projects, but neither correlation is significantly different from zero (Columns 1); (ii) the unique number of project types assigned to an organization is also not significantly related to the management practices in place (Column 2). Taken together, the results in Tables A6 and A7 suggest that whatever the bargaining game between various stakeholders in the lengthy Parliamentary procedure that assigns projects to civil service organizations, there is not much evidence to suggest it is based on management practices in those organizations.

5.2 Unobservables

The second class of econometric concern is that our measures of management practice are correlated to unobserved factors captured in ϵ_{ijn} in (1), and these unobservables directly determine project completion rates thus biasing our coefficients of interest. To develop some intuition on the direction of bias, we consider the simplest case where we only condition on the constant term and the two CS- measures of interest in (1). It is then straightforward to show that the parameters of interest are functions of the variances of the two management practices, denoted $\sigma^2_{CS\text{-}a}$ and $\sigma^2_{CS\text{-}i}$, their covariance with each other ($\sigma_{CS\text{-}a,CS\text{-}i}$) and with the outcome ($\sigma_{CS\text{-}a,y}$, $\sigma_{CS\text{-}i,y}$):

$$\widehat{\gamma}_1 = \frac{\sigma_{CS\text{-}a,y}.\sigma^2_{CS\text{-}i} - \sigma_{CS\text{-}i,y}\sigma_{CS\text{-}a,CS\text{-}i}}{\sigma^2_{CS\text{-}a}.\sigma^2_{CS\text{-}i} - [\sigma_{CS\text{-}a,CS\text{-}i}]^2}. \tag{3}$$

$\widehat{\gamma}_2$ can be analogously defined. Given both CS- measures are standardized z-scores and their covariance is .24, the denominator in (3) is .9424. Substituting in for y in $\sigma_{CS\text{-}a,y}$ and $\sigma_{CS\text{-}i,y}$ in the numerator in (3), the bias depends on the following term:

$$\widehat{\gamma}_1 - \gamma_1 = cov(CS\text{-}a, \epsilon) - \sigma_{CS\text{-}a,CS\text{-}i}cov(CS\text{-}i, \epsilon). \tag{4}$$

To fix ideas we assume the unobserved factor is positively correlated with management practices. In this case the impact of CS-autonomy management practices on project completion rates is actually underestimated (so $\widehat{\gamma}_1 \leq \gamma_1$), and so in truth, even more positive than our baseline findings

suggest, if $\frac{cov(CS\text{-}a,\epsilon)}{cov(CS\text{-}i,\epsilon)} \leq \sigma_{CS\text{-}a,CS\text{-}i} = .2366$. Following a similar logic, the impact of CS-incentives management practices on project completion rates is actually overestimated ($\widehat{\gamma}_2 \geq \gamma_2$), and so, in truth, even more negative than our baseline findings suggest, if $\frac{cov(CS\text{-}i,\epsilon)}{cov(CS\text{-}a,\epsilon)} \geq \sigma_{CS\text{-}a,CS\text{-}i} = .2366$. Figure A1 shows the parts of the $(cov(CS\text{-}i,\epsilon), cov(CS\text{-}a,\epsilon))$ parameter space where each condition can be met. This makes precise that if both management practices are approximately equally positively correlated with the unobserved factor (so $cov(CS\text{-}a,\epsilon) \approx cov(CS\text{-}i,\epsilon) > 0$), $\widehat{\gamma}_1 > \gamma_1$ and we would overestimate the positive impact of CS-autonomy, but it would still be the case that $\widehat{\gamma}_2 > \gamma_2$ so that, in reality, the impact of CS-incentives is even more negative than estimated. For the remainder of the discussion, we classify such unobserved factors as arising from other features of organizations, management practices, and bureaucrats.

5.2.1 Organization Features

We address this concern in Table A7 by continuing to probe the data at the organizational level using specification (2) to check whether management practices in place correlate to other organizational outcomes beyond project completion rates. Two natural checks follow. First, in Column 3 we construct the average complexity of projects assigned to organization n as our dependent variable, and then regress this against our measures of management practice and other organizational characteristics. Neither CS- measure is significantly correlated to the average complexity of projects the organization is tasked to implement. Hence it is not the case that organizations with better practices related to autonomy are assigned easier projects, and those with more performance incentives are assigned harder to implement projects. Second, in Column 4 we use the log of the organization's aggregate budget as our dependent variable: again we find no correlation between the management practices in place and the resources provided to an organization. Hence it is not the case that organizations that have better practices towards their bureaucrats also command larger budgets overall, that might otherwise have indicated it was easier for the organization to complete their assigned projects.

5.2.2 Management Practices

Our management survey tool elicited management practices along nine dimensions in each organization: roles, flexibility, incentives, monitoring, culture, targeting, facilities, skills and staffing. We have thus far aggregated these into three measures. We now estimate a specification analogous to (1) but where we condition on the nine separate dimensions of management practice. This disaggregated specification serves two purposes: (i) by breaking up the CS-other measure into its components, it highlights if there are likely to be wider aspects related to how organizations are functioning, that the CS-autonomy and CS-incentives measures might pick up; (ii) it underpins the validity of the three-way classification of management practices in our baseline specification.

To begin with in Table A8 we focus on estimating the impact of each dimension separately. The roles and flexibility subcomponents that formed our CS-autonomy measure are individually positive and significantly different from zero (Columns 1 and 2). Columns 3 and 4 show that

each subcomponent that formed the CS-incentives index is negatively and significantly related to project completion rates. Columns 5 to 9 then split the remaining management practices that were previously subsumed within the CS-other measure. We see that most of these have no significant impact on project completion rates.

In Column 10 we simultaneously control for all nine CS- components, and find: (i) both roles and flexibility components remain positive and individually significantly correlated to project completion rates, with the impact of roles being significantly larger than for flexibility; (ii) the incentives component is negative but not significantly different from zero (the point estimate hardly changes from Column 3 but the standard error almost doubles), while the monitoring component remains negative and significantly correlated with project completion rates; (iii) the impacts of the other dimensions of management practice remain rather weak, with four out of five of these measures not being significantly different from zero.[23]

The fact that different elements of management practice have positive, zero, or negative impacts on project completion rates helps allay the concern that the CS- measures used in our core results simply pick up some unobserved element of management practice. If organizations that employed the most resourceful and skilled bureaucrats were also those best able to provide autonomy to their bureaucrats and adapt to new scenarios, then we would be concerned that the partial correlation between CS-autonomy and project completion rates was being partly confounded by these other (potentially unobserved) factors. That other dimensions of management practice related to culture or staffing are uncorrelated with project completion rates, suggests this is not the case.

In Column 11 we aggregate all nine indices into a single metric for 'good management', as considered by BVR for manufacturing firms. In our setting such aggregation leads to the wrong conclusions being drawn: we find no significant relationship between the aggregate CS-management measure and project completion rates. The reason for this is clear: the underlying components of the aggregate index do not all have the same signed impacts on project completion rates. Hence management of bureaucrats does matter, but different dimensions of management practice have very different impacts on outcomes in this civil service setting.

5.2.3 Bureaucrat Characteristics

If bureaucrats *sort* into organizations based on management practices in place, then the impacts of management practices are confounded by any direct relation between bureaucrat characteristics and project delivery. In Section 2 we highlighted there are frictions in the labor market for bureaucrats that limit the scope for individuals to freely join or move across organizations, and hence this also limits the ability of organizations to use management practices to attract certain types of bureaucrat. Such selection mechanisms have been argued to be important when evaluating the impacts of compensation schemes in private sector settings.

In Table A10 we present additional characteristics of bureaucrats elicited in our civil servant

---

[23]Another split of practices we considered was to subdivide the practices in CS-incentives into those that refer to rewards and punishments separately: we find conditional on CS-autonomy and CS-other measures, these subcomponents both have significantly negative partial correlations with project completion rates.

survey to shed further light on this issue. We aggregate each response to an organizational average. Column 1 shows the mean and standard deviation of the bureaucrat characteristic for the average organization. Columns 2 and 3 then show some of the regression coefficients from a specification analogous to (2), so estimated at the organization level, where the dependent variable is the organizational average for the reported bureaucrat characteristic, and the reported coefficients are $(\widehat{\theta}_1, \widehat{\theta}_2)$, the implied effect sizes from the CS-autonomy and CS-incentives management practices, and their corresponding robust standard errors in parentheses.

The first batch of characteristics probe further the specific social connections bureaucrats might have with senior staff in the organization to which they are appointed: almost no bureaucrats report knowing their direct boss, or other managers, prior to their appointment. The lack of variation in this response precludes it being regressed against the CS- practices. Where there is more variation is in terms of the number of close colleagues that were known to the individual prior to their assignment: on average, bureaucrats report knowing 4.40 other colleagues. However, we find this degree of linkage across bureaucrats, when averaged to the organization level, to be uncorrelated with the management practices in place related to autonomy and performance incentives, as shown in Columns 2 and 3.

Our next question uses a more subtle means through which to elicit information on personal connections among bureaucrats. We asked individuals whether they were the first family member to join the civil service: 48% of bureaucrats in the average organization report being the first in their family to be part of the civil service. We find no correlation between this measure of bureaucrat networks and the management practices in place.

5.3 Endogenous Management Practices

The final set of concerns we address relate to endogenous management practices. If for example some organization completes projects because they employ more dedicated civil servants, they might then endogenously choose to give those bureaucrats more autonomy, and be less reliant on the use of incentives/monitoring. On the other hand, in those organizations staffed by less reliable bureaucrats, the organization might choose to monitor and incentivize them along some margins (even if those turn out to be poorly targeted), and allocate them less autonomy. This is the econometric concern that most plausibly generates the pattern of positive *and* negative partial correlations in our core results. We use four strategies to make some headway on this issue.

Our first approach is to consider additional bureaucrat characteristics that might be indicative of the 'reliability' of bureaucrats, as shown in the lower half of Table A10. Throughout, we split bureaucrat characteristics into those of senior and low-tier bureaucrats, as these different tiers might differentially impact the management practices in place. Two points are of note: (i) there is no significant correlation between management practices related to autonomy or performance incentives and the average tenure of bureaucrats, the proportions of them that report being intrinsically motivated, or that report observing corrupt practices on projects; (ii) this is the case

for characteristics of senior and lower-tier bureaucrats.[24]

A second approach to measuring bureaucrat 'types' and examining whether the distribution of types in organizations correlate to the management practices in place, is based on responses to *vignettes* on scenarios bureaucrats face in service. For example, one vignette we use relates to a bureaucrat being faced with the following scenario: an official in an organization like theirs is told by her manager to take funds from a certified project and give them to a contractor/supplier for projects that the government has yet to complete due process on, against public service rules. The contractor then does not do the work he was paid for. We asked bureaucrats whether they agreed/disagreed with various statements related to the vignette: whether the bureaucrat acted correctly, whether the bureaucrat's manager acted correctly, whether bureaucrats should sometimes go against rules and so forth. We find little evidence of any significant relationship between the proportion of bureaucrats that agree with each such statement and the management practices in place in the organization. If such vignettes are informative of bureaucrat 'types', this evidence suggests that typically harder to observe measures of bureaucrat reliability are not much correlated with management practices in place.[25]

Our third approach identifies those organizations that have the greatest incentives to fine tune their management practices to maximize project completion rates. We think of these organizations as being those in which the completion of projects is relatively more important in the set of overall activities the organization is tasked with. We measure this using the share of the organization's total budget that is assigned to capital expenditures (such as those on OPEN-style projects). We then estimate a specification analogous to our baseline model in (1) and additionally interact each CS- measure with this measure of 'project importance' (for expositional ease, this interaction term is defined in terms of its deviation from mean), also conditioning on the organization's log total budget. If management practices are endogenously determined, their impacts might be attenuated for organizations in which it is more important to fine tune practices to maximize completion rates. The result in Column 1 of Table A10 finds no evidence of such heterogenous impacts.

Our final approach builds on the results in Figure 1 and identifies those organizations that have the most similar portfolio of projects to implement, as measured by the projects ambiguity. Such organizations might have the greatest ability to set their management practices optimally. For

---

[24]These results suggest that it is not the case that organizations limit the provision of performance-based incentives to attract intrinsically motivated workers [Delfgaauw and Dur 2010], or that the matching of intrinsically motivated workers to public sector organizations limits the need to use performance-based incentives [Besley and Ghatak 2005]. Both findings are partly driven by the rigidities in the assignment of bureaucrats to organizations.

[25]Enumerators stressed there were no right answers to the vignettes, and the question was not an exam, but an exercise in obtaining the opinions of bureaucrats. The vignette was phrased in reference to a bureaucrat in the third person (Ngozi). Officials were asked whether they agreed or disagreed with the following statements: (i) 'Ngozi should have refused to divert the funds and paid the contractors/suppliers/consultant for the certified project' (76% agreed); (ii) 'The contractors/suppliers/consultant should not have accepted the funds given that Due Process had not been followed' (68%); (iii) 'The boss was wrong to direct Ngozi to share the funds of a project with others' (81%); (iv) 'Sometimes it is right to go against strict rules, as Ngozi did, so to make things fairer or more efficient' (28%); (v) 'Bureaucrats should be given more space to interpret what the best action to take in a given situation is' (59%). We then check whether there is a significant correlation between the proportion of bureaucrats that agree with each of the five statements and the two management practices in place: we find that nine out of ten correlations are not significantly different from zero.

each organization we construct the variance in project ambiguity over all projects the organization is tasked to implement. We then estimate whether the impacts of the two CS- measures of interest are attenuated among those organizations that have the most similar projects in terms of their design and implementation complexity. The result in Column 2 of Table A10 shows there to be no evidence of such heterogenous impacts.

6   Discussion

We have studied whether management practices for bureaucrats correlate with effective public service delivery in an important developing country context: Nigeria.[26] We do so by combining novel project level data measuring the completion, quality and complexity of over 4700 projects implemented by various civil service organizations, with a management survey in each organization. Our primary contribution is thus to provide among the first evidence on whether and how the management of bureaucrats matters for public service delivery. The relevance of such investigations is first order, given the large number of developing countries engaged in reforming public bureaucracies along the lines of the 'good governance' agenda of the World Bank and United Nations [Goldfinch *et al.* 2012, Hasnain *et al.* 2012].

Our results confirm that the two dimensions of management practice we focus on: autonomy and performance incentives/monitoring, do indeed correlate to the quantity and quality of public services delivered, as emphasized by the public administration and economics literatures respectively. Our findings provide support to the notion that public agencies ought to delegate some decision making to bureaucrats, relying on their professionalism and resolve to deliver public services [Simon 1983]. The evidence is less supportive of the notion that when bureaucrats have more agency or organizations are more flexibly structured, then they are more likely to pursue their own objectives that diverge from societal interests, resulting in fewer public services being delivered.

The interpretation of the negative correlation of project completion rates with incentives/monitoring is more nuanced because we measure management *practices* related to the provision of incentives and monitoring: we do not necessarily measure the use of explicit compensation schemes based on project completion. Our extended results provide evidence consistent with two interpretations of why such management practices have detrimental impacts in this setting: bureaucrats operate in a multi-tasking environment, and the measured management practices capture an incentive system that places excessive regulatory burden or 'red tape' on bureaucrats (say through the use of mis-targeted key performance indicators), that has long been argued to lead bureaucrats to mis-allocate effort towards processing activities [Kelman 1990]. Alternatively, our management practices related to performance incentives and monitoring pick up elements of *subjective* performance evaluation, that lead to other dysfunctional responses among bureaucrats, especially engagement in influencing activities to curry favor with senior management.

To the extent that the day-to-day work of bureaucrats in other contexts is also characterized as

---

[26]We also add to a nascent literature utilizing the BVR approach to measure management practices in non-profits including hospitals [Bloom *et al.* 2013] and universities [McCormack *et al.* 2013].

multi-tasking environments in which processing and productive efforts can be exerted, our results will have some external validity to those settings. Moreover, there will nearly always be scope for the use of SPE in bureaucratic environments in which it is difficult to measure outputs, or to attribute outputs to the inputs of specific individuals. As such, our results sound a word of caution to the good governance agenda: the provision of incentives and monitoring will be effective only if such schemes are well tailored to the specific bureaucratic context. The simple import of practices from private sector settings might lead to dysfunctional responses among bureaucrats of the nature we document. Moreover, the broad design of management practices needs to be jointly considered with reforms to wider civil service practices along margins that might prevent such practices being fine tuned to their optimal designs, as discussed below.

Our results point to new directions for theoretical research to better understand the contracting environment in public bureaucracies, as well as highlighting specific areas in which better measurement of inputs and outputs can aid our understanding of public service provision in the developing world.[27] In this final Section we discuss the implications of our findings for understanding optimal management practices in public bureaucracies, and highlight two important directions for future research: on the methodology of collecting management practices for bureaucrats, and the interplay between such practices and other mechanisms to improve public service delivery.

6.1 Optimal Management Practices

Our core results suggest there are potentially huge gains to be made from marginal changes in management practices. This naturally begs the question of why civil service organizations are not optimizing over management practices to begin with. In our context, a fundamental source of this inefficiency arises from organizations being tasked to implement many project types (Table 2). As Figure 1 highlights, project types have very different characteristics and so it is unlikely that there exists a unique set of optimal management practices any organization could have in place. This lack of specialization in tasks across civil service organizations is an underlying reason why management practices still matter on the margin.

A second explanation is based on a Weberian view, that organizations might well be optimizing their management practices according to whatever is their true objective, and our evidence merely suggests this objective is only weakly aligned with maximizing project completion rates. To check for this, we have investigated whether the management practices correlate to the frequency with which bureaucrats report performing other functions, such as engaging with politicians, or engaging with citizens/civil society groups, but we find no such correlations. Moreover, we have previously noted that the type of projects we study constitute almost half the total budgets of these civil service organizations, so their completion is likely to be weighted to some extent in the objectives of the organization, but it is by no means their sole objective.[28]

---

[27]Dixit [2002] discusses that optimal management practices can differ in public and private sectors because of the existence of multiple principals, multiple tasks, a lack of competition, and motivated agents.

[28]To further assess the degree of alignment in organization with project completion rates, we regressed the log of project budgets on each of the 16 subcomponents of the project complexity indicator, as shown in Table A3. The

As discussed by BVR for private sector firms, suboptimal management practices might also persist in equilibrium despite the resultant loss of efficiency because: (i) the fixed costs of adopting better practices; (ii) best management practices might be heterogeneous across organizations. On the first point, to derive an estimate of the required fixed cost of changing management practices, we note the sum of capital expenditures for all projects from the 63 organizations studied is $3,694mn. Focusing on the impacts of management practices on those 38% of projects that never start, assuming a linear relationship between costs and completion rates, and taking the marginal impacts from Table 3 of a one standard deviation change in each management practice, implies there would need to be fixed cost per organization of around $(.38 \times .32 \times 3,694)/63 =7.13mn for such a move not to occur for this reason alone.[29]

The second point has been examined when we explored the heterogeneous impacts of management practices in Section 4.2. Those results imply the optimal performance incentives in place should partly reflect the complexity and ambiguity of projects in an organization's portfolio, the use of IT facilities, as well as the tenure and intrinsic motivation of its bureaucrats. However, in contrast to this evidence, these issues were not at the fore during the semi-structured interviews we conducted at organizations to better our understanding of what drives management practices in reality. These all highlighted how practices evolve slowly over time as a function of ground rules laid out in the Public Service Rules of the Nigerian civil service, the history of senior management staff in an organization, and external events such as trade union demands. As further discussed in BVR, inefficient management practices might also persist for dynamic reasons: learning and adjustment costs might cause best practice to diffuse over time. This is in line with the evidence discussed throughout on the frictions in the labor market for bureaucrats related to the initial assignment and immobility across organizations of bureaucrats.

Finally, a particularly acute concern is that a lack of competitive pressure enables poorly managed public sector organizations to survive. This might especially be true in developing country contexts where mechanisms are rarely in place to allow citizens to choose across alternative public providers of a given good or service. We return to this issue below when discussing alternative mechanisms through which public service delivery can be improved.[30]

6.2 Other Mechanisms To Improve Public Sector Delivery

Our analysis fits within a broader literature documenting mechanisms to improve public service delivery in developing countries. There are at least two additional mechanisms that can interplay with the management practices for bureaucrats within organizations. First, there remains much

---

residuals from this regression, that might capture the rents to be gained from the project if it is completed, are found to be weakly positively correlated with actual project completion rates, with a correlation of .13.

[29]This is a lower bound estimate because, as discussed in the Appendix, the impacts of management practice on project completion rates apply not only at the margin of ensuring projects are started, but have similar impacts at all thresholds of project completion rate.

[30]In the private sector, Bloom *et al.* [2012] provide evidence that product market competition drives innovations towards more better management practices. Bloom *et al.* [2013] find evidence that hospitals that face competition for patients from rival hospitals do indeed adopt better management practices.

scope to combine and jointly test models of *selection* into, and turnover of bureaucrats *between*, public sector organizations. On selection into the civil service, our findings highlight there can be an interplay between the intrinsic motivation of bureaucrats and how they respond to management practices. On turnover, bureaucratic tenure is important both because longer serving bureaucrats shape the management practices in place, but they might also react differently to any given set of incentives, as highlighted in Table 4. A recent contribution in this direction is Dal Bo *et al.* [2013] who present evidence from Mexico, exploiting experimental variation in salaries to identify their impact on the selection of public sector officials. They find higher wages attracted more able workers, and that there are no adverse selection impacts in terms of motivation. This evidence complements our findings on the impact of management practices once bureaucrats are hired.

Second, the role of grass roots monitoring or accountability of public service providers is important and likely interplays with management practices in place. Bjorkman and Svensson [2009] present evidence from an RCT in Uganda on how community based monitoring of primary health care providers led to large increases in utilization of services and improved health outcomes: in part these findings are driven by changes in the behavior of health providers themselves. Reinikka and Svensson [2011] and Duflo *et al.* [2012] provide other examples of the gains to public service delivery to be had from bottom-up monitoring of frontline public sector workers. On the other hand, such findings are not uniform: Olken [2007] finds more muted impacts of community based monitoring on reducing corruption in road infrastructure projects in Indonesia. Rather, top-down audits are effective in reducing corruption on such locally organized public works projects. Along the same lines of top-down monitoring, Besley and Burgess [2002] and Ferraz and Finan [2008] have documented the significant role that media scrutiny can play for public service provision.[31]

We view there to be a rich future agenda for understanding public service delivery – in rich and poor countries – that effectively links incentives and practices in civil service organizations, the selection and retention of bureaucrats, and public monitoring and accountability of organizations.

A   Appendix

A.1   Measuring Project Complexity

Data on the complexity of government projects is not directly part of the OPEN data set. We thus worked with a pair of Nigerian engineers familiar with the OPEN projects and a number of international researchers working on technical complexity to define a relevant set of indicators based on the technical documents for each OPEN project. We followed the perspectives on complexity suggested by Remington and Pollack [2007], by asking the engineer-assessors to individually assess projects along the following five topics, each with their own set of indicators.

---

[31]Evidence from our civil servants survey confirms that bureaucrats do regularly engage with community groups and so such bottom-up approaches might well interplay with management practices in place. For example, 24% of surveyed bureaucrats reported personally engaging with community/religious groups; 11% reported engaging with members of the National Assembly.

Structural complexity stems from the scale of different interconnected tasks and activities. The indicators associated with this topic capture structural aspects such as project size and the number of inputs required for production. They also capture issues in raw material and labour supply, and the ease with which any necessary specialized skills and equipment can be sourced. *Temporally complex* projects are those whose production involves uncertainties. Hence there are indicators for uncertainties in design and implementation. *Technically complex* projects are those whose production have ambiguous risks, namely their uncertainties are not well understood. Hence some indicators capture ambiguities in design and implementation. *Directional complexity* refers to the potential for preferences over the project to diverge. The engineer assessors are thus asked to rate the managerial complexities of the project. Finally, there is a subjective assessment as to the overall complexity of the project. This allows any unassessed aspects of complexity to be measured and provides a coherent picture of project complexity.

Two qualified and independent Nigerian engineers were then contracted to assess each project in the OPEN data set along these margins. The process of aggregation between engineers used in this project aimed to build a consensus. The first engineer coded indicators for the entire data set. The codings of the first engineer were then provided to the second engineer who then constructed his own codings with reference to the codings of the first. The aim was to anchor the coding of the second engineer in that of the first but give him freedom to disagree where he felt the coding was incorrect. Other methods would have been to have them code independently and average the two data sets or to have them work together. We decided our approach was a balance between consensus and subjectivity.

The two engineers were provided with project details and documents and asked to code a value for each indicator. The documents only contained information available *before* implementation such that there was no bias from the coding being done after the projects were implemented.

Table A3 provides descriptive statistics for all 16 indicators from which the complexity index is constructed, as well as how each is correlated with the other indicators. Aggregate complexity is a subjective assessment of the overall complexity of the projects by the two engineers, that includes 'all factors that might influence the difficulty of implementing the project, not only those assessed [by the other indicators]'. We asked the engineers to take the distribution of complexity in the OPEN data set as a whole, with the least complex project in the data having an aggregate complexity of zero and the most complex project having an aggregate complexity of 100, and place each project within this distribution.

We undertook a number of measures to check the complexity of the OPEN indicators coded by the engineers. First, we inserted 200 randomly chosen repeated projects into the data set provided to the engineers. Since the project characteristics of the original and repeat projects are identical, we would expect that the codings of the two sets of projects would be similar. Reassuringly, we find that in general the original and duplicate projects are coded in similar ways. We compare the differences between these two sets by looking at group and paired means, and distributional tests for each variable. The differences are only statistically significant at conventional levels in a few cases, and the magnitude of the differences are relatively small. For example, the only variable

that is statistically significantly different below the 10% level in the mean-comparison t-test relates to raw material storage. Here, despite a standard deviation of 0.2 in the originals, the difference is 0.07 between the originals and the duplicates.

Second, we looked at the similarity of the codings of the two engineers. We find that the second engineer's codings are not dramatically different from the first engineer's efforts. Whilst there are a small number of differences, they are relatively small and rarely significant, indicating that the re-coding left the overall picture relatively stable.

Finally, over a year after he had completed the prompted codings, we asked the second engineer to re-code a sub-sample of projects from scratch, this time without prompting. The differences between these independent codings and the consensus data we rely on are again relatively minor. It seems that once he had become accustomed to the broad parameters of the coding framework, the second engineer's coding was not dissimilar to the consensus generated by the two engineers working one after the other.

We therefore have evidence of similar projects within the data set being coded in a similar way, of the two engineers coding in similar ways both when prompted and unprompted, and when there were deviations, of the deviations not being particularly quantitatively large. Taken together these checks reassure us that the complexity measures pick up meaningful variation across projects, rather than merely picking up noise that should have led to the multiple reports (either across engineers or over time) being uncorrelated.

A.2   Standard Errors

Our baseline specification assumes the disturbance term is clustered by project type-organization. This might capture best the nature of unobserved factors that relate to project completion rates, such as the quality of the sub-departments in each organization that are tasked with the same project types. In Table A4 we show the robustness of our results to alternative assumptions on the error structure. The table is structured to cover the same specifications as in Columns 1 to 4 of Table 3, but where we now report standard errors based on alternative assumptions: (i) robust standard errors; (ii) errors clustered by project-type and organization as in Table 3; (iii) standard errors clustered by organization. One concern with such clustered standard errors is that they may be downwards biased when the number of clusters is small (and in our specification the number of clusters corresponds to 63) [Cameron *et al.* 2008]. They propose various asymptotic refinements using bootstrap techniques, finding the wild cluster bootstrap-t technique performs particularly well in Monte Carlo simulations. We have implemented this method on our baseline specifications and show in brackets in Columns 1 to 4 the resulting p-values. This correction does not alter the significance of any of the coefficients from our preferred specification with project fixed effects shown in Column 4 of Table A3.

A.3 Robustness Checks

We conduct a range of robustness checks on our core result on project completion rates, shown in Column 4 of Table 3. To begin with we note that the results in Table 3 used our continuous measure of project completion rates (from zero to one) as the dependent variable. However, an alternative approach is to define a threshold of completion that would deem the project usable. To do so we consider all potential thresholds from 1% to 100% in increments of 1% and then estimate a specification analogous to (1) where y_{ijn} is defined as a dummy variable equal to one if the project completion rate is above the given threshold $p\%$, and zero otherwise. For any given threshold p the coefficients of interest are denoted γ_{1p} and γ_{2p}. Figure A2 then plots each pair of $(\widehat{\gamma}_{1p}, \widehat{\gamma}_{2p})$ estimates, for each threshold and their associated 95% confidence interval.

Two points are of note. First, at the extreme left of the figure where we consider a 1% threshold, we are essentially using a linear probability model to assess the relationship between management practices in civil service organizations and whether projects are *started* in some way. Here we find marginal impacts of each type of management practice to be qualitatively similar to those documented earlier for the average completion rate, although the point estimate on CS-autonomy seems to be lower than the baseline specification.[32] Second, we cannot reject the null that the sign and significance of the coefficients are the same for every threshold of project completion. Taken together, the results imply that managerial incentives along both margins have similar impacts on the extensive margin of public service delivery (namely whether projects are started at all) and the intensive margins of project completion (namely the extent to which projects are completed).

The checks detailed in Table A5 all verify the robustness of our main finding to alterative samples of organizations and projects: Column 1 excludes those projects implemented by the largest organization in terms of total expenditures. Column 2 excludes projects implemented by the largest organization in terms of total expenditures. Columns 3 and 4 remove the 10 smallest organizations by expenditures and number of projects respectively. Columns 5 and 6 exclude organizations at the top and bottom of the CS-autonomy and CS-incentives management scales respectively. In each cut of the data, the core results remain stable, so that our findings do not appear to be drive by outlier or specific organizations.

The third series of robustness checks in all probe the core results in alternative subsamples of data (with the caveat that such splits of the data inevitably lead to less precise estimates). Columns 7 and 8 in Table A5 break down the project types listed in Table 1 into two groups: those related to construction projects (borehole, building, electrification, dam, road and canal) and those related to non-construction projects. We find that for construction projects, the results closely replicate the baseline findings. Indeed, for construction projects as a whole the point estimates on each management practice index is slightly larger in absolute value than the baseline result in Column 4 of Table 3. For non-construction projects, the results show that CS- autonomy

---

[32]The result needs to be interpreted carefully. There is not continuous mass in project completion rates over the [0,1] interval: hence we do not expect the marginal impact of the each management practice to be sensitive to marginal changes in threshold p where there is little mass in project completion rates. This explains why the marginal impacts shown in Figure A2 jump at a small number of points.

continues to have a positive and significant impact on project completion rates; CS-incentives has a negative correlation with project completion rates that is significant at the 10% level. As suggested by Figure 1, this might be because non-construction projects such as training and procurement, are inherently less ambiguous in design.

The next split considered in Columns 9 and 10 is between centralized and decentralized civil service organizations, where centralized agencies correspond to government ministries. We see the impacts of CS-autonomy to be similar in both, although the CS-index index only has negative impacts in the centralized agencies. This latter result is worth exploring in future work.

The next robustness check addresses the concern that different management practices might be spread across the country in a way that is correlated with characteristics of the organization's local geographies. Indeed, it is well understood that the characteristics of local populations interplay with them being able to solve collective action problems, and thus are an important driver of public goods provision [Banerjee *et al.* 2007]. To address this issue, Column 11 additionally controls for a wide variety of state-level controls for each project and finds almost no change in the coefficients of interest. In short, local area characteristics do not seem to be driving our results.[33] Columns 12 and 13 split the sample into projects that are located in the Northern and Southern regions of Nigeria respectively, that characterizes the first order cultural divide in Nigeria, and provides a relatively even split in the number of projects. We find the impacts of both management practices to remain significant in both regions. This suggests that cultural differences across regions do not much interact with responses to these management practices.

The final two robustness checks relate to methodological issues. First, we consider alternative constructions of the CS- management practice indices. As described in Section 3.3, following BVR, we aggregated responses to individuals questions to construct our indices of management practices giving equal weight to all questions. A natural alternative is to cluster the variables into the various management topics described in Table A2 and weight each *topic* (rather than each variable) equally. We re-construct our CS-autonomy and CS-incentives measures along these lines and reestimate our preferred specification (1). The result, in Column 14 of Table A5 is qualitatively in line with our baseline results, although the absolute magnitude of each measure of public sector management is *larger*: a one standard deviation increase in CS-autonomy corresponds to a significantly higher project completion rate of 23%, and a one standard deviation increase in CS-incentives corresponds to a significantly lower project completion rate of 14%.

The analysis has so far estimated (1) using OLS. The final robustness check estimates this specification using a fractional regression model that accounts for the dependent variable being a continuous variable between zero and one. To do so, we utilize Papke and Wooldridge's [1996] fractional logit model in which the conditional expectation function is modelled as a logistic function that can take all values in the unit interval. The interpretation of the marginal effects are the same as in the binary logit model and evaluated at sample averages, the partial effects are approximately comparable to the coefficients from a linear regression. The result in Column 15 of

---

[33]The sample drops slightly in this specification because the state in which the project is located (or should have been located for those projects that are never started) is missing for around 450 projects.

Table A5 shows our core findings to be robust to this alternative estimation model.

A.4 The Assignment of Projects to Organizations

A central econometric concern addressed in Section 5.1 is that the assignment of projects to organizations might correlate to the management practices in place. For example, projects for which there is concerted political pressure to complete might be allocated to better managed organizations, all else equal. To investigate the issue we use a conditional logit model to directly estimate the likelihood of project i being assigned to organization n conditional on the management practices for bureaucrats in place in the organization, and other project and organizational characteristics exploited in our analysis.

To do so we first reshape our data as follows: for each project we created a binary variable with 63 values corresponding to our 63 organizations. The variable, denoted D_{in}, takes the value one for the organization at which that project is actually assigned, and zero otherwise. Thus, the data-set is at the project-organization pair level (in), with a total of $4721 \times 63 = 297,423$ project-organization paired observations. To each observation, we attach the relevant organization-level characteristics used in our analysis (that were denoted OC_n in (1)), such as our management indices, capital controls and organizational averages of bureaucrat characteristics. We then also consider whether specific project-organization interactions, denoted Z_{in}, correlate with the assignment of projects to organizations. We estimate a conditional logit specification for $\text{Prob}(D_{in} = 1)$, based on both sets of characteristics;

$$prob[D_{in} = 1] = \frac{\exp(\beta' OC_n + \gamma' Z_{in})}{\sum_n \exp(\beta' OC_n + \gamma' Z_{in})}. \tag{5}$$

Note that in this modelling framework project characteristics play no role as these do not vary within a given project i over the organizations n it could potentially have been assigned to.

We run three sets of specifications. The first takes the perspective that each project could have been assigned to *any* of our 63 organizations. The second takes the perspective that projects can only be implemented by organizations of the same sector. Thus, health projects could only be implemented by health sector organizations for example. When we impose this restriction on the permissible project-organization pairs, we lose 847 projects as there is only one organization of that sector in our data, and there are $44,429$ potential within sector project-organization pairs. The third specification further imposes the assumption that projects can only potentially be assigned to organizations in the same sector that are observed being tasked to implement a project of the same complexity. This further reduces the sample to $33,480$ feasible project-organization pairs. In all specifications we cluster standard errors by organization as there are likely to be unobserved characteristics of organizations that determine project assignments.

Table A6 presents the results. Column 1 utilizes the entire set of project-organization combinations and uses only our management scores in the regression. Unconditional on other organization characteristics, neither management practice related to autonomy or performance-based incentives

significantly predicts the assignment of projects to organizations. When we restrict the sample so that each project can only feasibly be assigned to organizations in the same sector, Column 2 shows the coefficient on CS-autonomy to become significant at the 10% level, but this result is not robust. In Column 3 where we restrict the feasible pairs also using information on the complexity of projects, we find the probability of projects being assigned to organizations do not significantly correlate with the management practices in place in the organization.

Column 4 extends the specification to include all the organizational controls utilized in our baseline specification in Table 3 (capital, general and noise), as well as the full set of organizational controls that we use elsewhere in the paper. The coefficients on the management practices remain insignificantly different from zero at the usual levels. Moreover, we find no evidence that any of these other organizational characteristics predict the assignment of a given project to that particular organization rather than other organizations it could feasibly have been assigned to.

In Column 5 we additionally control for a series of interactions between project characteristics (scale, complexity) and organizational characteristics (total staff, total budget). Some of these project-organization interactions do predict the assignment of projects. As is intuitive, we find that the interactions between the number of staff at an organization and the project budget, and the organization's total budget and the project complexity, both are positive and significant predictors of project assignment. Over and above these interactions, we continue to find no impact of management practices on project assignment even once we allow for specific matched pairs to be assigned in this way.

Overall, these results suggest our results are not reflecting the non-random selection of projects to organizations based on their management practices. While there is no doubt some complex bargaining process between Parliament, civil service organizations and other stakeholders that determined the assignment of projects to locations and organizations, on the margin, this assignment is uncorrelated with the management practices in place in implementing organizations.

References

[1] AGHION.P AND J.TIROLE (1997) "Real and Formal Authority in Organizations," *Journal of Political Economy* 105: 1-29.

[2] ALABI.M.O.A AND J.Y.FASHAGBA (2010) "The Legislature and Anti-corruption Crusade under the Fourth Republic of Nigeria: Constitutional Imperatives and Practical Realities," *International Journal of Politics and Good Governance* 1.

[3] ASHRAF.N, O.BANDIERA AND K.JACK (2012) No Margin, No mission? A Field Experiment on Incentives for Pro-social Tasks, mimeo, London School of Economics.

[4] BAKER.G.P (2002) "Distortion and Risk in Optimal Incentive Contracts," *Journal of Human Resources* 37: 728-51.

[5] BANDIERA.O, A.PRAT AND T.VALLETTI (2009) "Active and Passive Waste in Government Spending: Evidence from a Policy Experiment," *American Economic Review* 99: 1278-308.

[6] BANERJEE.A, L.IYER AND R.SOMANATHAN (2007) "Public Action for Public Goods," in *Handbook of Development Economics* Vol. 4, T.P.Schultz and J.Strauss (eds.), Amsterdam: Elsevier.

[7] BENABOU.R AND J.TIROLE (2006) "Intrinsic and Extrinsic Motivation," *Review of Economic Studies* 70: 489-520.

[8] BESLEY.T (2004) "Paying Politicians: Theory and Evidence," *Journal of the European Economic Association* 2: 193-215.

[9] BESLEY.T AND R.BURGESS (2002) "The Political Economy of Government Responsiveness: Theory and Evidence from India," *Quarterly Journal of Economics* 117: 1415-52.

[10] BESLEY.T AND M.GHATAK (2005) "Competition and Incentives with Motivated Agents," *American Economic Review* 95: 616-36.

[11] BESLEY.T AND T.PERSSON (2010) "State Capacity, Conflict and Development," *Econometrica* 78: 1-34.

[12] BJORKMAN.M AND SVENSSON.J (2009) "Power to the People: Evidence from a Randomized Field Experiment of a Community-Based Monitoring Project in Uganda", *Quarterly Journal of Economics* 124: 735-69.

[13] BLACK.S AND L.LYNCH (2001) "How to Compete: The Impact of Workplace Practices and Information Technology on Productivity," *Review of Economics and Statistics* 88: 434-45.

[14] BLOOM.N AND J.VAN REENEN (2007) "Measuring and Explaining Management Practices Across Firms and Countries," *Quarterly Journal of Economics* 122: 1351-1408.

[15] BLOOM.N AND J.VAN REENEN (2010) "New Approaches to Surveying Organizations," *American Economic Review* 100: 105-9.

[16] BLOOM.N, R.SADUN AND J.VAN REENEN (2012) "The Organization of Firms Across Countries," *Quarterly Journal of Economics* 127: 1663-1705.

[17] BLOOM.N, B.EIFERT, A.MAHAJAN, D.MCKENZIE AND J.ROBERTS (2013) "Does Management Matter: Evidence from India," *Quarterly Journal of Economics* 128: 1-51.

[18] BLOOM.N, C.PROPPER, S.SEILER AND J.VAN REENEN (2013) The Impact of Competition on Management Practices in Public Hospitals, mimeo, Stanford University.

[19] BUURMAN.M, J.DELFGAAUW, R.DUR AND S.VAN DEN BOSSCHE (2012) "Public Sector Employees: Risk Averse and Altruistic?," *Journal of Economic Behavior and Organization* 83: 279-91.

[20] CAMERON.C, J.GELBACH AND D.MILLER (2008) "Bootstrap-Based Improvements for Inference with Clustered Errors", *Review of Economics and Statistics* 90: 414-27.

[21] DAL BO.E, F.FINAN AND M.ROSSI (2013) "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service," forthcoming, *Quarterly Journal of Economics*.

[22] DELFGAAUW.J AND R.DUR (2010) "Managerial Talent, Motivation, and Self-Selection into Public Management," *Journal of Public Economics* 94: 654-60.

[23] DIJKSTRA.G, B.AKANJI, C.HIDDINK, S.SANGARABALAN AND F.DE MEVIUS (2011) *Mutual Interests - Mutual Benefits: Evaluation of the 2005 Debt Relief Agreement Between the Paris Club and Nigeria,* The Hague: Ministry of Foreign Affairs of the Kingdom of the Netherlands.

[24] DIXIT.A (2002) "Incentives and Organizations in the Public Sector: An Interpretive Review," *Journal of Human Resources* 37: 696-727.

[25] DUFLO.E, R.HANNA AND S.P.RYAN (2012) "Incentives Work: Getting Teachers to Come to School," *American Economic Review* 102: 1241-78.

[26] EBOH.E (2010) *MDGs-based Planning in Africa: Lesson, Experiences and Challenges: A Case Study of Nigeria*, Addis Ababa: United Nations Economic Commission for Africa.

[27] FEDERAL GOVERNMENT OF NIGERIA (2007) *The Story of OPEN*, Federal Government of Nigeria.

[28] FEDERAL GOVERNMENT OF NIGERIA (2008) *Monitoring and Evaluation Report of the DRG-Funded MDG Projects and Programmes in Nigeria 2006*, Federal Government of Nigeria.

[29] FEDERAL GOVERNMENT OF NIGERIA (2009) *Monitoring and Evaluation Report of the DRG-Funded MDG Projects and Programmes in Nigeria 2007*, Federal Government of Nigeria.

[30] FERRAZ.C AND F.FINAN (2008) "Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes," *Quarterly Journal of Economics* 123: 703-45.

[31] FLODGREN.G, M.P.ECCLES, S.SHEPPERD, A.SCOTT, E.PARMELLI AND F.R.BEYER (2011) "An Overview of Reviews Evaluating the Effectiveness of Financial Incentives in Changing Healthcare Professional Behaviours and Patient Outcomes," *Cochrane Database of Systematic Reviews* Issue 7, Art. No.: CD009255.

[32] FRANCOIS.P AND M.VLASSOPOULOS (2008) "Pro-social Motivation and the Delivery of Social Services," *CESifo Economic Studies* 54: 1-33.

[33] FRYER.R.G (2013) "Teacher Incentives and Student Achievement from New York Public Schools," *Journal of Labor Economics* 31: 373-407.

[34] GAGLIARDUCCI.S AND T.NANNICINI (2013) "Do Better Paid Politicians Perform Better? Disentangling Incentives from Selection," *Journal of the European Economic Association* 11: 369-98.

[35] GLEWWE.P, N.ILIAS AND M.KREMER (2010) "Teacher Incentives," *American Economic Journal: Applied Economics* 2: 205-27.

[36] GOLDFINCH.S, K.DEROUEN AND P.POSPIESZNA (2012) "Flying Blind? Evidence for Good Governance Public Management Reform Agendas, Implementation and Outcomes in Low Income Countries," *Public Administration and Development* 33: 50-61.

[37] HASNAIN.Z, N.MANNING AND J.H.PIERSKALLA (2012) Performance-related Pay in the Public Sector, World Bank Policy Research Working Paper 6043.

[38] HOLMSTROM.B AND P.MILGROM (1991) "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design", *Journal of Law, Economics and Organization* 7: 24-52.

[39] ICHNIOWSKI.C, K.SHAW AND G.PRENNUSHI (1993) "Effects of Human Resource Management Practices," *American Economic Review* 87: 291-313.

[40] KELMAN.S (1990) *Procurement and Public Management: The Fear of Discretion and the Quality of Public Performance*, Washington DC: American Enterprise Initiative.

[41] MCCORMACK.J, S.SMITH AND C.PROPPER (2013) Herding Cats? Management and University Performance, CMPO Working Paper 13/308.

[42] MILGROM.P.R (1988) "Employment Contracts, Influence Activities, and Efficient Organization Design", *Journal of Political Economy* 96: 42–60.

[43] MILGROM.P.R AND J.ROBERTS (1988) "An Economic Approach to Influence in Organizations", *American Journal of Sociology* 94: S154-79.

[44] MILLER.G AND K.S.BABIARZ (2013) Pay-For-Performance Incentives in Low- and Middle-Income Country Health Programs, NBER WP 18932.

[45] MURALIDHARAN.K AND V.SUNDARARAMAN (2011) "Teacher Performance Pay: Experimental Evidence from India," *Journal of Political Economy* 119: 39-77.

[46] MURALIDHARAN.K (2012) Long-Term Effects of Teacher Performance Pay: Experimental Evidence from India, mimeo UCSD.

[47] NKONJO-IWEALA.N AND P.OSAFO-KWAAKO (2007) Nigeria's Economic Reforms: Progress and Challenges, Brookings Institute Global Working Papers Washington: Brookings Institute.

[48] OLKEN.B (2007) "Monitoring Corruption: Evidence from a Field Experiment in Indonesia," *Journal of Political Economy* 115: 200-49.

[49] OGUNDIYA.I.S (2011) "Anti-corruption Reforms in Nigeria: Challenges and Failures," in I.S.Ogundiya, S.Ilufoye, A.Olanrewaju Olutayo and J.Amzat (eds.) *Assessment Of Democratic Trends In Nigeria,* Delhi: Gyan Publishing House.

[50] PAPKE.L.E AND J.M.WOOLDRIDGE (1996) "Econometric Methods for Fractional Response Variables with an Application to 401(k) Plan Participation Rates," *Journal of Applied Econometrics* 11: 619-32.

[51] PERRY.J.L AND L.R.WISE (1990) "The Motivational Basis of Public Service," *Public Administration Review* 50: 367-73.

[52] PERRY.J.L (1996) "Measuring Public Service Motivation: An Assessment of Construct Reliability and Validity," *Journal of Public Administration Research and Theory* 6: 5-2.

[53] PERRY.J.L, T.A.ENGERS AND S.Y.JUN (2009) "Back to the Future? Performance-related Pay, Empirical Research, and the Perils of Persistence," *Public Administration Review* 69: 39-51.

[54] PRENDERGAST.C (2002) "The Tenuous Trade-Off between Risk and Incentives", *Journal of Political Economy* 110: 1071-102.

[55] REINIKKA.R AND J.SVENSSON (2011) "The Power of Information in Public Services: Evidence from Education in Uganda," *Journal of Public Economics* 95: 956-66.

[56] REMINGTON.K AND J.POLLACK (2007) *Tools for Complex Projects*, Aldershot: Gower.

[57] ROSE-ACKERMAN.S (1986) "Reforming Public Bureaucracy Through Economic Incentives?," *Journal of Law, Economics and Organization* 2: 131-61.

[58] SCOTT.A, P.SIVEY, D.AIT OUAKRIM, L.WILLENBERG, L.NACCARELLA, J.FURLER AND D.YOUNG (2011) "The Effect of Financial Incentives on the Quality of Health Care Provided by Primary Care Physicians," *Cochrane Database of Systematic Reviews* Issue 9. Art.No.: CD008451.

[59] SIMON.H (1951) "A Formal Theory of the Employment Relation," *Econometrica* 19: 293-305.

[60] SIMON.W (1983) "Legality, Bureaucracy, and Class in the Welfare System," *Yale Law Journal* 92: 1198-269.

[61] TULLOCK.G (1965) *The Politics of Bureaucracy*, Washington: Public Affairs Press.

[62] WILSON.J (1989) *Bureaucracy: What Government Agencies Do and Why They Do It*, Basic Books: New York.

Table 1: Descriptive Evidence on Project Types

Project Type	(1) Number of Projects [Proportion]	(2) Number of Implementing Organizations	(3) Median Budget Allocation (US$000s)	(4) Proportion Never Started	(5) Average Completion Rate	(6) Proportion Completed Conditional on Being Started	(7) Proportion Fully Completed	(8) Proportion With Satisfactory Quality
Borehole	1348 [0.29]	18	29	0.44	0.47	0.84	0.37	0.85
Building	806 [0.17]	32	120	0.37	0.50	0.79	0.34	0.81
Electrification	751 [0.16]	2	93	0.14	0.56	0.65	0.25	0.87
Dam	624 [0.13]	14	18	0.79	0.15	0.74	0.10	0.50
Procurement	345 [0.07]	41	87	0.30	0.58	0.83	0.47	0.85
Road	217 [0.05]	4	167	0.12	0.52	0.59	0.22	0.79
Training	189 [0.04]	26	80	0.20	0.60	0.74	0.42	0.84
Financial project	157 [0.03]	8	17	0.38	0.49	0.79	0.35	0.84
Research	122 [0.03]	21	67	0.11	0.63	0.72	0.52	0.99
Advocacy	86 [0.02]	23	49	0.24	0.61	0.80	0.47	0.94
Canal	76 [0.02]	12	347	0.70	0.14	0.45	0.05	0.92

Notes: The "project type" classification refers to the primary classification for each project. Other project classifications exist. The median budget allocation in Column 3 is in thousands of US Dollars (assuming an exchange rate of US$1: Naira 150). The sample of projects covers those which have a positive budget allocation and for which the proportion completed evaluation variable and management scores are available. The project quality variable in Column 8 is not available for all projects. Standard deviations are in parentheses. Figures are rounded to two decimal places where relevant.

Table 2: Descriptive Evidence on Largest Civil Service Implementing Organizations

Civil Service Organization	(1) Number of Projects	(2) Number of Unique Project Types	(3) Budget Allocation (US$mn)	(4) Proportion Never Started	(5) Proportion Completed	(6) Proportion Completed Conditional on Being Started	(7) Proportion Fully Completed	(8) Proportion With Satisfactory Quality
Federal Ministry of Agriculture and Rural Development	797	9	144	0.54	0.29	0.63	0.14	0.76
Federal Ministry of Power and Steel	750	1	490	0.14	0.56	0.25	0.65	0.87
Federal Ministry of Water Resources	520	4	426	0.95	0.04	0.77	0.03	0.69
National Primary Health Care Development	447	4	56	0.19	0.64	0.79	0.42	0.75
Sokoto Rima River Basin Development Authority	277	2	23	0.22	0.66	0.85	0.51	0.76
Upper Benue River Basin Development Authority	169	3	13	0.11	0.89	1.00	0.89	0.25
Ogun/Oshun River Basin Development Authority	165	4	22	0.55	0.32	0.71	0.24	0.89
Chad Basin River Basin Development Authority	148	3	16	0.43	0.56	1.00	0.56	1.00
Lower Benue River Basin Development Authority	143	3	16	0.45	0.42	0.77	0.17	0.86
Nigerian Agricultural Cooperative and Rural Development Bank	133	2	8	0.42	0.46	0.80	0.34	0.81

Notes: The sample covers the ten largest civil service organizations ranked by number of projects from our overall sample of projects. The "project type" classification refers to the primary classification for each project. Other project classifications exist. The budget allocation in Column 3 is in millions of US Dollars (assuming an exchange rate of US$1: Naira 150). The sample of projects covers those which have a positive budget allocation and for which the proportion completed evaluation variable and management scores are available. The project quality variable in Column 8 is not available for all projects. Standard deviations are in parentheses. Figures are rounded to two decimal places where relevant.

Table 3: Management Practices and Public Sector Service Delivery

Standard Errors: Clustered by Project Type Within Organization

OLS Estimates

	Project Completion Rates					
	(1) Unconditional	**(2) Organization Controls**	**(3) Project Controls**	**(4) Project Type Fixed Effects**	**(5) Quality-Adjusted Completion Rate**	**(6) Interaction**
CS-Autonomy	0.11**	0.18***	0.17***	0.18***	0.11**	0.23***
	(0.05)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)
CS-Incentives	-0.06*	-0.11***	-0.11***	-0.14***	-0.08***	-0.15***
	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
CS-Other	0.10***	0.05	0.05	0.08***	0.08***	0.08***
	(0.04)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
CS-Autonomy x CS-Incentives						0.06**
						(0.02)
Organization Controls (capital, general, noise)	No	Yes	Yes	Yes	Yes	Yes
Project Controls	No	No	Yes	Yes	Yes	Yes
Fixed Effects	None	None	None	Project Type	Project Type	Project Type
Observations (clusters)	4721 (201)	4721 (201)	4721 (201)	4721 (201)	2206 (144)	4721 (201)

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. Standard errors are in parentheses, and are clustered by project type within organization throughout. All columns report OLS estimates. The dependent variable in Columns 1 to 4 and 6 is the proportion of the project completed (that is a continuous measure between zero and one). The dependent variable in Column 5 is the product of the proportion completed variable and the dummy variable for quality. The sample of projects in Column 5 is limited to those for which project completion and quality data is available. Project Type fixed effects relate to whether the primary classification of the project is as a financial, training, advocacy, procurement, research, electrification, borehole, dam, building, canal or road project. Project controls comprise project-level controls for the project budget, whether the project is new or a rehabilitation, and an assessment of its aggregate complexity by Nigerian engineers. Capital controls comprise organization-level controls for the logs of number of employees, total budget, and capital budget. General controls comprise organization-level controls for the share of the workforce with degrees, and the share of the workforce with postgraduate qualifications. Noise controls are four interviewer dummies, indicators of the seniority, gender, and tenure of the managers who responded, the day of the week the interview was conducted, the time of day the interview was conducted, a dummy variable indicating whether the interview was conducted during Ramadan, the duration of the interview, and an indicator of the reliability of the information as coded by the interviewer. Note that no quality information is available for organizations surveyed on a Saturday, and thus the dummy variable indicating a survey took place on a Saturday is omitted in Column 5. Total and capital budget figures are an average of organization budget figures for the years 2006-10. Figures are rounded to two decimal places.

Table 4: Heterogeneous Impacts of Management Practices Related to Incentives and Monitoring

Dependent Variable: Project Completion Rate
**Standard Errors: Clustered by Project Type Within Organization**
Interactions in Deviation from Mean in Columns 1, 3, 4, 5 and 6
OLS Estimates

	(1) Project Complexity	(2) Non-modal Project	(3) Facilities	(4) Tenure	(5) Intrinsic Motivation	(6) Observe Corrupt Practices
CS-Autonomy	0.19***	0.19***	0.20***	0.20***	0.21***	0.16***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
CS-Incentives	-0.16***	-0.13***	-0.14***	-0.11***	-0.17***	-0.17***
	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)
CS-Other	0.08***	0.08***		0.06**	0.07***	0.07***
	(0.02)	(0.02)		(0.03)	(0.02)	(0.03)
CS-Incentives x Project Complexity	-0.19***					
	(0.06)					
CS-Incentives x Non-modal Project Type		-0.05*				
		(0.03)				
CS-Incentives x CS-Facilities			0.03*			
			(0.02)			
CS-Other (Without Facilities)			0.09***			
			(0.03)			
CS-Facilities			0.01			
			(0.02)			
CS-Incentives x Average Tenure of Bureaucrats				-0.02***		
				(0.004)		
CS-Incentives x Proportion of Bureaucrats Intrinsically Motivated					0.54**	
					(0.22)	
CS-Incentives x Proportion of Projects that Bureaucrats Report Observing Corrupt Practices On						0.25
						(0.28)
Project Complexity	-0.01	0.05				
	(0.11)	(0.12)				
Project of Non-modal Type for Organization		-0.04				
		(0.03)				
Average Tenure of Bureaucrats				-0.01		
				(0.01)		
Proportion of Bureaucrats Intrinsically Motivated					-0.47	
					(0.33)	
Proportion of Projects that Bureaucrats Report Observing Corrupt Practices On						-1.09***
						(0.37)
Organization Controls (capital, general, noise)	Yes	Yes	Yes	Yes	Yes	Yes
Project Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Project Type	Project Type	Project Type	Project Type	Project Type	Project Type
Observations (clusters)	4721 (201)	4721 (201)	4721 (201)	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 project type within organization throughout. All columns report OLS estimates. The dependent variable is the proportion of the project completed (that is a continuous measure between zero and one). Project Type fixed effects relate to whether the primary classification of the project is as a financial, training, advocacy, procurement, research, electrification, borehole, dam, building, canal or road project. Project controls comprise project-level controls for the project budget, whether the project is new or a rehabilitation, and an assessment of its aggregate complexity by Nigerian engineers. Capital controls comprise organization-level controls for the logs of number of employees, total budget, and capital budget. General controls comprise organization-level controls for the share of the workforce with degrees, and the share of the workforce with postgraduate qualifications. Noise controls are four interviewer dummies, indicators of the seniority, gender, and tenure of the managers who responded, the day of the week the interview was conducted, the time of day the interview was conducted, a dummy variable indicating whether the interview was conducted during Ramadan, the duration of the interview, and an indicator of the reliability of the information as coded by the interviewer. Total and capital budget figures are an average of organization budget figures for the years 2006-10. In Column 1, the aggregate complexity is a project-level subjective assessment by Nigerian engineers of the relative difficulty of the project within the population of OPEN projects. In Column 2, the non-modal project type is a binary indicator as to whether the project is of a different project type (as defined above) as the modal project type at the organisation. In Column 4, tenure is an organization-level average for the number of years officials have worked at the implementing organization. In Column 5, intrinsic motivation is an organization-level proportion of employees at an organization that answered 'The chance to serve Nigeria' to the question 'What most influenced you to take up a career in the service?' in the Civil Servants Survey. In Column 6, observation of corrupt practices is an organization-level average of the proportion of projects officials at an organization stated on which 'I observed others breaking the service rules for their own benefit' in the Civil Servants Survey. Figures are rounded to two decimal places.

Figure 1: Project Ambiguity and the Response to CS-Incentive Management Practices
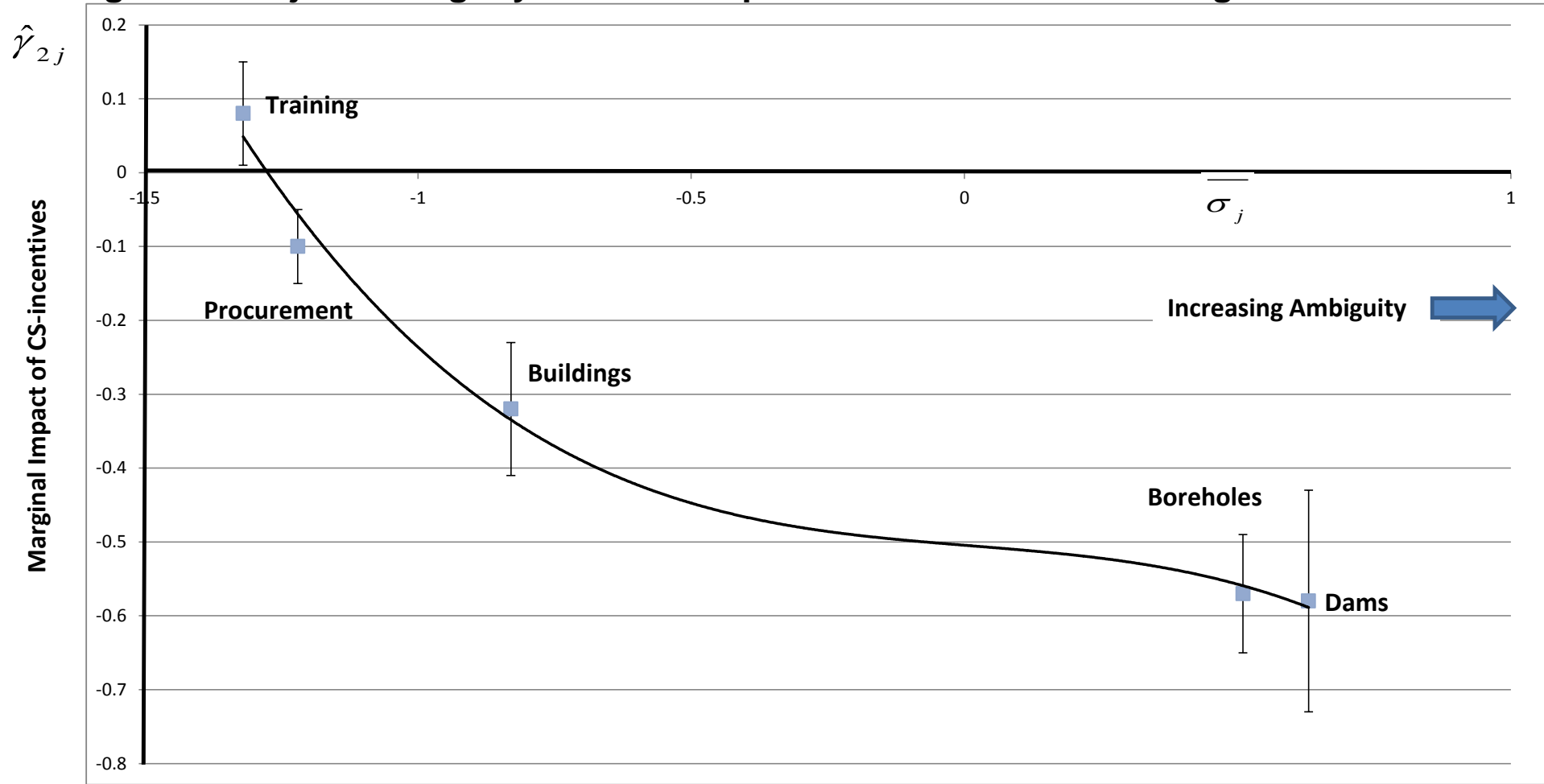
Notes: The horizontal axis measures the ambiguity of projects of a given type. We measure this design ambiguity using subcomponents of the complexity indicator described in Table A3. In particular we construct a z-score based on the design uncertainty, implementation uncertainty, design ambiguity and implementation ambiguity components of the project complexity metric. We then take the average of this over all projects of type j, denoting the average ambiguity of projects of type j by σ_j. We then estimate a specification analogous to our baseline regression model for a given project type j (excluding project fixed effects). We do so for five project types: boreholes, buildings, dams, procurement and training. For each project type j we then obtain an estimate of the partial correlation between CS-incentives and project completion rates, γ_{2j}, conditional on CS-autonomy. Figure 1 then plots the five (γ_{2j}, σ_j) pairs, as well as a cubic best fit.

Table A1: Federal Civil Service Organizations Under Study

Civil Service Organization	Sector	Annual Budget (US$)	Number of Staff	Level of Centralization
Anambra/Imo River Basin Development Authority	Water	26,651,696	324	Deconcentrated
Benin Owena River Basin Development Authority	Water	17,637,829	333	Deconcentrated
Chad Basin River Basin Development Authority	Water	15,781,353	399	Deconcentrated
Cross River River Basin Development Authority	Water	18,823,211	318	Deconcentrated
Citizenship and Leadership Training Centre	Youth	3,510,409	601	Deconcentrated
Federal College of Education, Gombe	Education	5,319,472	608	Deconcentrated
Federal College of Education, Gusau	Education	4,665,009	379	Deconcentrated
Federal College of Education, Omuku	Education	5,887,740	699	Deconcentrated
Federal Government Girls College, Gboko	Education	1,233,030	161	Deconcentrated
Federal Government Girls College, Lejja	Education	1,325,661	122	Deconcentrated
Federal Medical Centre, Abeokuta	Health	6,459,959	1,646	Deconcentrated
Federal Medical Centre, Asaba	Health	4,957,423	777	Deconcentrated
Federal Medical Centre, Bayelsa State	Health	5,026,215	725	Deconcentrated
Federal Medical Centre, Bida	Health	4,135,214	709	Deconcentrated
Federal Medical Centre, Ebute Metta	Health	5,358,665	958	Deconcentrated
Federal Medical Centre, Gombe	Health	5,640,897	1,518	Deconcentrated
Federal Medical Centre, Kebbi State	Health	3,560,097	528	Deconcentrated
Federal Medical Centre, Makurdi	Health	7,120,460	955	Deconcentrated
Federal Medical Centre, Nasarawa State	Health	4,572,968	785	Deconcentrated
Federal Medical Centre, Owerri	Health	8,709,623	1,722	Deconcentrated
Federal Medical Centre, Owo	Health	8,219,773	1,119	Deconcentrated
Federal Medical Centre, Umuahia	Health	7,157,419	1,306	Deconcentrated
Federal Ministry of Agriculture and Rural Development	Agriculture	144,055,160	5,789	Concentrated
Federal Ministry of Education	Education	44,055,244	2,776	Concentrated
Federal Ministry of Environment	Environment	20,324,048	2,093	Concentrated
Federal Ministry of Federal Capital Territory Administration	Education	473,280,702	18,987	Concentrated
Federal Ministry of Health	Health	170,406,214	3,871	Concentrated
Federal Ministry of Housing and Urban Development	Housing	58,610,300	7,837	Concentrated
Federal Ministry of Intergovernmental Affairs, Youth Development and Special Duties	Youth	11,233,012	392	Concentrated
Federal Ministry of Power and Steel	Power	490,123,985	580	Concentrated
Federal Ministry of Water Resources	Water	425,805,770	740	Concentrated
Federal Ministry of Women Affairs	Women	14,934,361	577	Concentrated
Federal Polytechnic, Bida	Education	4,766,557	1,025	Deconcentrated
Federal Polytechnic, Ede	Education	5,111,209	706	Deconcentrated
Federal Polytechnic, Idah	Education	7,643,274	987	Deconcentrated
Federal Polytechnic, Nasarawa	Education	6,431,211	810	Deconcentrated
Federal Polytechnic, Nekede	Education	9,104,663	1,282	Deconcentrated
Federal Polytechnic, Offa	Education	5,397,664	673	Deconcentrated
Federal Polytechnic, Oko	Education	11,266,398	1,627	Deconcentrated
Federal Staff Hospital, Apo-Abuja	Health	2,439,537	471	Deconcentrated
Hadejia-Jama'are River Basin Development Authority	Water	32,758,143	589	Deconcentrated
Jos University Teaching Hospital	Health	19,008,930	2,261	Deconcentrated
Lower Benue River Basin Development Authority	Water	16,307,146	347	Deconcentrated
Lower Niger River Basin Development Authority	Water	18,954,248	436	Deconcentrated
Mass Literacy Commission	Education	6,706,812	154	Deconcentrated
National Action Committee on Aids	Health	12,554,883	243	Deconcentrated
National Arbovirus and Vector Research	Health	301,502	207	Deconcentrated
National Board for Technical Education	Education	10,956,835	460	Deconcentrated
National Centre for Women Development	Women	1,430,538	135	Deconcentrated
National Commission for Colleges of Education	Education	5,633,851	222	Deconcentrated
National Commission for Nomadic Education	Education	4,159,863	195	Deconcentrated
National Primary Health Care Development	Health	56,036,340	646	Deconcentrated
National Tuberculosis and Leprosy Referred Hospital and Training Centre, Zaria	Health	3,000,639	153	Deconcentrated
National Teachers Institute	Education	33,613,921	728	Deconcentrated
National Youth Service Corps	Youth	121,389,773	103,686	Deconcentrated
Niger Delta River Basin Development Authority	Water	18,600,953	406	Deconcentrated
Nigerian Agricultural Cooperative and Rural Development Bank	Agriculture	7,798,094	474	Deconcentrated
Ogun/Oshun River Basin Development Authority	Water	22,259,121	285	Deconcentrated
Sokoto Rima River Basin Development Authority	Water	23,430,400	566	Deconcentrated
Specialist Hospital, Gwagwalada	Health	10,953,476	1,275	Deconcentrated
Universal Basic Education Commission	Education	50,549,049	393	Deconcentrated
Upper Benue River Basin Development Authority	Water	12,822,263	272	Deconcentrated
Upper Niger River Basin Development Authority	Water	21,951,907	239	Deconcentrated

Notes: The budget figures are averages for 2006 to 2010. They are in US Dollars exchanged at a rate of US$1: Naira 150. Staff numbers come from administrative data for 2010. In the few cases we do not have the staff numbers explicitly, we estimate them from the personnel expenditures, which have are correlated with staff numbers with a coefficient of over 0.9. Concentrated organizations refer to the central organizing authority for the sector, with a direct line of responsibility to the President and the National Assembly. Deconcentrated organizations refer to those whose day-to-day running is largely independent of the central authority. They have boards of governors that make decisions over policy and operation and a separate budget line to the central ministries.

Table A2: Defining Management Practices Using the CS Indices

Management Practice	Topic	Specific Questions Related to this Topic
CS-Autonomy	**Roles**	Can most staff above SGL 7 in your organization make substantive contributions to the policy formulation and implementation process?
		Can most staff above SGL 15 in your organization make substantive contributions to the policy formulation and implementation process?
		To what extent do the employees in this organization have the ability to determine how they carry out the assignments in their daily work?
	Flexibility	Does your organization make efforts to redefine its standard procedures in response to the specific needs and peculiarities of a community?
		How flexible would you say your organization is in terms of responding to new practices, new techniques, and regulations?
		At your organization, how efficiently is best practice shared between departments?
		Given past experience, how effectively would a conflict within your organization be dealt with?
CS-Incentives	**Performance Incentives**	Given past experience, how would under-performance be tolerated?
		Given past experience, what happens if there is a part of your organization that isn't achieving agreed results?
		What percentage of *workers* were rewarded when targets were met?
		What percentage of *managers/directors* were rewarded when targets were met?
		Given past experience, are members of this organization disciplined for breaking the Public Service Rules?
		Given past experience, what would most likely happen to a person in this organization who accepted money or a present from someone who came to them with a problem?
	Monitoring	In what kind of ways does your organization track how well it is delivering services?
		If have performance indicators, how often are these indicators collected?
		If have performance indicators, how often are these indicators reviewed by Minister or Permanent Secretary?
		If have performance indicators, how often are these indicators reviewed by non managerial staff?
		Does the organization use performance or quality indicators for tracking the performance of its employees?
		At your organization, how highly regarded is the collection and use of data in planning and implementing projects?
CS-Other	**Facilities**	During a typical working day (8 hours from 8am to 4pm), how many hours is there electricity (PHCN or generator)?
		Out of the five [5] working days, how many days is the network (GSM) coverage working for 50% of calls or more?
		Out of the five [5] working days, how many hours is their internet access good enough to check e-mail?
		Out of every ten [10] officers above SGL 7, how many have access to a computer (desktop or laptop)?
		Out of every ten [10] officers above SGL 7, how many have access to a vehicle (privately owned or otherwise) that can be used for work?
	Skills	Out of every ten [10] officers above SGL 7, how many can use a computer to write a memo?
		Out of every ten [10] officers above SGL 7, how many can use a computer to create a PowerPoint presentation?
		Out of every ten [10] officers above SGL 7, how many can use a computer to create an Excel spreadsheet?
		On which topics have trainings been performed at your organization in the last five [5] years? Technical trainings.
		On which topics have trainings been performed at your organization in the last five [5] years? Laws and regulations.
		On which topics have trainings been performed at your organization in the last five [5] years? Legal rights of the public.
		On which topics have trainings been performed at your organization in the last five [5] years? Good relations with the public.
		On which topics have trainings been performed at your organization in the last five [5] years? Ethics.
		On which topics have trainings been performed at your organization in the last five [5] years? What to do with presents.
		Out of every ten [10] officers above SGL 7 at your organization, how many have had some form of training over the last five [5] years?
	Staffing	Do you think the most senior staff of your organization talk about attracting and developing talented people?
		Do you think the most senior staff of your organization then actually goes about attracting and developing talented people?
		If two people both joined your organization five years ago and one was much better at their work than the other, would he/she be promoted through the service faster?
		Given past experience, if there is a 'top performing' civil servant, does your organization do their best to keep him/her?
		Is the burden of achieving the organization's targets evenly distributed across its different departments, or do some groups consistently shoulder a greater burden than others?
		How do you feel the number of staff in your organization relates to the activities undertaken there?
		What percentage of staff is doing most of the work at your organization?
		Thinking about all the projects that your organization has been involved in since your appointment here, would you say that senior staff try to use the right staff for the right job?
	Targeting	Does your organization have a clear set of targets derived from its mission and goals?
		How tough are the targets of the organization?
		When you arrive at work each day, do you and your colleagues know what your organization is trying to achieve on that particular day?
	Culture	How effectively would you say your organization is in making the bulk of its staff feel valued?
		To what extent would you say employees of your organization trust each other?
		If you think about the way that employees of this organization respond to a standard work challenge, would you say that there is a set of 'shared values' amongst all the staff?
		Out of every ten [10] officers above SGL 7, how many people from this organization participate in groups, committees and activities with other people from this organization outside of the formal structure of government (for example, in community or social organizations)?

Table A3: Correlation of Subcomponents of the Project Complexity Indicator

	Mean	Standard deviation	Project size	Number of inputs	Number of methods	Interdependencies	Access to raw materials	Storage of raw materials	Requires local labor	Requires skilled labor	Access to construction equipment	Design uncertainty	Implementation uncertainty	Design ambiguity	Implementation ambiguity	Difficulty to manage	Number of agencies involved	Aggregate complexity
Project size	0.27	0.45	1.00															
Number of inputs	6.82	4.13	0.06	1.00														
Number of methods	5.04	2.29	0.33	0.61	1.00													
Interdependencies	0.65	0.48	-0.03	0.13	0.07	1.00												
Access to raw materials	0.25	0.43	-0.11	-0.24	-0.09	0.04	1.00											
Storage of raw materials	0.04	0.21	0.19	-0.05	0.07	0.07	-0.05	1.00										
Requires local labor	0.45	0.50	0.31	-0.03	0.48	0.04	0.23	0.14	1.00									
Requires skilled labor	0.45	0.50	-0.16	-0.02	-0.21	0.53	0.21	-0.03	0.00	1.00								
Access to construction equipment	0.21	0.41	-0.01	-0.24	0.06	0.36	0.62	0.12	0.53	0.53	1.00							
Design uncertainty	0.70	0.46	0.10	0.18	0.12	0.77	-0.08	0.02	0.26	0.45	0.31	1.00						
Implementation uncertainty	0.78	0.41	0.06	0.26	0.20	0.58	-0.16	-0.01	0.35	0.42	0.24	0.78	1.00					
Design ambiguity	0.66	0.47	-0.08	0.08	-0.01	0.84	-0.01	0.08	0.04	0.60	0.35	0.73	0.63	1.00				
Implementation ambiguity	0.65	0.48	0.00	0.06	0.02	0.85	0.02	0.11	0.07	0.56	0.36	0.75	0.58	0.89	1.00			
Difficulty to manage	0.28	0.45	0.15	-0.16	0.24	0.38	0.47	0.19	0.64	0.37	0.81	0.36	0.27	0.43	0.43	1.00		
Number of agencies involved	3.54	0.51	-0.01	0.12	-0.05	0.21	0.11	-0.02	0.25	0.55	0.21	0.39	0.54	0.30	0.26	0.22	1.00	
Aggregate complexity	24.98	17.92	0.36	0.24	0.50	0.25	-0.21	0.16	0.39	-0.09	0.03	0.35	0.37	0.29	0.32	0.33	-0.05	1.00
Observations (projects)	4721	4721	4721	4721	4721	4721	4721	4721	4721	4721	4721	4721	4721	4721	4721	4721	4721	4721

Notes: The sample used is those projects in our core analysis for which we have complexity and project completion data. 'Project size' is a binary variable that aims to gauge the physical size of the project. It takes the value 1 if it is classified as equivalent to a medium scale build or larger. 'Number of inputs' counts the number of distinct product classes the finished project contains. 'Number of methods' counts the number of distinct disciplines or methods involved in implementing the project. 'Interdependencies' is a binary variable reflecting the extent of interdependencies between the activities involved in the project. It takes a value of 1 if the project is classified as highly interdependent. 'Access to raw materials' is a binary variable that takes the value 1 if raw materials could not be sourced within the state of implementation. 'Storage of raw materials' is a binary variable that takes the value 1 if some of the raw materials could not be easily stored or transported. 'Requires local labor' is a binary variable that takes the value 1 if local labor was useful or critical. 'Requires skilled labor' is a binary variable that takes the value 1 if specialized skills were necessary and difficult to obtain. 'Access to construction equipment' is a binary variable that takes the value 1 if the equipment required is difficult to obtain, heavy duty, or difficult to transport to the site. 'Design uncertainty' is a binary variable that takes on the value 1 if the design of the project is context specific. 'Implementation uncertainty' is a binary variable that takes on the value 1 if there are substantial risks involved in implementation. 'Design ambiguity' is a binary variable that takes on the value 1 if there is a risk of redesign late on in the project. 'Implementation ambiguity' is a binary variable that takes on the value 1 if the technical risks of the project cannot be fully understood at implementation. 'Difficulty to manage' is a binary variable that takes the value 1 if the project is seen have elements that require project management skills of above average level. 'Number of agencies involved' is simply a count of the estimated number of agencies involved in the project cycle. 'Aggregate complexity' is a subjective assessment as to the overall complexity of the project by the coding engineers. This variable is an assessment of the interaction of the other variables as well as any unassessed aspects of complexity and provides a coherent picture of the complexity of the projects by a specialist. Figures are rounded to two decimal places.

Table A4: Standard Errors

Dependent Variable: Project Completion Rate
OLS Estimates

	(1) Unconditional	(2) Organization Controls	(3) Project Controls	(4) Project Type Fixed Effects
CS-Autonomy	0.11	0.18	0.17	0.18
Robust standard errors	(0.01)	(0.01)	(0.01)	(0.01)
Errors clustered by project type within organisation level	(0.05)	(0.03)	(0.03)	(0.03)
Errors clustered by organisation	(0.06)	(0.03)	(0.03)	(0.03)
P-value of clustered coefficients	[0.07]	[0.00]	[0.00]	[0.00]
Wild cluster bootstrap p-value of clustered coefficients	[0.08]	[0.02]	[0.02]	[0.00]
CS-Incentives	-0.06	-0.11	-0.11	-0.14
Robust standard errors	(0.01)	(0.01)	(0.01)	(0.01)
Errors clustered by project type within organisation level	(0.03)	(0.02)	(0.02)	(0.02)
Errors clustered by organisation	(0.03)	(0.03)	(0.03)	(0.02)
P-value of clustered coefficients	[0.08]	[0.00]	[0.00]	[0.00]
Wild cluster bootstrap p-value of clustered coefficients	[0.20]	[0.00]	[0.00]	[0.00]
CS-Other	0.10	0.05	0.05	0.08
Robust standard errors	(0.01)	(0.02)	(0.02)	(0.02)
Errors clustered by project type within organisation level	(0.04)	(0.03)	(0.03)	(0.02)
Errors clustered by organisation	(0.03)	(0.04)	(0.03)	(0.03)
P-value of clustered coefficients	[0.00]	[0.20]	[0.17]	[0.00]
Wild cluster bootstrap p-value of clustered coefficients	[0.01]	[0.28]	[0.26]	[0.01]
Organization Controls (capital, general, noise)	No	Yes	Yes	Yes
Project Controls	No	No	Yes	Yes
Fixed Effects	None	None	None	Project Type
Observations	4721	4721	4721	4721

Notes: Standard errors are in round parentheses, and are robust, clustered by project type within organization, and by organization respectively under the coefficients corresponding to a management index. Standard p-values and wild cluster bootstrap p-values are in square parentheses below these. All columns report OLS estimates. The dependent variable is the proportion of the project completed (that is a continuous measure between zero and one). Project Type fixed effects relate to whether the primary classification of the project is as a financial, training, advocacy, procurement, research, electrification, borehole, dam, building, canal or road project. Project controls comprise project-level controls for the project budget, whether the project is new or a rehabilitation, and an assessment of its aggregate complexity by Nigerian engineers. Capital controls comprise organization-level controls for the logs of number of employees, total budget, and capital budget. General controls comprise organization-level controls for the share of the workforce with degrees, and the share of the workforce with postgraduate qualifications. Noise controls are four interviewer dummies, indicators of the seniority, gender, and tenure of the managers who responded, the day of the week the interview was conducted, the time of day the interview was conducted, a dummy variable indicating whether the interview was conducted during Ramadan, the duration of the interview, and an indicator of the reliability of the information as coded by the interviewer. Total and capital budget figures are an average of organization budget figures for the years 2006-10. Figures are rounded to two decimal places.

Table A5: Robustness Checks

Dependent Variable: Project Completion Rate
Standard Errors: Clustered by Project Type Within Organization
OLS Estimates

	(1) Largest Org. by Total Exp.	(2) Org. With Most Projects	(3) Ten Orgs. with Smallest Total Exp.	(4) Ten Orgs. with Smallest No. of Projects	(5) Orgs. Below 5% or Above 95% of CS-Autonomy Scale	(6) Orgs. Below 5% or Above 95% of CS-Performance Scale	(7) Construction Projects	(8) Non-Construction Projects	(9) Centralized	(10) Decentralized	(11) State-level Controls	(12) Northern Projects	(13) Southern Projects	(14) Weighted Topics in CS Indices	(15) Fractional Regression
CS-Autonomy	0.17***	0.17***	0.19***	0.19***	0.23***	0.20***	0.23***	0.16***	0.17***	0.14***	0.17***	0.19***	0.22***	0.23***	1.02***
	(0.02)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.05)	(0.05)	(0.03)	(0.03)	(0.03)	(0.06)	(0.05)	(0.03)	(0.17)
CS-Incentives	-0.15***	-0.13***	-0.15***	-0.14***	-0.24***	-0.18***	-0.21***	-0.06*	-0.24***	-0.01	-0.14***	-0.12**	-0.30***	-0.14***	-0.96***
	(0.02)	(0.03)	(0.02)	(0.02)	(0.04)	(0.03)	(0.05)	(0.03)	(0.04)	(0.03)	(0.02)	(0.06)	(0.10)	(0.02)	(0.16)
CS-Other	0.09***	0.05	0.09***	0.08***	0.09***	0.06**	0.11**	0.06*	0.18***	-0.03	0.08***	-0.03	0.10*	0.08***	0.49***
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)	(0.03)	(0.05)	(0.04)	(0.03)	(0.08)	(0.05)	(0.02)	(0.15)
Organization Controls (capital, general, noise)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Project Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Project Type	Project Type	Project Type	Project Type	Project Type	Project Type	Project Type	Project Type	Project Type	Project Type	Project Type	Project Type	Project Type	Project Type	Project Type
Observations (clusters)	4201 (197)	3924 (192)	4601 (176)	4711 (191)	3810 (145)	4209 (181)	3822 (82)	899 (119)	2434 (53)	2287 (148)	4269 (63)	2049 (91)	1895 (104)	4721 (201)	4721 (201)

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. Standard errors are in parentheses, and are clustered by project type within organization throughout. All columns bar 15 report OLS estimates. Column 15 reports estimates from a fractional regression model. The dependent variable is the proportion of the project completed (that is a continuous measure between zero and one). Column 1 excludes those projects implemented by the largest organization in terms of total expenditures. Column 2 excludes projects implemented by the largest organization in terms of number of projects. Columns 3 and 4 remove the 10 smallest organizations by expenditures and number of projects respectively. Columns 5 and 6 exclude organizations at the top and bottom of the CS-autonomy and CS-performance management scales respectively. Column 7 includes only construction projects (borehole, building, electrification, dam, road and canal) and Column 8 includes only non-construction projects. Column 9 restricts our specification to those projects implemented by centralized ministries. Column 10 restricts our specification to those projects implemented by decentralized agencies. In Column 11, "State-level Controls" comprise 'poverty controls', the proportion of households in a state who have difficulty meeting their food needs and the proportion of households that self-classify as poor; 'educational controls', literacy rates for adults and youth, and primary and secondary net enrolment and completion rates; and 'infrastructure controls', indicators of access to secure tenure, safe water, safe sanitation, improved waste disposal, medical services, vaccinations, electricity, computers, mobile phones and credit facilities, as well as the unemployment rate for over 15s. Column 3 restricts our specification to those projects for which we have location data and that are implemented in Northern Nigeria, where we exclude the Saturday dummy. Column 4 restricts our specification to those projects for which we have location data and that are implemented in Southern Nigeria. In Column 14 we construct CS-Autonomy and CS-Performance measures by weighting each topic (rather than each variable) equally. Project Type fixed effects relate to whether the primary classification of the project is as a financial, training, advocacy, procurement, research, electrification, borehole, dam, building, canal or road project. Project controls comprise project-level controls for the project budget, whether the project is new or a rehabilitation, and an assessment of its aggregate complexity by Nigerian engineers. Capital controls comprise organization-level controls for the logs of number of employees, total budget, and capital budget. General controls comprise organization-level controls for the share of the workforce with degrees, and the share of the workforce with postgraduate qualifications. Noise controls are four interviewer dummies, indicators of the seniority, gender, and tenure of the managers who responded, the day of the week the interview was conducted, the time of day the interview was conducted, a dummy variable indicating whether the interview was conducted during Ramadan, the duration of the interview, and an indicator of the reliability of the information as coded by the interviewer. Total and capital budget figures are an average of organization budget figures for the years 2006-10. Figures are rounded to two decimal places.

Table A6: Management Practices and the Assignment of Projects to Organizations

Dependent Variable: Binary Variable Indicating Organization Implementing Project
Standard Errors: Clustered by Project Type Within Organization
Conditional Logit Model

	(1) Unconstrained Pairs	(2) Feasible Pairs by Sector	(3) Feasible Pairs by Complexity	(4) Organization Controls	(5) Project Interactions
CS-Autonomy	0.35	0.53*	0.37	0.35	0.40
	(0.22)	(0.27)	(0.24)	(0.29)	(0.31)
CS-Incentives	-0.17	-0.33	-0.26	-0.26	-0.34
	(0.26)	(0.33)	(0.36)	(0.45)	(0.45)
CS-Other	-0.36	-0.02	0.37	0.57	0.63
	(0.30)	(0.23)	(0.31)	(0.45)	(0.48)
Senior Bureaucrat's Span of Control				-0.22	-0.34
				(0.35)	(0.40)
Tenure of Senior Bureaucrats				0.03	0.05
				(0.17)	(0.18)
Tenure of Low-tier Bureaucrats				-0.03	-0.04
				(0.09)	(0.09)
Proportion of Senior Bureaucrats That Report Being Intrinsically Motivated				-4.82	-4.6
				(4.17)	(4.37)
Proportion of Low-tier Bureaucrats That Report Being Intrinsically Motivated				-0.35	0.06
				(3.49)	(3.22)
Proportion of Projects that Senior Bureaucrats that Report Observing Corrupt Practices On				-0.01	0.00
				(0.04)	(0.04)
Proportion of Projects that Low-tier Bureaucrats that Report Observing Corrupt Practices On				0.02	0.02
				(0.03)	(0.04)
Organization Controls (capital, general, noise)	No	No	No	Yes	Yes
Interactions of Project and Organizational Characteristics	No	No	No	No	Yes
Project-Organization Paired Observations (clusters)	297423 (63)	44429 (59)	33480 (59)	33370 (58)	33370 (58)

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. Standard errors are in parentheses, and are clustered by project type within organization throughout. All Columns report conditional logit estimates. The data is set up as follows. For each project, we associate a binary variable with 63 values corresponding to the 63 organizations in the core analysis of the paper. This variable takes the value one if the project is implemented at that organization, and zero otherwise. Thus, for each of our 4721 projects, we have 63 'project-organization' observations. These 297,423 observations are what makes up the 'unrestricted pairs' specification in Column 1. The 'feasible pairs' specification in Column 2 restricts the set of organizations associated with a project to only those in the same sector. Those 847 projects implemented at the only organization in the sector drop out in this specification, leaving us with 44,429 project-organization observations. In Columns 4 and 5 as we do not have data for one organization on the proportion of senior bureaucrats that report observing corrupt practices, this organization drops out, leaving 33,370 project-organization pairs. In Column 5 we restrict the feasible pairs further by assuming projects can only potentially be assigned to organizations in the same sector that are observed being tasked to implement at least one project of similar or greater complexity. Capital controls comprise organization-level controls for the logs of number of employees, total budget, and capital budget. General controls comprise organization-level controls for the share of the workforce with degrees, and the share of the workforce with postgraduate qualifications. Noise controls are four interviewer dummies, indicators of the seniority, gender, and tenure of the managers who responded, the day of the week the interview was conducted, the time of day the interview was conducted, a dummy variable indicating whether the interview was conducted during Ramadan, the duration of the interview, and an indicator of the reliability of the information as coded by the interviewer. Total and capital budget figures are an average of organization budget figures for the years 2006-10. We follow the grading system of the Federal Government by defining senior bureaucrats as those on grade level 12 and above. The proportion of staff intrinsically motivated refers to the fraction of employees at an organization that answered 'The chance to serve Nigeria' to the question 'What most influenced you to take up a career in the service?' in the Civil Servants Survey. The percentage of staff who observed corruption refers to the average proportion of projects officials at an organization stated on which 'I observed others breaking the service rules for their own benefit' in the Civil Servants Survey. Figures are rounded to two decimal places.

Table A7: Management Practices and Public Sector Service Delivery at the Organization Level

Robust Standard Errors

OLS Estimates

	(1) Number of Projects Assigned to Organization	(2) Number of Unique Project Types Assigned to Organization	(3) Project Complexity	(4) Log Organizational Budget
CS-Autonomy	22.17	0.21	1.56	0.03
	(19.57)	(0.27)	(1.89)	(0.05)
CS-Incentives	10.07	-0.01	-0.47	0.05
	(24.87)	(0.40)	(1.53)	(0.07)
CS-Other	-19.92	-0.01	-2.10	0.03
	(28.30)	(0.37)	(1.66)	(0.05)
Organization Controls (capital, general, noise)	Yes	Yes	Yes	Yes
Observations	63	63	63	63

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. Robust standard errors are in parentheses. All columns report OLS estimates. The dependent variable in Column 1 is a count measure of the number of projects assigned to an organization. In Column 2 it a count variable of the number of unique project types assigned to the organization, where the projects are assigned a primary classification of the following type: a financial, training, advocacy, procurement, research, electrification, borehole, dam, building, canal or road project. The dependent variable in Column 3 is the average complexity of projects assigned to the organization. The dependent variable in Column 4 is the log of the average total organizational budget, averaged over 2006 to 2010. In Columns 1 to 3, capital controls comprise organization-level controls for the logs of number of employees, total budget, and capital budget. In Column 4, we drop total organizational budget as a control as the log of organizational budget is the dependent variable. General controls comprise organization-level controls for the share of the workforce with degrees, and the share of the workforce with postgraduate qualifications. Noise controls are four interviewer dummies, indicators of the seniority, gender, and tenure of the managers who responded, the day of the week the interview was conducted, the time of day the interview was conducted, a dummy variable indicating whether the interview was conducted during Ramadan, the duration of the interview, and an indicator of the reliability of the information as coded by the interviewer. Total and capital budget figures are an average of organization budget figures for the years 2006-10. Figures are rounded to two decimal places.

Table A8: Decomposing the Impacts of Management Practices

Dependent Variable: Proportion Project Completed
Standard Errors: Clustered by Project Type Within Organization
OLS Estimates

	(1) Roles	(2) Flexibility	(3) Performance incentives	(4) Monitoring	(5) Culture	(6) Targeting	(7) Facilities	(8) Skills	(9) Staffing	(10) All	(11) Management
CS-Roles	0.27***									0.33***	
	(0.06)									(0.04)	
CS-Flexibility		0.18***								0.11**	
		(0.05)								(0.04)	
CS-Performance Incentives			-0.08*							-0.07	
			(0.04)							(0.06)	
CS-Monitoring				-0.21***						-0.28***	
				(0.05)						(0.06)	
CS-Culture					0.02					0.01	
					(0.04)					(0.03)	
CS-Targeting						0.03				0.07	
						(0.04)				(0.05)	
CS-Facilities							0.13***			0.06	
							(0.04)			(0.04)	
CS-Skills								0.08		0.14**	
								(0.07)		(0.06)	
CS-Staffing									-0.06	0.08	
									(0.07)	(0.06)	
CS-Management											0.03
											(0.03)
H₀: CS-Autonomy Components Equal [p-value]										[0.00]	
H₀: CS-Incentives Components Equal [p-value]										[0.06]	
Organization Controls (capital, general, noise)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Project Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Project Type	Project Type	Project Type	Project Type	Project Type	Project Type	Project Type	Project Type	Project Type	Project Type	Project Type
Observations (clusters)	4721 (201)	4721 (201)	4721 (201)	4721 (201)	4721 (201)	4721 (201)	4721 (201)	4721 (201)	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 project type within organization throughout. All columns report OLS estimates. The dependent variable in all columns is the proportion of the project completed (that is a continuous measure between zero and one). All of the index component variables in the table are z-scores which are centered at 0. The variable 'facilities' measures the quality of infrastructure at an organization and the equipment available to staff. The variable 'skills' measures the skills and training opportunities embodied in the staff body. The variable 'staffing' measures the effective utilization of that body of staff. The variable 'targeting' examines the extent of use of targets. The variable 'flexibility' measures the extent to which the organization is able to respond to best practice and project peculiarities. The variable 'roles' measures the extent to which staff can play a role in defining the direction of the organization. The variable 'culture' measures whether the organization inculcates a productive work culture. Together these components make up our autonomy variable. The variable 'monitoring' focuses on the tracking of performance of individuals and projects. The variable 'performance' examines the use of incentives both to reward success and punish failure. Project Type fixed effects relate to whether the primary classification of the project is as a financial, training, advocacy, procurement, research, electrification, borehole, dam, building, canal or road project. Project controls comprise project-level controls for the project budget, whether the project is new or a rehabilitation, and an assessment of its aggregate complexity by Nigerian engineers. Capital controls comprise organization-level controls for the logs of number of employees, total budget, and capital budget. General controls comprise organization-level controls for the share of the workforce with degrees, and the share of the workforce with postgraduate qualifications. Noise controls are four interviewer dummies, indicators of the seniority, gender, and tenure of the managers who responded, the day of the week the interview was conducted, the time of day the interview was conducted, a dummy variable indicating whether the interview was conducted during Ramadan, the duration of the interview, and an indicator of the reliability of the information as coded by the interviewer. Total and capital budget figures are an average of organization budget figures for the years 2006-10. At the foot of the table we report the p-value on the null that the coefficients in Column 10 on the variables that make up the CS-Autonomy measure (those associated with roles and flexibility) are of equal magnitude. We also report the p-value on the null that the coefficients in Column 10 on the variables that make up the CS-Incentives measure (those associated with performance incentives and monitoring) are of equal magnitude. Figures are rounded to two decimal places.

Table A9: Bureaucrat Characteristics and Management Practices
Means, standard deviations and regression coefficients

	(1) Mean (s.d.)	Joint Regression Coefficients	
		(2) Coefficient on CS-Autonomy	(3) Coefficient on CS-Incentives
Proportion of Bureaucrats Who Knew Direct Boss Before Starting Posting	0.00	-	-
	(0.01)		
Proportion of Bureaucrats Who Knew Managers Before Starting Posting	0.00	-	-
	(0.01)		
Average Number of Close Colleagues Known Before Starting Posting	4.40	-0.06	-0.04
	(3.18)	(0.42)	(0.42)
Proportion of Bureaucrats Being the First Family Member to Join Service	0.48	0.02	-0.01
	(0.1)	(0.01)	(0.01)
Tenure of Senior Bureaucrats	15	0.42	-0.49
	(5)	(0.65)	(0.65)
Tenure of Low-tier Bureaucrats	10	0.00	-0.48
	(4)	(0.59)	(0.59)
Proportion of Senior Bureaucrats Intrinsically Motivated	0.31	0.00	-0.02
	(0.12)	(0.02)	(0.02)
Proportion of Low-tier Bureaucrats Intrinsically Motivated	0.39	0.01	-0.02
	(0.13)	(0.02)	(0.02)
Proportion of Projects that Senior Bureaucrats that Report Observing Corrupt Practices On	0.35	-0.02	0.01
	(0.10)	(0.01)	(0.01)
Proportion of Projects that Low-tier Bureaucrats that Report Observing Corrupt Practices On	0.41	-0.02	-0.01
	(0.11)	(0.01)	(0.01)

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. Columns 2 to 3 report OLS estimates. In Column 1, standard deviations are in parentheses. In all other columns, standard errors are in parentheses. Proportion of bureaucrats who knew their direct boss before starting their posting is an organization-level average of those bureaucrats who answered positively to a question outlining possible connections to the officials boss before they came to work at the organization. Proportion of bureaucrats who knew their managers before starting their posting is an organization-level average of those bureaucrats who answered positively to a question outlining possible connections to the officials' managers before they came to work at the organization. Average number of close colleagues known before starting their posting is an organization-level average of the number of colleagues an official stated that they had connections with before they came to work at the organization. Proportion of bureaucrats being the first family member to join service is an organization-level average of those bureaucrats who answered 'Yes' to the question 'Are you the first member of your family to enter the Federal service'. Tenure of bureaucrats refers to the number of years they have served in the organization for. The proportion of staff intrinsically motivated refers to the fraction of employees at an organization that answered 'The chance to serve Nigeria' to the question 'What most influenced you to take up a career in the service?'. The percentage of staff who observed corruption refers to the average proportion of projects officials at an organization stated on which 'I observed others breaking the service rules for their own benefit'. These variables are all elicited from the Civil Servants Survey we conducted. When we control for the proportion of projects that senior bureaucrats that report observing corrupt practices on, we lose one organization (that implemented a single project) in which no senior bureaucrat answered the question. Figures are rounded to two decimal places.

Table A10: Endogenous Management Practices

Dependent Variable: Project Completion Rate
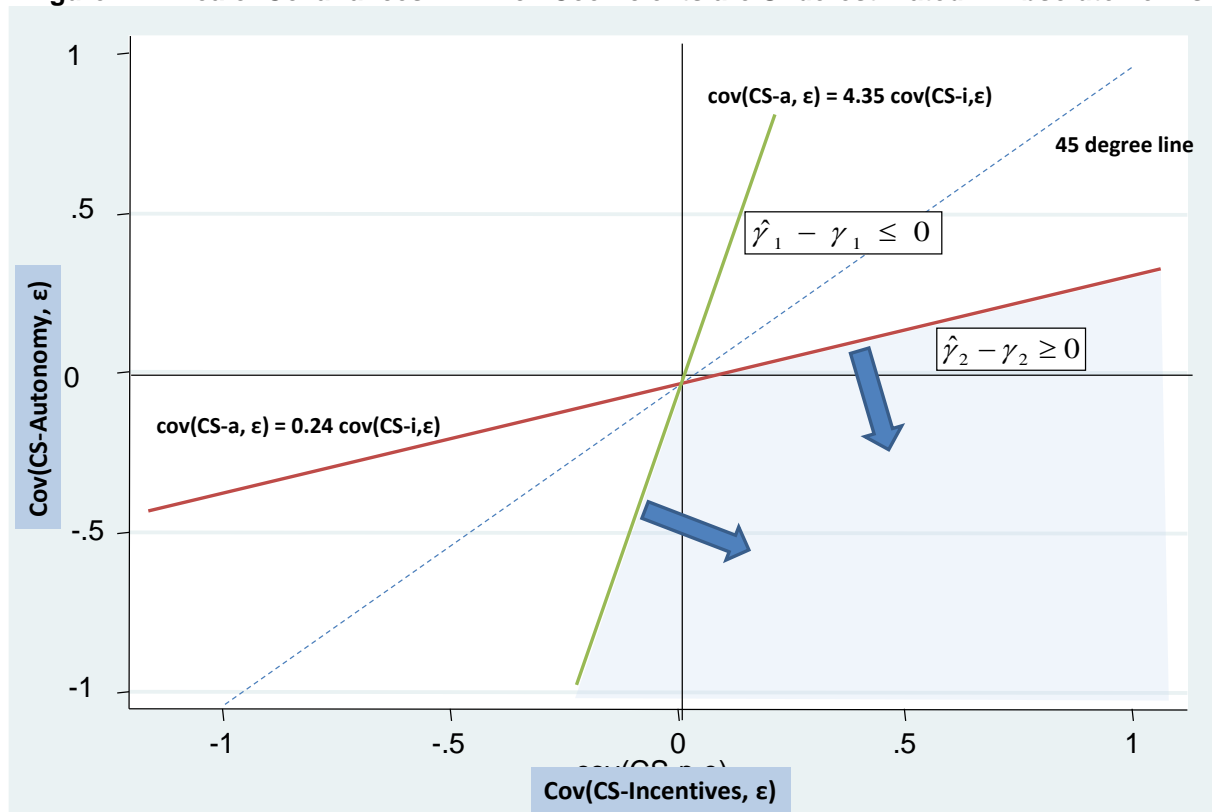**Standard Errors: Clustered by Project Type Within Organization**
**Interactions in Deviation from Mean**
OLS Estimates

	(1) Capital to Total Budget Ratio	(2) Variance of Ambiguity in Portfolio of Projects
CS-Autonomy	0.17***	0.20***
	(0.04)	(0.03)
CS-Incentives	-0.17***	-0.17***
	(0.04)	(0.04)
CS-Other	0.07***	0.04
	(0.02)	(0.04)
CS-Autonomy x Capital to Total Budget Ratio	0.07	
	(0.10)	
CS-Incentives x Capital to Total Budget Ratio	0.10	
	(0.12)	
CS-Autonomy x Var (Project Ambiguity)		0.00
		(0.02)
CS-Incentives x Var (Project Ambiguity)		-0.04
		(0.03)
Capital to Total Budget Ratio	0.70	
	(0.54)	
Var (Project Ambiguity)		-0.01
		(0.03)
Organization Controls (capital, general, noise)	Yes	Yes
Project Controls	Yes	Yes
Fixed Effects	Project Type	Project Type
Observations	4721 (201)	4721 (201)

Notes: *** denotes significance at 1%, ** at 5%, and * at 10% level. Standard errors are in parentheses, and are clustered by project type within organization throughout. All columns report OLS estimates. The dependent variable is the proportion of the project completed (that is a continuous measure between zero and one). Project Type fixed effects relate to whether the primary classification of the project is as a financial, training, advocacy, procurement, research, electrification, borehole, dam, building, canal or road project. Project controls comprise project-level controls for the project budget, whether the project is new or a rehabilitation, and an assessment of its aggregate complexity by Nigerian engineers. Capital controls comprise organization-level controls for the logs of number of employees, total budget, and capital budget. General controls comprise organization-level controls for the share of the workforce with degrees, the share of the workforce with postgraduate qualifications, and the span of control at the organization. Noise controls are four interviewer dummies, indicators of the seniority, gender, and tenure of the managers who responded, the day of the week the interview was conducted, the time of day the interview was conducted, a dummy variable indicating whether the interview was conducted during Ramadan, the duration of the interview, and an indicator of the reliability of the information as coded by the interviewer. Total and capital budget figures are an average of organization budget figures for the years 2006-10. The capital to total budget Ratio is a variable that divides an organization's capital budget by its total budget. Figures are rounded to two decimal places.

Figure A1: Area of Covariances in Which Coefficients are Underestimated in Absolute Terms

cov(CS-a, ε) = 4.35 cov(CS-i,ε)

45 degree line

$\hat{\gamma}_1 - \gamma_1 \leq 0$

$\hat{\gamma}_2 - \gamma_2 \geq 0$

cov(CS-a, ε) = 0.24 cov(CS-i,ε)

Cov(CS-Autonomy, ε)

Cov(CS-Incentives, ε)

Figure A2: Impact of Management Practices For Different Thresholds of Project Success

Coefficient on CS-Autonomy

Coefficient on CS-Incentives

Threshold of project success (% completed)
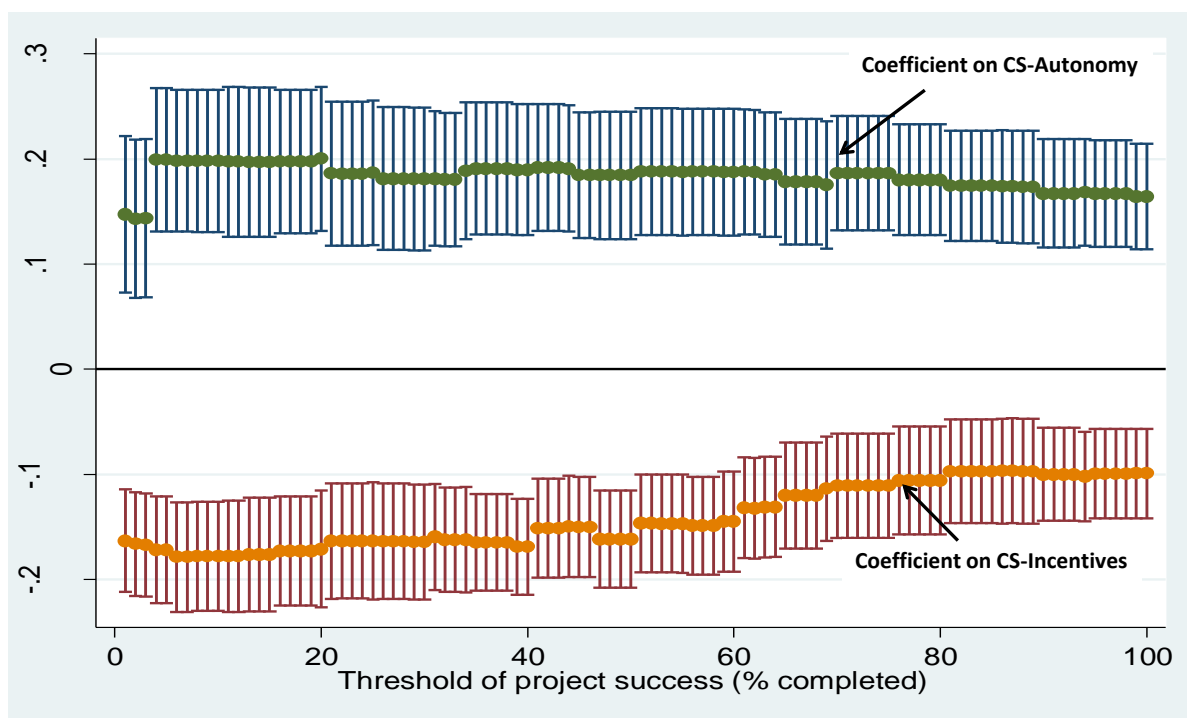
Notes: We define a threshold completion rate of p% and then consider all potential thresholds from 1% to 100% in increments of 1%. We then use this to define a dummy variable for our main empirical specification, where the outcome is a dummy variable equal to one if the project completion rate is above the threshold p%, and zero otherwise. For any given threshold p the coefficients of interest on CS-Autonomy and CS-Incentives are shown above and their associated 95% confidence interval.