The Allocation of Authority in Organizations: A Field Experiment with Bureaucrats*

Oriana Bandiera, Michael Carlos Best, Adnan Q. Khan, & Andrea Prat

April 2019
To view the latest version, please go to
https://bit.ly/2UItDXP

Abstract

We design a field experiment to study the allocation of authority in organizations where autonomy leverages the agents' private information and rules prevent them from extracting private benefits. A simple model illustrates that monitoring of adherence to rules creates a second set of agents subject to their own agency problems, and hence the optimal allocation of authority and incentives depends on the relative alignment of frontline workers and their monitors with organizational goals. The experiment, run with the government of Punjab, Pakistan, creates exogenous variation in the autonomy and incentives of 600 procurement officers. We find that increasing procurement officers' autonomy vis-à-vis their auditors reduces prices by 9% without reducing quality and the effect is stronger when the auditor is more extractive. In contrast, performance pay only reduce prices when the auditor is not extractive and is close to zero on average. The results suggest auditors are less concerned with saving public money than procurement officers are. This has implications for organizational design and anti-corruption policies.

*Bandiera: London School of Economics o.bandiera@lse.ac.uk, Best: Columbia University and NBER michael.best@columbia.edu, Khan: London School of Economics a.q.khan@lse.ac.uk, Prat: Columbia University andrea.prat@columbia.edu. This experiment was preregistered in the Social Science Registry at https://www.socialscienceregistry.org/trials/610. We are grateful to the World Bank, the International Growth Centre, and the JPAL Governance Initiative for financial support. We thank XXXX and seminar participants at the Canadian Institute for Advanced Research, Chicago Booth, Columbia, Georgetown, Harvard/MIT, the London School of Economics, Manchester, Rosario-Los Andes, Simon Fraser, Tepper School of Business, University of Southern California, and Warwick for helpful comments and discussion. All remaining errors are our own.

1 Introduction

The allocation of authority between bosses and subordinates is a crucial choice for all organizations. Classic theory models the problem as a trade-off between the principal retaining control to ensure the agent acts in the interest of the organization or granting autonomy to allow the agent to use their private information at the risk of them exploiting it for their private benefit. In most organizations, however, control over rules that regulate agents' behavior resides with other agents at a higher level of the hierarchy rather than the principal itself, and these agents might also be prone to act in their own interest.

Guided by a simple model, we design a field experiment to provide evidence on the trade-off organizations face when allocating authority between agents whose interests are misaligned with the organization's. The model makes precise how the response to an exogenous shift in autonomy allows us to infer the relative strength of misalignment of the two sets of agents, and how the effectiveness of performance pay for agents at the lower level depends on the degree of misalignment of the agent at the top.

We engineer an exogenous increase in autonomy in a statewide experiment with 600 offices of the government of Punjab, Pakistan. Within these organizations we focus on the procurement unit where procurement officers buy the generic goods required by the office and are supervised by an independent auditor office. Procurement of generic goods is ideal for two reasons: performance is measurable as goods are homogeneous and agency issues are severe as officers buy goods they do not use with money they do not own. As in most bureaucracies, the behavior of frontline workers is heavily regulated to prevent corruption so in the status quo procurement officers have little autonomy and the auditors hold them accountable to many rules.

We model the interaction between officers and auditors, whose type determines how aligned they are with the organization. Officers and auditors choose a mark up to maximize their utility which depends on their type, the allocation of authority, and the financial incentives they face. The equilibrium price is a function of both types, whose weight depends on the allocation of authority. When rules are many, the auditors' type matter more, whereas when the agent has more autonomy her type matters more. The comparative static with respect to the policy parameters that we vary with the experiment are as follows. First an increase in autonomy (equivalent to a reduction in rules) lowers prices if and only if the auditor is sufficiently misaligned relative to the procurement officer, and the reduction is larger the more misaligned the auditor. Conversely, performance pay for the agent decreases prices only if the auditor is sufficiently aligned. If she is not, the agent cannot do much to reduce prices as these are mostly kept high by the supervisor's mark up. The model thus yield two independent predictions on the effect of autonomy and performance pay. that we can use to back out the relative misalignment of auditors and officers in our context.

To create variation in the policy parameters we randomly allocate 600 procurement offices to four groups: a control group, an autonomy group, a pay for performance group and a group that gets both. The experiment lasts two years in total and we stagger the introduction of the two treatments so that performance pay. is offered from the first year whilst autonomy only kicks in the second year. This allows us to use the control group in the first year as a benchmark for the status quo and to build a proxy for the auditor's type because each district and department has their own auditors.

Our findings are as follows. First, autonomy reduces prices by 9% on average either on its own or in combination with performance pay. whilst the effect of performance pay is close to zero. In light of the model this indicates that the auditors are relatively more misaligned on average. Next we use the share of transactions approved at the end of the fiscal year to proxy for the auditors' type and test whether indeed the cross-derivative of autonomy and performance pay have opposite signs. We find that performance pay reduces prices by 11% when the auditor is aligned, that is when he approves transactions smoothly over the year. Under these circumstances autonomy has a smaller effect as it does not limit the power of a bad type.Conversely, performance pay has no effect when the auditor is misaligned, which is true in average. Taken together the results indicate that the two policy instruments are effective under different circumstances: giving autonomy to the agent is desirable when it means taking it away from an extractive auditor while incentives are ineffective in this case because the agent has limited control over prices. The fact that autonomy reduces prices on average whilst performance pay does not indicate that the rules put in place to curb the agents' corruption are counterproductive because they give authority to another set of agents who can be equally corrupt.

Our paper contributes to the empirical literature on organizations by providing the first experimental evidence on the effect of autonomy on bureaucratic performance. They are consistent with the positive correlation between survey measures of autonomy and the performance of public works projects in Nigeria and Ghana (Rasul & Rogger, 2018; Rasul *et al.* , 2019) as well as the findings that more regulated public bodies are more inefficient at procurement in Italy (Bandiera *et al.* , 2009). The fact that rules meant to curtail corruption backfire in countries which, like Pakistan, have high corruption indexes is particularly striking and suggests that most organizations might be giving too little autonomy to their frontline employees. Our findings are also consistent with Duflo *et al.* (2018) who experimentally decrease the autonomy of environmental inspectors by providing them with a list of firms to inspect rather than letting them use their local knowledge to choose their own. This lowers performance as the inspectors visits more firms but the same number of heavy polluters. Their experiment is not designed to identify the effect of the allocation of authority between different levels of the hierarchy because the researchers themselves are the principal, as it is them who provide the list to the frontline agents.

Our findings highlight the importance of the motivation of agents in contiguous steps of the hierarchy to understand agents responses to performance pay as this can be ineffective if the actions of the agent whose pay is linked to performance can be undone by the response of their superiors.

The remainder of the paper proceeds as follows. In section 2 we present the empirical context for our experiment, and section 3 describes the experimental design. Section 4 develops our conceptual framework to guide our empirical analysis. Section 5 presents our results, and our conclusions are in section 6.

2 Context & Data

In this section we present the context for our empirical application in section 2.1 and our approach to measuring bureaucratic performance in section 2.2.

2.1 Procurement in Punjab

This study takes place in Punjab, Pakistan's most populous province: its population of over 110 million would rank 12th in the world were it a country. The province of Punjab is divided into 36 administrative districts, of which our study took place in 26, covering 80% of the population and the largest districts. These districts were chosen on the basis of logistical feasibility being geographically contiguous and ruling out the remote districts.[1] In this study we work with different agencies in the government of Punjab. These include the Punjab Procurement Regulatory Authority (PPRA) and the Punjab Information Technology Board

[1] Appendix figure

(PITB). We also worked with four administrative agencies - the departments of Education, Health, Agriculture and Communication and Works.

Each office of the government of the province of Punjab has one employee who is designated as the Procurement Officer (PO) who has the legal authority to conduct small and medium sized public procurement purchases.[2] Procurement officers manage procurement on behalf of Public Bodies (PBs) that are allocated budgets under a range of accounting headers (salary, repairs, utilities etc.) including procurement, by the finance department. However, before making payments to vendors, the POs are required to submit their purchases for pre-audit approval by an independent agency of the federal government known as the Accountant General's office (AG). The AG has offices in each of the districts of the province, monitoring the purchase of offices in that district.

A typical procurement process for the purchase of a generic item like the ones we study proceeds in five steps. First, an employee of the of the office makes a request for the purchase of an item (for example, a teacher might request the purchase of pens for the classroom). Second, the PO approves the purchase and surveys the market for vendors who can supply the required item and solicits quotes for the item. Once the PO has received enough quotes for the item, he/she chooses which vendor to allocate the contract to.[3] Third, the vendor delivers the items to the public body and the PO verifies receipt of the items. Fourth, the PO prepares the necessary documentation of the purchase and presents it to the AG office. Fifth, the AG reviews the paperwork. If the AG is satisfied with the documentation, he/she sanctions the payment and gives the PO a check made out to the vendor. If the AG is not satisfied, he/she can demand more thorough documentation that the purchase was made according to the rules.

2.2 Measuring Bureaucratic Performance

The government of Punjab considers that the primary purpose of public procurement transactions is to ensure that "...the object of procurement brings value for money to the procuring agency..." (Punjab Procurement Regulatory Authority, 2014). In line with this, we developed a measure of bureaucratic performance that seeks to measure value for money in the form of the quality-adjusted unit prices paid for the items being purchased by POs. The backbone of our approach is to collect detailed data on the attributes of the items being purchased with which to measure the quality of the items being purchased.

To achieve this, we proceed in two steps. First, we restrict attention to relatively homogeneous goods for which we believe that by collecting detailed enough data we will be able to adequately measure the quality of the item being purchased (similar to the approach taken in Bandiera *et al.* 2009).[4] Second, for these homogeneous goods, we partnered with the Punjab IT Board (PITB) to build an e-governance platform named the Punjab Online Procurement System (POPS). This web-based platform allows offices to enter detailed data on the attributes of the items they are purchasing. We trained over a thousand civil servants in the use of POPS and the departments we worked with required the offices in our experimental sample (as described below) to enter details of their purchases of generic goods into the POPS system.

After running the POPS platform for the two years of the project and cleaning the data entered by the officers, our analysis dataset consists of the 25 most frequently purchased goods: a total of 21,503 purchases of 25 homogeneous goods. Dropping the top and bottom 1% of unit prices results in a dataset of 21,183

[2]The title of this position is known as the "Drawing and Disbursement Officer" of the office.

[3]For very small purchases, only one quote is needed. For most of the purchases we consider, POs must obtain three quotes and then choose the cheapest one.

[4]To do this, we chose accounting codes from the government's chart of accounts that we expected to contain mostly or exclusively generic goods. The list of accounting codes is contained in appendix table A.1.

observations.[5] Figure 1 shows summary statistics of the purchases in the POPS dataset. The 25 items are remarkably homogeneous goods such as printing paper and other stationery items, cleaning products, and other office products. While each individual purchase is small, these homogeneous items form a significant part of the procurement budgets of our offices.

Despite the homogeneous nature of the items being purchased, the prices paid display a remarkable degree of variation. Figure 1 shows this variation for each product, and figure 2 shows the full distribution of prices paid for printing paper. The blue bars show the distribution of raw prices. The orange bars show the distribution of residualized prices using the method described in section section 5.2. Both distributions display dramatic variation, suggesting different bureaucrats are paying very different amounts for identical products.

3 Experimental Design

3.1 Design of Experimental Treatments

To design our treatments, we conducted a series of focus groups and a baseline survey of procurement officers. Our goals were twofold. First, to understand what they perceived to be their performance incentives. Second, to understand what they perceived as the barriers that stopped them from being able to improve the value for money achieved in public procurement.

To elicit procurement officers' perceptions of their incentives to perform procurement well, we asked officers what types of errors would be detrimental to their career progress. Since civil servants in Punjab are not typically paid based on their performance, the main incentive they face to perform well is that their performance is used when evaluating their applications to transfer posts and to progress up the civil service hierarchy. Specifically, we asked officers how detrimental overpaying in their procurement purchases would be, and how detrimental failing to complete the required documentation would be. Figure [XXX] shows the results. While the officers respond that both transgressions would be detrimental for their careers, they report that having incomplete documentation is a severe impediment much more often than overpaying. This stands in clear contrast to the government's stated goal when conducting public procurement—to achieve value for money (Punjab Procurement Regulatory Authority, 2014), and motivates our two treatments.[6]

Our *incentives* treatment aimed to align procurement officers' incentives with the government's by providing officers with financial incentives to improve value for money. To achieve this, we paid procurement officers in this treatment arm bonuses as a function of their performance. Officers' performance was evaluated by a committee established for this purpose. The committee was co-chaired by a well-respected, senior, private-sector auditor and the director of the Punjab Procurement Regulatory Authority (PPRA). Delegates from each of the line departments, the finance department, and the research team rounded out the committee. Based on common practice in the private sector, the committee was tasked with ranking the procurement officers' performance by applying a wholistic assessment to the officer's performance at achieving the aims of public procurement. To seed the discussions, the research team provided an initial ranking of the procurement officers according to our measure of value added described in section 2.2, though the committee were told they had absolute freedom to alter the ranking.

[5] The majority of these outliers are the result of officers adding or omitting zeros in the number of units purchased.

[6] Paragraph 4 of Punjab's procurement rules (Punjab Procurement Regulatory Authority, 2014) states "Principles of procurements.– A procuring agency, while making any procurement, shall ensure that the procurement is made in a fair and transparent manner, the object of procurement brings value for money to the procuring agency and the procurement process is efficient and economical."

Based on the committee's ranking, bonuses were paid. The *gold* group, comprising the top 7.5% of officers, received two months' salary. The *silver* group, the next 22.5% of officers, received one month's salary. The *bronze* group, the next 45% of officers received half of a month's salary. Finally, the remaining 25% of officers did not receive an honorarium. The committee met twice a year. Based on the interim rankings at the middle of the year, bonuses of half the amounts were paid to the officers, which were then credited against the bonuses received in the final ranking at the end of the year.

We made two design choices to increase the salience and credibility of this treatment that are worth noting. First, we chose prize structure that meant that 75% of officers received a prize so that many officers would experience receiving a prize, making our promises of the bonuses more credible. Second, we chose to have the committee meet twice a year. This gave our treatment greater credibility during the second half of the year when the bulk of procurement expenditure takes place. Moreover, the incentive treatment was in place during the pilot year to build credibility so officers already had experience with the treatment when the experimental year began.

Our *autonomy* treatment sought to shift decision-making power towards procurement officers. In our baseline survey we asked respondents why they thought procurement officers don't achieve good value for money. Figure3 shows the results. The three most popular answers were *"Budgets are released late so DDOs [procurement officers] cannot plan appropriately"*; *"AG/DAO requirements are not clear and they do not clear bills without inside connections or payment of speed money"*; and *"DDOs [procurement officers] do not have enough petty cash to make purchases quickly."* Motivated by these responses, these were the three obstacles our treatment targeted.

The autonomy treatment had three parts. First, each office's petty cash balance was increased to Rs. 100,000 (USD 1,000). Petty cash can be used to make payments to vendors without having to seek pre-audit approval from the AG. Circumventing the AG gives more autonomy to procurement officers to make timely payments, use vendors who don't have connections to the AG office, and generally avoid markups imposed by the AG.[7] It also shifts government liquid funds from the finance department down to the spending offices. On the other hand, having more petty cash available means that a procurement officer seeking to embezzle funds has more available on hand, increasing the risk to the taxpayer. Second, the finance department released budget to offices in this treatment arm in two timely installments in August and January, rather than the usual four installments (due each quarter but typically delayed by several weeks).[8] Having more budget available to spend allows offices to plan their spending more effectively, but also gives officers more scope to make larger corrupt deals if they are corrupt. Third, offices were given a checklist prepared by the AG of the documents required to be presented for pre-audit of purchases of different types. This made the pre-audit procedure more transparent and predictable for officers, aiming to improve their bargaining position vis-a-vis the AG, giving them more decision power in the procurement process.

3.2 Experimental Population and Randomization

The experiment was conducted in collaboration with a range of organs of the government of Punjab. The finance department and the Accountant General's (AG) office implemented the autonomy treatment together with the four line departments from which our sample was drawn. We sampled offices from the four largest departments in the government, the departments of education, health, agriculture, and communication &

[7] Petty cash is still subject to all the same legal scrutiny and documentary requirements as ordinary spending during post audit after the conclusion of the financial year. The only difference is that it does not require pre-audit approval by the AG.

[8] Online Appendix figure XXXXX shows the timing of budget distributions in the treatment and control groups

works. Since 10 districts were thought to be considering secession to form their own province at the time of the development of the experiment, we restricted ourselves to 26 of the 36 districts in Punjab, covering over 80% of the population of the province, or over 110 million people. Within these departments and districts we sampled from offices with procurement budgets in the 2012-13 fiscal year of at least Rs. 250,000 (USD 2,500). The experiment was conducted in collaboration with a number of organs of the government of Punjab. The finance department and the Accountant General's (AG) office implemented the autonomy treatment together with the four line departments from which our sample was drawn.

In June 2014, we randomized 688 offices into the four treatment arms, stratifying by district $\times$ department to ensure balance on geographical determinants of prices and composition of demand. Offices were told by their departments that they were part of a a study to evaluate the impact of policy reforms under consideration for rollout across the province and that their participation was mandatory, including entering data into the POPS system and cooperating with occasional survey team visits. With this backing, 587 offices, or 85% of the sample, participated in trainings on the POPS system and on the implications of their treatment status for how they conduct procurement.

Table 2 presents summary statistics on a range of variables in the participating offices. The table shows that the participation rate is balanced across the treatment arms, as are the vast majority of office characteristics and budgetary variables available in the finance department's administrative data. Of the 27 variables presented and six pairwise comparisons across treatment arms per variable (for a total of 162 pairwise comparisons), only 8 are significant at at least the 10% level, of which two at the 5% level, consistent with what we would expect by pure chance. Overall, we conclude that the randomization produced a balanced sample and that compliance was high and balanced across the treatment arms.

4 Conceptual Framework

The starting point of our model is Shleifer & Vishny's(1993) analysis of the institutional determinants of misbehavior of public agencies. They show that how decision-making power is distributed among agencies is an important determinant of the overall level of corruption. In a nutshell, we model the interaction between a (potentially corrupt) purchasing manager and a (potentially corrupt) auditor who is supposed to monitor the purchasing manager. We wish to understand what happens when we give the purchasing manager an extrinsic reason to save money (the incentive treatment) and when we make the manager more independent of the auditor. The predictions that are derived from this simple model will be used to guide the remainder of our empirical analysis.

The goal of this model is not to develop a general theory of the"organization of corruption" (like the ones in Guriev, 2004 and Banerjee *et al.* , 2012). We offer a parsimonious framework that delivers highly stylized predictions.

4.1 Definitions

This simple model captures a situation where administrative decisions are taken by an *agent* and possibly monitored by a *supervisor* with veto power. There is a mass 1 of procurement decisions to be taken.

The agent selects a *markup*$x_a \geq 0$. The most direct interpretation of the markup is a bribe that the agent receives from the supplier of the procured good. The payoff of the agent is given by

$$v_a = u_a(x_a) - k_a(x_a; \theta_a, \lambda; b, r),$$

where the first term u_a is the benefit the agent receives from the markup, and is therefore increasing in x_a. The second term is the cost the agent incurs because of x_a, and is increasing in x_a. The cost can be thought of as a psychological cost of being dishonest or inefficient, or as the risk of being caught, or the disutility the agent experiences if the markup causes an increase in the price of the procured good. The underlying idea is that suppliers tend to incorporate the markup into the price of the procured good. The cost depends on two parameters: $\theta_a \in [0, \infty)$ is the agent's alignment, where $\theta_a = 0$ denotes an agent who experiences only utility from imposing a markup and $\theta_a \to \infty$ denotes an agent who suffers infinitely when the markup is anything but zero, while λ captures the homogeneity of the good being provided, with the idea that markups on more homogeneous products or services (high λ) may be easier to observe and therefore impose a higher cost on the agent. The cost also depends on two policy values. The first, b, represents a monetary incentive given to the agent for paying low procurement prices; a high b increases the cost of demanding a higher markup x_a because of the effect the markup has on prices. The second, r, is the supervisor's power, which may change the agent's markup cost by making him more likely to be subject to sanctions if he chooses a high markup level.

The supervisor too selects a *markup* $x_s \geq 0$. That too can be interpreted as a bribe that is ultimately paid by the supplier of the good, even though in practice it can be delivered by the agent. The supervisor's payoff is given by

$$v_s = u_s(x_s) - k_s(x_a, x_s; \theta_s, \lambda; r)$$

where the first term u_s is the benefit the supervisor receives from the markup. The second term is the cost the supervisor incurs. Importantly, it is increasing in the markups of both players: x_a and x_s. The supervisor cares about her own markup x_s for the same reasons the agent cares about x_a. In addition she is hurt by a high x_a because she may be held responsible for the agent's markup or because she cares about the price of the procured good, which depends on both markups. The supervisor's cost depends on her own alignment θ_s, as well as the homogeneity parameter λ, and the supervisor's power r.[9]

The price of the procured good$p$ depend on some baseline characteristics of the good, which are constant and therefore can be omitted, as well as the markups demanded by the two players, x_a and x_s. The price rises with higher markups.

One can imagine a number of extensive game forms given the payoffs above, perhaps a dynamic bargaining protocol between the two players, possibly including asymmetric information. However, the key predictions we wish to test are captured by a simple complete-information game where both players set their markups simultaneously. Furthermore, we can assume simple quadratic forms for the payoff functions above.

For the agent,

$$v_a = x_a - \frac{1}{2}\lambda(\theta_a + b + r)x_a^2,$$

For the supervisor, let us assume that that supervisory power$r$ is the probability she has veto power over the purchase (something that will have a natural interpretation in our setting). The agent chooses his markup knowing$r$ but not whether the supervisor can veto that particular purchase. The supervisor's payoff is

[9]One could also assume that the agent's cost depends on the supervisor's markupx_s. The results would be unchanged.

assumed to be:[10]

$$v_s = \begin{cases} x_s - \frac{1}{2}\theta_s (x_a + x_s)^2 & \text{with probability } r; \\ 0 & \text{with probability } 1-r. \end{cases}$$

We also assume that the final price of the procured good is linear in the two markups. With normalization, we have a price $p = x_a + x_s$ when the supervisor has veto power and $p = x_a$ when she does not. The average price is therefore

$$\bar{p} = x_a + r x_s.$$

4.2 Equilibrium Price

We begin by deriving the markups the two agents demand:

Proposition 1. *In equilibrium, the agent chooses*

$$x_a = \frac{1}{\lambda(\theta_a + b + r)};$$

the supervisor chooses

$$x_s = \frac{1}{\theta_s} - \frac{1}{\lambda(b + r + \theta_a)};$$

and the average price is

$$\bar{p} = \frac{1-r}{\lambda(b + r + \theta_a)} + \frac{r}{\theta_s}.$$

Proof. The agent's first-order condition yields x_a immediately. Instead, x_s is obtained by replacing x_a in the supervisor's first-order condition

$$\frac{d}{dx_s}\left(x_s - \frac{1}{2}\theta_s\left(\frac{1}{\lambda(\theta_a + b + r)} + x_s\right)^2\right) = 0$$

$$\frac{1}{\lambda(b + r + \theta_a)}(\theta_s - b\lambda - r\lambda - \lambda\theta_a + b\lambda\theta_s x_s + r\lambda\theta_s x_s + \lambda\theta_a\theta_s x_s) = 0$$

The average price is given by:

$$\begin{aligned} \bar{p} &= \frac{1}{\lambda(\theta_a + b + r)} + r\left(\frac{-\theta_s + b\lambda + r\lambda + \lambda\theta_a}{\lambda(b\theta_s + r\theta_s + \theta_a\theta_s)}\right) \\ &= \frac{\theta_s - r\theta_s + r^2\lambda + br\lambda + r\lambda\theta_a}{\lambda\theta_s(b + r + \theta_a)}. \end{aligned}$$

□

As one would expect, the average price is decreasing in the alignment parameters of the two agents. If θ_a or θ_s increase, expected price goes down. The effect of r is subtle. For now, notice that when the supervisor is powerless ($r = 0$), we have

$$\bar{p} = \frac{1}{\lambda(b + \theta_a)},$$

[10] Note that the agent's payoff can be seen as the expected payoff from:

$$v_a = \begin{cases} x_a - \frac{1}{2}\lambda(\theta_a + b + 1)x_a^2 & \text{with probability } r \\ x_a - \frac{1}{2}\lambda(\theta_a + b)x_a^2 & \text{with probability } 1-r \end{cases}$$

namely the average price is fully determined by the agent's incentive to keep his markup low. Instead, when the supervisor is omnipresent ($r = 1$), the average price is determined by the supervisor's characteristics:

$$\bar{p} \to \frac{1}{\theta_s}.$$

4.3 Comparative Statics

Given our field experiment, we are interested in predicting the effect of changes in the two policy variables (partially) under our control: supervisory power r and agent incentive b.

Let us begin with agent incentive:

Proposition 2. *(i) An increase in incentive intensity b reduces $\bar{p}$.*

(ii) An increase in b causes a higher percentage decrease in $\bar{p}$ if θ_s is large.

(iii) An increase in b causes a lower percentage decrease in $\bar{p}$ if λ is large.

Proof. (i):

$$\frac{\partial \bar{p}}{\partial b} = \frac{d}{db}\left(\frac{\theta_s - r\theta_s + r^2\lambda + br\lambda + r\lambda\theta_a}{\lambda\theta_s\left(b + r + \theta_a\right)}\right) = -\frac{1}{\lambda}\frac{1-r}{\left(b+r+\theta_a\right)^2} < 0$$

(ii):

$$\frac{\partial}{\partial \theta_s}\left(\frac{\frac{\partial \bar{p}}{\partial b}}{\bar{p}}\right) = \frac{d}{d\theta_s}\left(\frac{-\frac{1}{\lambda}\frac{1-r}{(b+r+\theta_a)^2}}{\frac{\theta_s - r\theta_s + r^2\lambda + br\lambda + r\lambda\theta_a}{\lambda\theta_s(b+r+\theta_a)}}\right) = -\frac{(1-r)\,r\lambda}{\left(\theta_s - r\theta_s + r^2\lambda + br\lambda + r\lambda\theta_a\right)^2} < 0$$

(iii):

$$\frac{\partial}{\partial \lambda}\left(\frac{\frac{\partial \bar{p}}{\partial b}}{\bar{p}}\right) = \frac{d}{d\lambda}\left(\frac{-\frac{1}{\lambda}\frac{1-r}{(b+r+\theta_a)^2}}{\frac{\theta_s - r\theta_s + r^2\lambda + br\lambda + r\lambda\theta_a}{\lambda\theta_s(b+r+\theta_a)}}\right) = r\theta_s\frac{1-r}{\left(\theta_s - r\theta_s + r^2\lambda + br\lambda + r\lambda\theta_a\right)^2}$$

□

All three parts of the proposition are intuitive. An agent who is incentivized to save will reduce his markup. The supervisor may take advantage of this reduction by increasing her own demand x_s, but the overall effect on price must be negative. The percentage effect on the average price depend on how large the markups were to start with, with θ_s and λ having opposite effects.

For supervisory power, we have a more complex result:

Proposition 3. *An increase in supervisory power r decreases average price $\bar{p}$ if and only if the supervisor is sufficiently aligned: $\theta_s > \bar{\theta}_s$ for some $\bar{\theta}_s$.*

Proof. Note that

$$\begin{aligned}\frac{d\bar{p}}{dr} &= \frac{d}{dr}\left(\frac{\theta_s - r\theta_s + r^2\lambda + br\lambda + r\lambda\theta_a}{\lambda\theta_s\left(b + r + \theta_a\right)}\right) \\ &= \frac{\lambda b^2 + 2\lambda br + 2\lambda b\theta_a - \theta_s b + \lambda r^2 + 2\lambda r\theta_a + \lambda\theta_a^2 - \theta_s\theta_a - \theta_s}{\lambda\theta_s\left(b + r + \theta_a\right)^2}.\end{aligned}$$

Therefore,

$$\frac{d\bar{p}}{dr} < 0 \iff \theta_s > \lambda\frac{b^2 + r^2 + \theta_a^2 + 2br + 2b\theta_a + 2r\theta_a}{b + \theta_a + 1}.$$

□

To understand the proposition, note that supervisory power has two effects. On the positive side, an increase in r induces the agent to lower her markup through the monitoring pressure in k_a. On the negative side, more supervisory power increases the supervisor's ability to extract a markup. The latter effect is akin to double marginalization in industrial organization or to Shleifer & Vishny's(1993) result that competition between veto-yielding corrupt agencies increases corruption. The positive side prevails when the supervisor is sufficiently aligned (θ_s is high). The negative side prevails when the supervisor is misaligned.

5 Empirical Analysis

With the conceptual framework of section 4 to guide the analysis, this section analyzes the impacts of the experiment on bureaucratic performance.

5.1 Measuring quality

Whilst goods are largely homogeneous there can be some vertical differentiation. To isolate the effect of treatment on prices we need to adjust for differences in quality, if any. We do so in three ways. First, we can control directly for the full set of attributes collected in POPS for each item. This measure of quality has the advantage of being very detailed, but comes at the cost of being high-dimensional. Our other two measures reduce the dimensionality of the quality controls. To do so, we run hedonic regressions using data from the control group to attach prices to each of the item's attributes. We run regressions of the form

$$p_{igto} = \mathbf{X}_{igto}\lambda_g + \rho_g s_{igto} + \gamma_g + \varepsilon_{igto}$$

where p_{igto} is the log unit price paid, s_{igto} is the size of the purchase, γ_g are good fixed effects, and $\mathbf{X}_{igto}$ are the attributes of good$g$.

Our second measure of quality uses the estimated prices for the attributes $\hat{\lambda}_g$ to construct a scalar measure of quality $q_{igto} = \sum_{j\in A(g)}\hat{\lambda}_j X_j$ where $A(g)$ is the set of attributes of item g. q_{igto} can therefore be interpreted as the expected price paid for a good with these attributes if purchased by the control group, aggregating the high-dimensional vector of attributes down to a scalar. Finally, our third measure of quality studies the estimated $\hat{\lambda}_g$s for each item and partitions purchases into high and low quality purchases based on the $\hat{\lambda}_g$s that are strong predictors of prices in the control group.

5.2 Identification

To estimate the treatment effects on bureaucratic performance we estimate equations of the form

$$p_{igto} = \alpha + \sum_{k=1}^{3}\eta_k \text{Treatment}_o^k + \beta q_{igto} + \rho_g s_{igto} + \delta_s \text{Department}_o \times \text{District}_o + \gamma_g + \varepsilon_{igto} \quad (1)$$

where p_{igto} is the log unit price paid in purchase i of good g at time t by office o; s_{igto} is the quantity purchased to capture good-specific bulk discounts; δ_s and γ_g are stratum and good fixed effects, respectively; and q_{igto} is the quality of the item. The regression is weighted by expenditure shares in the control group so that treatment effects can be interpreted as effects on expenditure, and the residual term ε_{igto} is clustered at the cost centre level.

The coefficients η_k estimate the causal effect of treatment k on unit prices under the assumption of stable unit treatment values (SUTVA). This might be violated if, for example, the AG extracts more from the offices in the control group because it is more difficult to extract from offices in the autonomy treatment. In practice, this is unlikely to affect our estimates because, as shown in Appendix figureA.2, AG officers have typically fewer than 20% of their cost centers in any treatment group.

The fact that we observe control cost centers before and after the roll out of autonomy also allows us to test SUTVA directly. To do so we estimate whether cost centers pay lower prices in year 1 (before the roll out of the autonomy treatment) than in year 2 (after the roll out) as a function of the share of autonomy cost centers in their same AG office. The evidence in Appendix table ZZ supports SUTVA as the prices in the control group are not affected by the share of autonomy cost centers that belong to the same AG office.

The coefficients η_k estimate the causal effect of treatment k on *quality-adjusted* unit prices under the assumption that treatment do not affect the quality of the items purchased. If this is violated quality is an endogenous outcome of the treatments and the η_k coefficients estimate a combination of the treatment effects on quality-adjusted prices and the composition of purchases. To see this, consider a simplified version of our setting. Suppose that purchase are associated with potential prices $p\left(D,q\right)$ depending on a binary treatment $D\in\{0,1\}$ and binary quality $q\in\{0,1\}$ and potential quality levels $q\left(D\right)$ depending on treatment. The random assignment in the experiment implies that the potential outcomes are independent of treatment status conditional on the randomization strata S_i:$\{p_i\left(D,q\right),q_i\left(D\right)\}\perp D_i|S_i$. We can now see that a comparison of expected prices between treated and control units conditional on quality combines a treatment effect on price with a potential composition effect coming from changes in the set of purchases at high or low quality in treatment versus control units:

$$\begin{aligned}\mathbb{E}\left[p|D=1,q=1\right]-\mathbb{E}\left[p|D=0,q=1\right]=&\underbrace{\mathbb{E}\left[p\left(1,1\right)|q\left(1\right)=1\right]-\mathbb{E}\left[p\left(0,1\right)|q\left(1\right)=1\right]}_{\text{treatment effect on price}}\\&+\underbrace{\mathbb{E}\left[p\left(0,1\right)|q\left(1\right)=1\right]-\mathbb{E}\left(p\left(0,1\right)|q\left(0\right)=1\right)}_{\text{composition effect}\neq 0?}\end{aligned} \quad (2)$$

To provide support for this assumption, we take two steps. First, we can directly estimate treatment effects on the quality of the items purchased using our quality measures described above to gauge the magnitude of the potential composition effect.Second, we exploit our baseline data from year 1 of the project to estimate treatment effects through a difference in differences approach, which allows us to control for office fixed effects so that we exploit only within-office changes in prices, holding constant the composition effect $\mathbb{E}\left[p\left(0,1\right)|q\right]$.

5.3 Average Treatment Effects

We begin by studying the overall impact of the experiment. Table 3 shows the average treatment effects estimated using equation (1). Below each coefficient we report its clustered standard error in parentheses and the p-value from randomization inference under the null hypothesis of no treatment effect for any office in square brackets.[11] Column 1 estimates equation (1) without controlling for the quality of the item. The remaining columns control for the item's quality in different ways. In column 2, we control directly for all

[11] We thank Young (2017) for producing the `randcmd` package for stata that greatly facilitates this.

the items' attributes. In column 3 we control for the scalar quality measure, and in column 5 we control for the coarse quality measure.

Three key findings emerge. First, the point estimates of the impacts of the treatments are negative for all three treatments. However, the average impact of the incentives treatment is statistically indistinguishable from zero. The autonomy treatment reduces average quality-adjusted prices paid by 8%, indicating that giving bureaucrats greater autonomy leads them to use it in the interests of the taxpayers. Viewed through the lens of the model in section 4, this implies that the accountant general is relatively more misaligned with the interests of the principal $\left(\theta_s < \tilde{\theta}\right)$. Second, columns 4 and 6 show no discernible impact of the incentives and autonomy treatments on the quality of the items being purchased. The combined treatment creates an increase of around 5% in the quality of the items being purchased. Third, the findings on the impact of the treatments on quality-adjusted prices paid are robust to alternative quality measures.

To address potential concerns about the estimation of treatment effects on quality-adjusted prices when the treatments also might affect quality, we also estimate the impact of the treatments using a difference in differences approach comparing the evolution of prices and quality between year 1 and year 2 of the experiment in the different treatment groups. This allows us to control for office fixed effects to control for the quality composition of purchases in year 1. We estimate

$$p_{igto} = \alpha + \sum_{k=1}^{3} \eta_k \text{Treatment}_o^k \times \text{Year2}_t + \beta q_{igto} + \rho_g s_{igto} + \gamma_g + \psi_o + \xi_t + \varepsilon_{igto} \tag{3}$$

which extends equation (1) to include year 1 data, time fixed effects, and office fixed effects. Table 4 presents the results. The findings are remarkably similar to those from estimating (1) in table 3. The point estimates on the treatment effects of autonomy and the combined treatment are larger, but not statistically different from their counterparts in table 3 suggesting that any changes in the quality composition of the items purchased does not affect the estimation of treatment effects on quality-adjusted prices paid.

The treatments lower quality-adjusted prices paid without affecting the quality of the items purchased. To investigate the impacts of the treatments on the quantities purchased and the composition of expenditure, we aggregate the purchases to the office-good-month level, valuing each purchase using the price predicted for the purchase if purchased by the control group in year 1 (as in the scalar quality measure). That is, for each purchase, the control group-weighted quantity is $e_{igto} = \exp\left(\hat{p}_{igto} + s_{igto}\right)$ where $\hat{p}_{igto}$ is the predicted log unit price for the purchase using the control group's prices, and s_{igto} is the log number of units purchased. Using the aggregated data we then estimate

$$e_{gto} = \gamma_g + \sum_{k=1}^{3} \eta_{kg} \text{Treatment}_o^k + \psi_o + \xi_t + \varepsilon_{gto} \tag{4}$$

where e_{gto} is the quantity purchased of good g in month t by office o; γ_g, ψ_o, and ξ_t are good, office and month fixed effects respectively; and the η_{kg} are good-specific treatment effects. Table5 shows the results. Of the 75 estimated η_{kg} treatment effects, two are statistically significant at the 5% level, consistent with what would be expected purely by chance. As a result, consistent with government offices having inelastic demand for their inputs, we conclude that there is no evidence that any of the treatments affected the composition of offices' expenditure or the overall amount they purchase.

5.4 Supervisor Alignment

Our conceptual framework in section 4 predicts that we should expect to see heterogeneity of the treatment effects with different effects according to the alignment of the accountant general θ_s. In this section we estimate heterogeneous treatment effects using a proxy for the alignment of the accountant general.

Our proxy for the alignment of the accountant general combines two elements. First, we note that the main power of the accountant general is to delay payments and require additional paperwork. Second, in Punjab, as is common around the world, government offices' budgets lapse at the end of the fiscal year if they remain unspent. As documented in Liebman & Mahoney (2017) in the US context, lapsing budgets lead to a rush to spend at the end of the year. Combined with the first element, we expect this end of year rush to be stronger in districts where the accountant general delays payments more. Our proxy for the accountant general's misalignment $\hat{\theta}_s$ is therefore the fraction of purchases in the district in year 1 that were approved in the last month of the fiscal year.

We augment equation (1) to include interactions with our proxy $\hat{\theta}_s$ as follows

$$p_{igto} = \alpha + \sum_{k=1}^{3} \left(\eta_k \text{Treatment}_o^k + \zeta_k \text{Treatment}_o^k \times \hat{\theta}_s \right) + \beta q_{igto} + \rho_g s_{igto} + \delta_s \text{Department}_o \times \text{District}_o + \gamma_g + \varepsilon_{igto}$$

Figure 4 and table 6 show the results. Three key findings emerge consistent with the predictions of the model. First, the incentives treatment does reduce prices when the supervisor is relatively more aligned (low$\hat{\theta}_s$), and the treatment effect of incentives shrinks as supervisors get less aligned. If purchases are approved evenly throughout the fiscal year incentives reduce prices by 10%. Second, the autonomy and combined treatments reduce prices more strongly when the supervisor is relatively misaligned (high$\hat{\theta}_s$) and the treatment effects are close to zero as supervisors are more aligned. Third, the effects of the incentives and autonomy treatments cancel out in the combined treatment when the supervisor is relatively aligned (low$\hat{\theta}_s$) but as the supervisor becomes less aligned the effect of autonomy dominates in the combined treatment.

5.5 Good Homogeneity

When the goods being purchased are homogeneous, deviations from market prices are easier to detect, making it harder for bureaucrats and supervisors to add markups. This means we should expect to see less padding of prices and also weaker treatment effects when the good being purchased is more homogeneous. In our conceptual framework of section 4, the (inverse of the) degree of homogeneity of the good being purchased is captured by the parameter λ. For larger values of λ our framework predicts that we will see larger treatment effects of both the incentives and autonomy treatments.

To construct an empirical counterpart for the theoretical parameter λ we again exploit the richness of our data on the attributes of the items being purchased. We consider goods to be more homogeneous when our attributes are able to predict a larger share of the variation in unit prices in the control group in year 1. Specifically, for each good we calculate $\phi_g = \mathsf{Var}\left(p_{ig}\right) / \mathsf{Var}\left(\hat{p}_{ig}\right)$ where p_{ig} is the log unit price paid, $\hat{p}_{ig}$ are the prices predicted using the item's attributes, and variances are taken across purchases of good g in year 1 in the control group.

Analogously to the analysis in section 5.5, we augment equation (1) to include interactions with our proxy ϕ_g as follows

$$p_{igto} = \alpha + \sum_{k=1}^{3} \left(\eta_k \text{Treatment}_o^k + \kappa_k \text{Treatment}_o^k \times \phi_g \right) + \beta q_{igto} + \rho_g s_{igto} + \delta_s \text{Department}_o \times \text{District}_o + \gamma_g + \varepsilon_{igto}$$

Figure 5 and table 7 show the results. Consistent with the predictions of the theory, we find that all three treatments have stronger effects on reducing prices when the good purchased is less homogeneous (higher ϕ_g).

6 Conclusion

[TBD]

References

BANDIERA, ORIANA, PRAT, ANDREA, & VALLETTI, TOMMASO. 2009. Active and Passive Waste in Government Spending: Evidence from a Policy Experiment. *American Economic Review*, **99**(4), 1278–1308.

BANERJEE, A, HANNA, R, & MULLAINATHAN, S. 2012. Corruption. *Chap. 27, pages 1109–1147 of: Handbook of Organizational Economics*.

DUFLO, ESTHER, GREENSTONE, MICHAEL, PANDE, ROHINI, & RYAN, NICHOLAS. 2018. The Value of Discretion in the Enforcement of Regulation: Experimental Evidence and Structural Estimates from Environmental Inspections in India. *Econometrica*, **86**, 2123–2160.

GURIEV, SERGEI. 2004. Red tape and corruption. *Journal of Development Economics*, **73**(2), 489–504.

LIEBMAN, JEFFREY B., & MAHONEY, NEALE. 2017. Do Expiring Budgets Lead to Wasteful Year-End Spending? Evidence from Federal Procurement. *American Economic Review*, **107**, 3510–3549.

PUNJAB PROCUREMENT REGULATORY AUTHORITY. 2014. *Punjab Procurement Rules No.ADMN(PPRA) 10–2/2013*.
https://ppra.punjab.gov.pk/system/files/Final%20Notified%20PPR-2014%20%28ammended%20upto%2006.01.2016%29.pdf.

RASUL, IMRAN, & ROGGER, DANIEL. 2018. Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service. *Economic Journal*, **128**, 413–446.

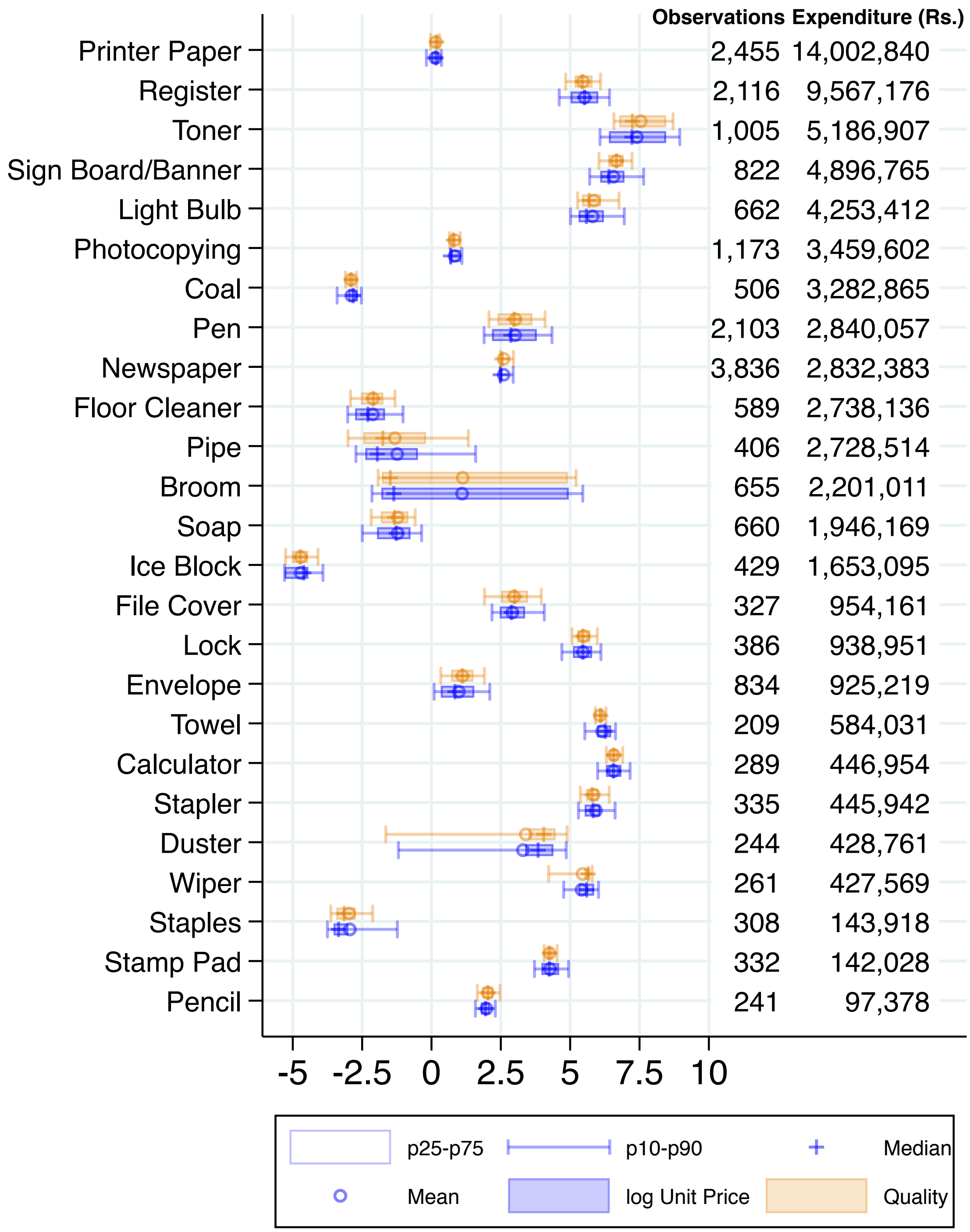
RASUL, IMRAN, ROGGER, DANIEL, & WILLIAMS, MARTIN J. 2019. *Management and Bureaucratic Effectiveness: Evidence from the Ghanaian Civil Service*. Mimeo: University College London.

SHLEIFER, ANDREI, & VISHNY, ROBERT W. 1993. Corruption. *Quarterly Journal of Economics*, **108**, 599–617.

YOUNG, ALWYN. 2017. *Channelling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results*. Mimeo: LSE.

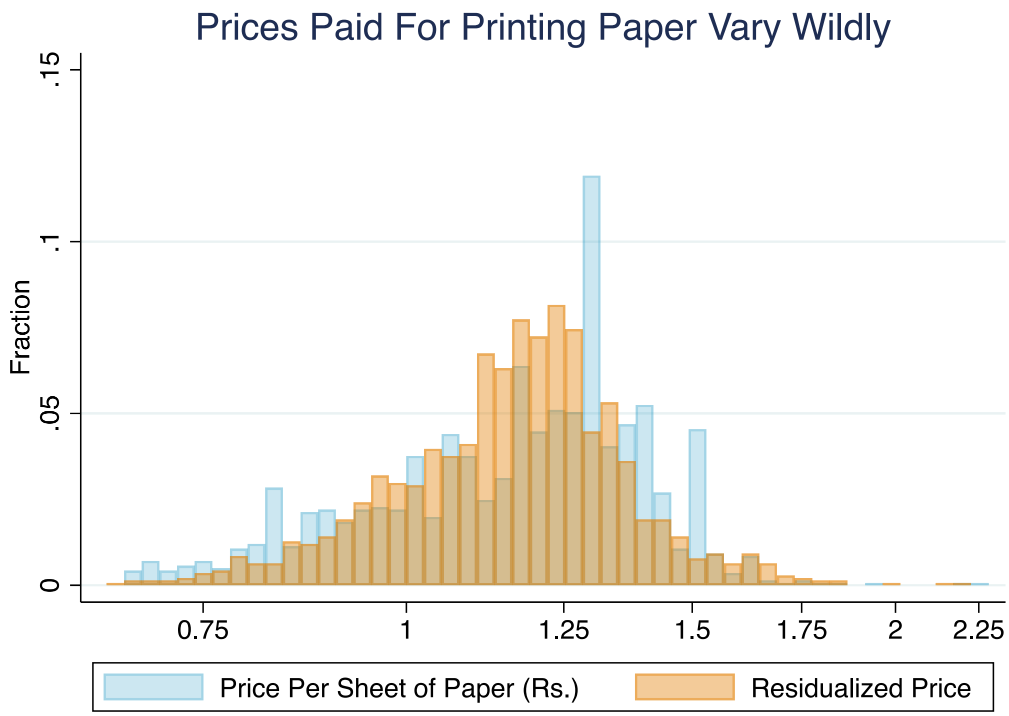
Figures & Tables

FIGURE 1: SUMMARY STATISTICS



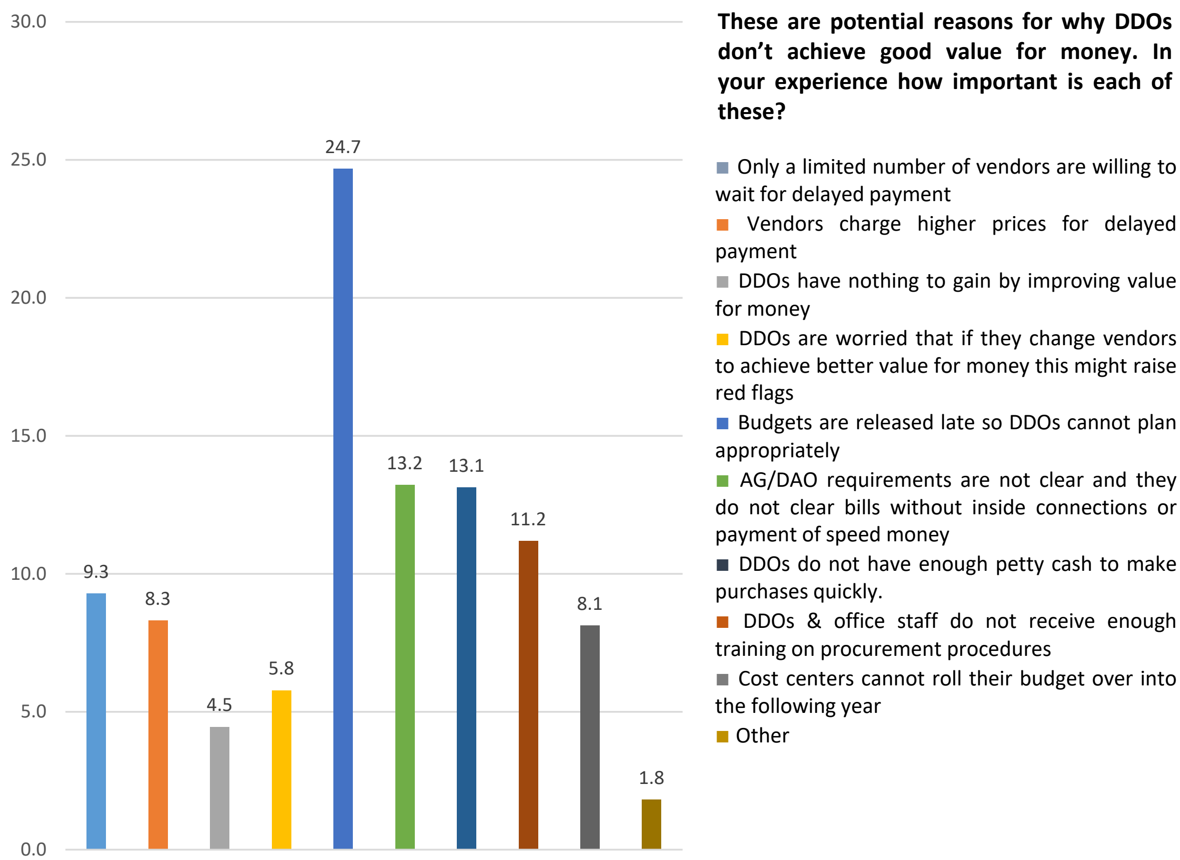
Notes: The figure displays summary statistics for the purchases of the goods in our cleaned purchase sample. The figure summarizes the log unit prices paid for the goods, the number of purchases of each good, and the total expenditure on the good (in Rupees) in the sample.

FIGURE 2: PRICES PAID FOR PRINTING PAPER VARY WILDLY



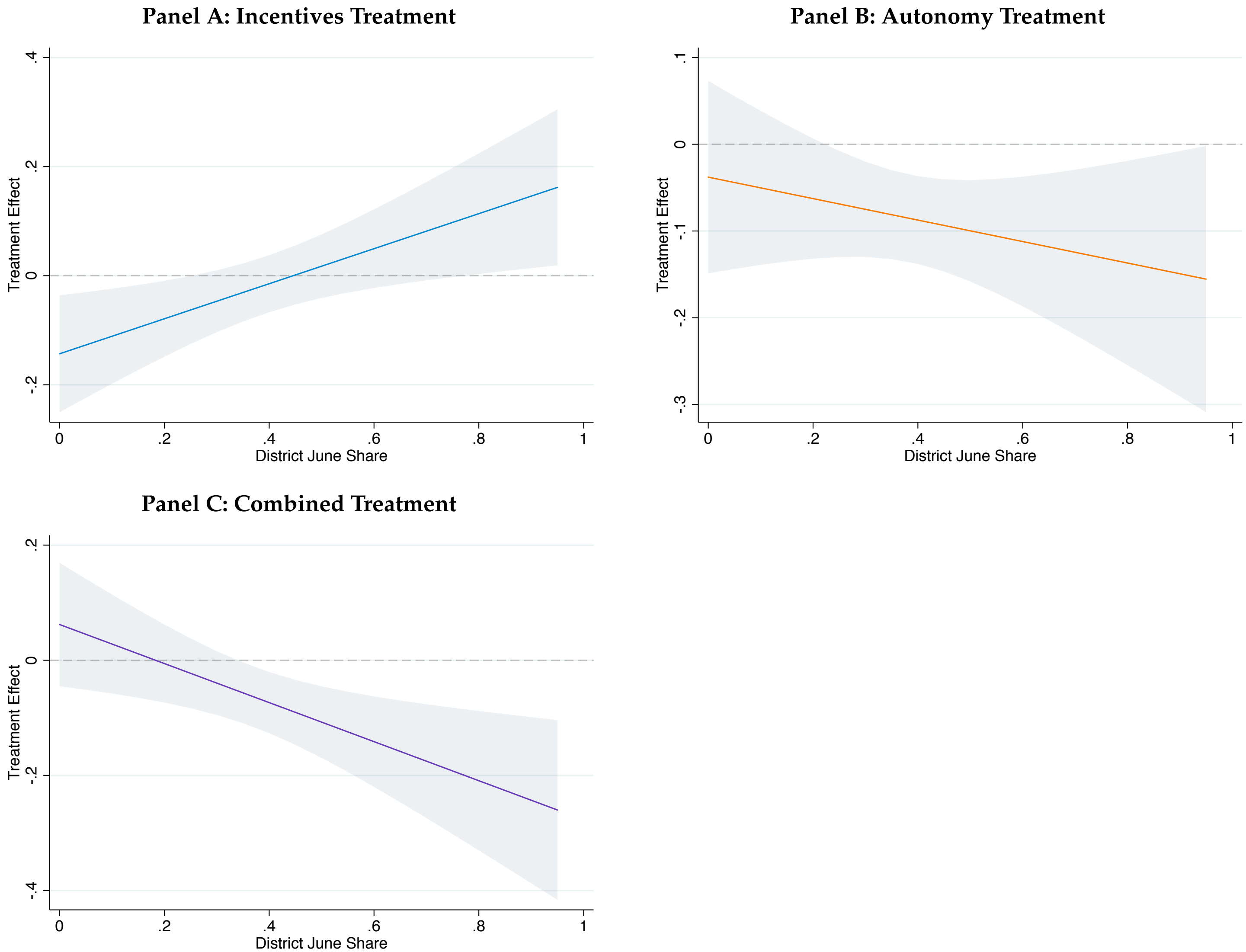
Notes: The figure shows the distribution of log unit prices paid for sheets of printing paper—an extremely homogeneous product—by procurement officers in our sample. The blue bars show the distribution of raw prices. The orange bars show the distribution of residualized prices using the method described in section XXX. Both distributions display dramatic variation, suggesting different bureaucrats are paying very different amounts for identical products.

FIGURE 3: PROCUREMENT OFFICERS' PERCEIVED OBSTACLES TO IMPROVING PERFORMANCE



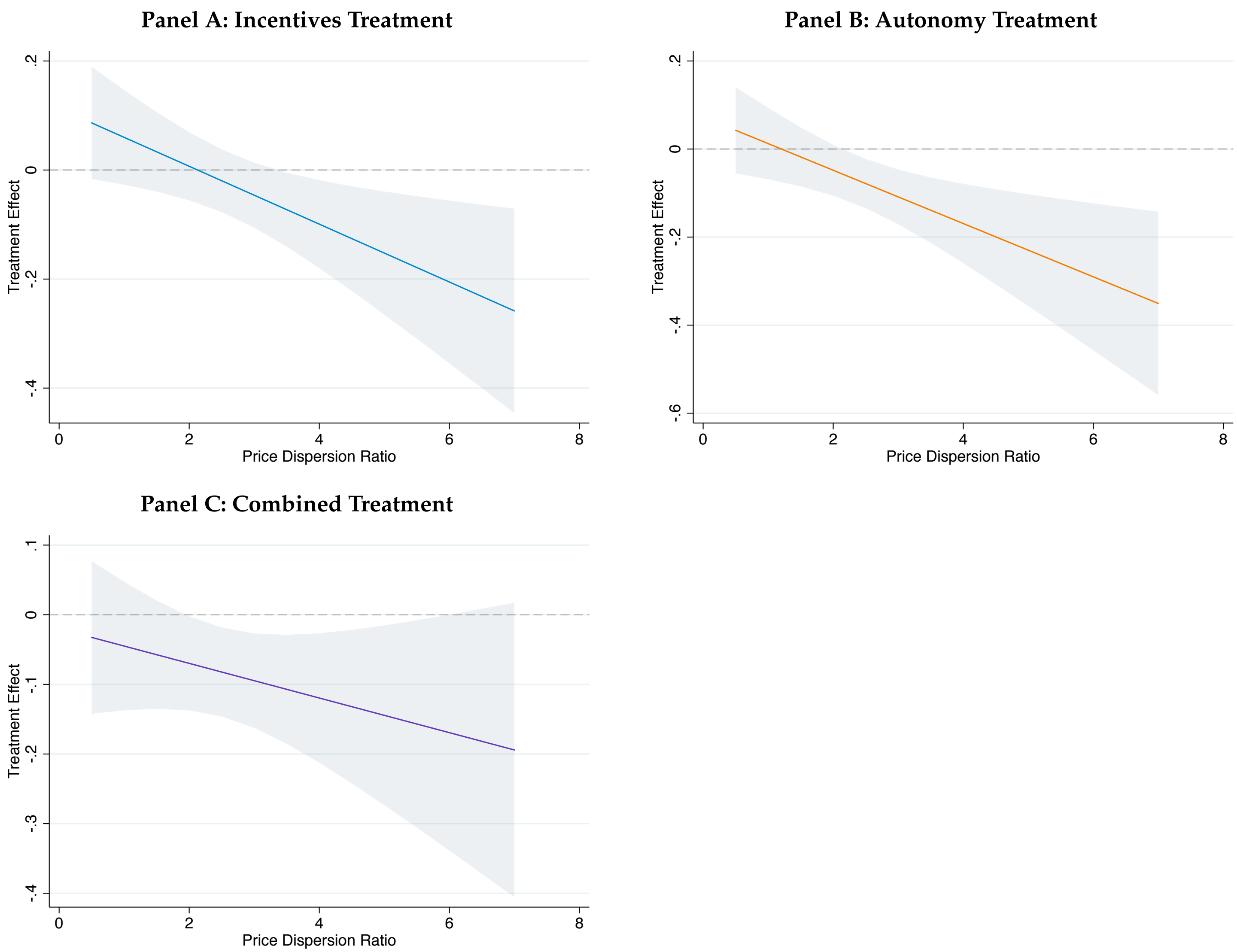
Notes: The figure displays procurement officers' responses to the question"*These are the potential reasons why DDOs [procurement officers] don't achieve good value for money. In your experience, how important is each of these?*" The height of each bar is the number of officers answering that the reason is"important" or"very important".

FIGURE 4: HETEROGENEITY OF TREATMENT EFFECTS BY SUPERVISOR ALIGNMENT



Notes: The figure shows heterogeneity of the three treatment effects by the degree of misalignment of the district's accountant general. Accountants general are classified according to the degree to which purchase approvals are bunched at the end of the fiscal year in June 2015 (year 1 of the project). Panel A shows heterogeneity of the incentives treatment effect. Panel B shows the autonomy treatment and panel C shows the combined treatment.

FIGURE 5: HETEROGENEITY OF TREATMENT EFFECTS BY GOOD HOMOGENEITY



Notes: The figure shows heterogeneity of the three treatment effects by the degree of homogeneity of the good being purchased. Goods g are classified according to the ratio of the total variance in their log unit prices to the variance explained by the item's attributes $\phi_g = \mathsf{Var}(p) / \mathsf{Var}(\hat{p})$, where $\hat{p}$ is the price as predicted using the item's attributes and variances are across purchases of that good in the control group in year 1 of the project. Panel A shows heterogeneity of the incentives treatment effect. Panel B shows the autonomy treatment and panel C shows the combined treatment.

TABLE 1: PROJECT TIMELINE

Year 1: July 2014 – June 2015	
06/14	Cost Centers allocated to treatment arms
07–08/14	Trainings on POPS and treatment brochures
08–09/14	Follow-up trainings on POPS
03–04/15	Baseline Survey
Year 2: July 2015 – June 2016	
07–10/15	Refresher trainings on treatments and POPS
10/15	Cash in Hand rolled out
03–04/16	Midline Survey
04/16	Performance Evaluation Committee Midline Meeting
06/16	Experiment Ends
Post-Experiment	
08-09/16	Endline Survey Part 1 & Missing Data Collection
02/17	Performance Evaluation Committee Endline Meeting
02–03/17	Endline Survey Part 2

TABLE 2: BALANCE ACROSS TREATMENT ARMS

	Control	Regression Coefficients			Joint Test
	mean/sd	Incentives	Autonomy	Both	All = 0
Office Characteristics					
	1.01	−0.007	0.033	0.012	2.360
Number of Public Bodies	{0.086}	(0.007)	(0.024)	(0.013)	[0.071]*
		[0.346]	[0.211]	[0.460]	[0.264]
	1.26	0.069	0.222	0.186	2.427
Number of Accounting Entities	{0.635}	(0.086)	(0.100)**	(0.087)**	[0.065]*
		[0.407]	[0.028]**	[0.038]**	[0.076]*
	0.30	0.059	−0.004	0.025	0.549
Agriculture	{0.461}	(0.056)	(0.055)	(0.055)	[0.649]
		[0.328]	[0.943]	[0.636]	[0.658]
	0.05	−0.031	0.009	0.007	1.826
Communication and Works	{0.222}	(0.022)	(0.027)	(0.027)	[0.141]
		[0.148]	[0.746]	[0.769]	[0.164]
	0.04	−0.011	0.003	−0.005	0.145
Health	{0.206}	(0.023)	(0.025)	(0.024)	[0.933]
		[0.624]	[0.968]	[0.813]	[0.934]
	0.60	−0.016	−0.008	−0.028	0.084
Higher Education	{0.491}	(0.058)	(0.058)	(0.058)	[0.969]
		[0.786]	[0.897]	[0.619]	[0.972]
Year-1 Budget & Expenditure Variables					
	13,073,482	789,962	−4,064,330	1,872,966	0.229
Budget: Total Nonsalary	{88,003,879}	(11,800,824)	(8,344,191)	(11,452,507)	[0.876]
		[0.965]	[0.708]	[0.908]	[0.949]

Continued on next page

Table 2 – *Continued from previous page*

	Control	Regression Coefficients			Joint Test
	mean/sd	**Incentives**	**Autonomy**	**Both**	**All = 0**
	10,240,016	2,705,330	−1,806,740	3,796,925	0.187
Expenditure: Total Nonsalary	{79,128,130}	(10,858,444)	(7,606,140)	(10,708,379)	[0.905]
		[0.847]	[0.858]	[0.788]	[0.955]
	9,231,007	970,924	−1,845,483	2,755,046	0.151
Budget: Operating Expenses	{70,918,641}	(8,729,335)	(6,912,254)	(9,193,999)	[0.929]
		[0.923]	[0.834]	[0.815]	[0.951]
	8,653,280	770,356	−1,659,969	2,734,579	0.139
Expenditure: Operating Expenses	{69,527,473}	(8,276,316)	(6,779,758)	(8,974,808)	[0.937]
		[0.931]	[0.845]	[0.805]	[0.947]
	1,273,849	−1,183,522	−859,849	−806,348	1.258
Budget: Physical Assets	{12,074,957}	(1,035,661)	(1,063,934)	(1,116,504)	[0.288]
		[0.324]	[0.615]	[0.681]	[0.522]
	802,783	−712,546	−401,845	−353,443	1.209
Expenditure: Physical Assets	{7,551,173}	(648,167)	(689,499)	(762,396)	[0.306]
		[0.386]	[0.736]	[0.791]	[0.598]
	168,896	530,296	110,288	268,101	1.030
Budget: Repairs & Maintenance	{634,550}	(441,244)	(112,243)	(263,385)	[0.379]
		[0.286]	[0.434]	[0.461]	[0.611]
	127,873	467,731	87,380	228,127	1.078
Expenditure: Repairs & Maintenance	{530,324}	(386,129)	(79,461)	(230,745)	[0.358]
		[0.265]	[0.340]	[0.495]	[0.516]
	8,040,553	−539,182	−2,292,622	740,120	0.125
Budget: Accounting Codes Potentially Including Generic Goods	{72,813,864}	(7,735,750)	(6,819,867)	(7,916,043)	[0.945]
		[0.943]	[0.775]	[0.939]	[0.951]

Continued on next page

Table 2 – *Continued from previous page*

	Control	Regression Coefficients			Joint Test
	mean/sd	Incentives	Autonomy	Both	All = 0
	7,675,569	−627,134	−2,128,920	766,982	0.114
Expenditure: Accounting Codes Potentially Including Generic Goods	{71,785,069}	(7,477,831)	(6,724,712)	(7,793,480)	[0.952]
		[0.931]	[0.792]	[0.939]	[0.955]
	1,147,625	−71,366	−8,890	257,414	0.057
Budget: Analysis Sample Accounting Codes	{8,265,221}	(843,746)	(798,445)	(977,030)	[0.982]
		[0.945]	[0.991]	[0.838]	[0.990]
	994,383	13,150	92,919	316,763	0.057
Expenditure: Analysis Sample Accounting Codes	{7,291,448}	(756,439)	(722,053)	(907,390)	[0.982]
		[0.984]	[0.926]	[0.787]	[0.992]
Year-2 Budget Variables					
	12,496,735	−7,238,563	−2,544,847	−5,429,337	0.550
Budget: Total Nonsalary	{98,045,551}	(8,515,884)	(9,950,058)	(8,722,782)	[0.648]
		[0.586]	[0.880]	[0.717]	[0.866]
	9,788,214	−5,759,401	−1,821,126	−4,042,359	0.512
Budget: Operating Expenses	{80,862,229}	(7,039,005)	(8,532,367)	(7,220,443)	[0.674]
		[0.625]	[0.898]	[0.757]	[0.870]
	1,386,751	−1,032,128	−1,058,950	−1,182,607	0.570
Budget: Physical Assets	{12,461,547}	(1,082,454)	(1,081,903)	(1,075,430)	[0.635]
		[0.459]	[0.420]	[0.288]	[0.743]
	264,305	−92,015	101,472	−55,864	0.472
Budget: Repairs & Maintenance	{1,540,930}	(135,432)	(241,530)	(156,835)	[0.702]
		[0.656]	[0.805]	[0.838]	[0.902]
	9,173,648	−5,851,536	−2,307,954	−4,923,463	0.420
Budget: Accounting Codes Potentially Including Generic Goods	{84,904,043}	(7,357,602)	(8,615,731)	(7,460,924)	[0.739]

Continued on next page

Table 2 – *Continued from previous page*

	Control	Regression Coefficients			Joint Test
	mean/sd	Incentives	Autonomy	Both	All = 0
		[0.650]	[0.883]	[0.728]	[0.968]
	1,485,532	−427,719	−291,308	−617,152	0.242
Budget: Analysis Sample Accounting Codes	{11,087,235}	(1,010,792)	(1,056,830)	(984,993)	[0.867]
		[0.777]	[0.837]	[0.639]	[0.947]
Number of Offices	136	150	148	153	

Notes: The table shows balance of a range of covariates across the treatment arms.

TABLE 3: OVERALL TREATMENT EFFECTS

	(1)	(2)	(3)	(4)
Autonomy	-0.085	-0.086	-0.080	-0.082
	(0.038)	(0.032)	(0.031)	(0.034)
	[0.045]	[0.018]	[0.020]	[0.024]
Incentives	-0.016	-0.026	-0.022	-0.020
	(0.038)	(0.030)	(0.033)	(0.034)
	[0.705]	[0.453]	[0.540]	[0.594]
Both	-0.070	-0.083	-0.072	-0.086
	(0.041)	(0.032)	(0.033)	(0.039)
	[0.141]	[0.030]	[0.059]	[0.057]
Item Type Control	None	Attribs	Scalar	Coarse
p(Incentives $\geq$0)	0.705	0.453	0.540	0.594
p(Autonomy $\geq$0)	0.045	0.018	0.020	0.024
p(Both $\geq$0)	0.141	0.030	0.059	0.057
Observations	11,771	11,771	11,771	11,771

Notes: The table shows the overall treatment effects of the three treatments.

TABLE 4: DIFFERENCE IN DIFFERENCES TREATMENT EFFECTS

	(1)	(2)	(3)	(4)
Autonomy × Year 2	-0.128	-0.130	-0.122	-0.132
	(0.047)	(0.043)	(0.043)	(0.045)
	[0.007]	[0.003]	[0.006]	[0.005]
Incentives × Year 2	-0.051	-0.064	-0.057	-0.057
	(0.047)	(0.040)	(0.041)	(0.043)
	[0.305]	[0.117]	[0.191]	[0.208]
Both × Year 2	-0.098	-0.117	-0.112	-0.102
	(0.050)	(0.042)	(0.043)	(0.045)
	[0.033]	[0.002]	[0.007]	[0.021]
Item Type Control	None	Attribs	Scalar	Coarse
p(Incentives $\geq$0)	0.305	0.117	0.191	0.208
p(Autonomy $\geq$0)	0.007	0.003	0.006	0.005
p(Both $\geq$0)	0.033	0.002	0.007	0.021
Observations	21,183	21,183	21,183	21,183

Notes: The table shows the overall treatment effects of the three treatments.

Table 5: Treatment effects on Demand for Goods

Item	Treatment Effect		
	Incentives	Autonomy	Both
Broom	90.1	52.6	-2.6
	(73.31)	(61.97)	(69.01)
Calculator	-2.2	-10.6	-49.7
	(35.39)	(41.03)	(63.40)
Coal	69.3	-12.5	29.2
	(64.82)	(63.13)	(92.82)
Duster	34.4	-6.8	-46.1
	(38.75)	(44.16)	(65.46)
Envelope	18.2	24.3	-33.3
	(39.41)	(43.89)	(65.91)
File Cover	-13.9	32.1	-23.7
	(43.71)	(53.65)	(70.20)
Floor Cleaner	22.1	10.9	23.9
	(49.47)	(51.13)	(53.46)
Ice Block	-24.8	6.8	-46.4
	(44.69)	(50.37)	(69.77)
Light Bulb	-10.1	77.6	10.9
	(60.49)	(102.21)	(89.13)
Lock	26.3	12.8	-44.5
	(43.25)	(45.12)	(60.96)
Newspaper	0.1	13.4	-52.3
	(43.44)	(47.84)	(68.33)
Pen	35.6	-14.5	-57.7
	(52.19)	(43.76)	(64.08)
Pencil	12.0	7.7	-39.8
	(38.55)	(43.54)	(64.95)
Photocopying	69.5	-18.2	30.4
	(79.27)	(72.76)	(106.00)
Pipe	116.4**	51.3	-3.5
	(54.21)	(56.53)	(69.95)
Printer Paper	372.1**	178.8	94.6
	(170.21)	(206.15)	(167.04)
Register	-471.0	-488.0	-56.5
	(453.84)	(453.99)	(637.88)
Sign Board/Banner	-206.2	-133.0	-252.9
	(236.33)	(239.77)	(241.09)
Soap/Detergent	48.4	61.6	1025.2
	(40.93)	(56.66)	(985.36)
Stamp Pad	19.8	8.9	-36.6
	(37.67)	(42.94)	(65.04)
Stapler	4.9	-6.9	-47.0
	(39.61)	(44.30)	(65.66)
Staples	13.0	2.0	-34.0
	(38.55)	(43.55)	(65.06)
Toner	350.0	116.3	188.0
	(343.82)	(118.08)	(239.41)
Towel	19.4	-2.1	-46.6
	(40.51)	(44.80)	(65.59)
Wiper	34.5	2.9	-42.6
	(40.26)	(44.07)	(65.14)

Notes: The table shows the overall treatment effects of the three treatments on the demand for different goods.

Table 6: Treatment Effect Heterogeneity by District

	(1)	(2)	(3)	(4)
Incentives	-0.121	-0.126**	-0.143**	-0.098
	(0.078)	(0.062)	(0.065)	(0.068)
Autonomy	-0.019	-0.065	-0.038	-0.013
	(0.083)	(0.067)	(0.067)	(0.074)
Both	0.107	-0.009	0.062	0.083
	(0.078)	(0.066)	(0.065)	(0.077)
Incentives × District June Share	0.279*	0.257*	0.321**	0.204
	(0.169)	(0.147)	(0.145)	(0.149)
Autonomy × District June Share	-0.187	-0.071	-0.124	-0.196
	(0.195)	(0.157)	(0.156)	(0.171)
Both × District June Share	-0.447**	-0.192	-0.339**	-0.430**
	(0.186)	(0.162)	(0.154)	(0.187)
Item Type Control	none	Attribs	Scalar	Coarse
Observations	11666	11666	11666	11666

Notes: The table shows heterogeneity of treatment effects by the degree of misalignment of the district's accountant general. Accountants general are classified according to the degree to which purchase approvals are bunched at the end of the fiscal year in June 2015 (year 1 of the project).

Table 7: Treatment Effect Heterogeneity by Item

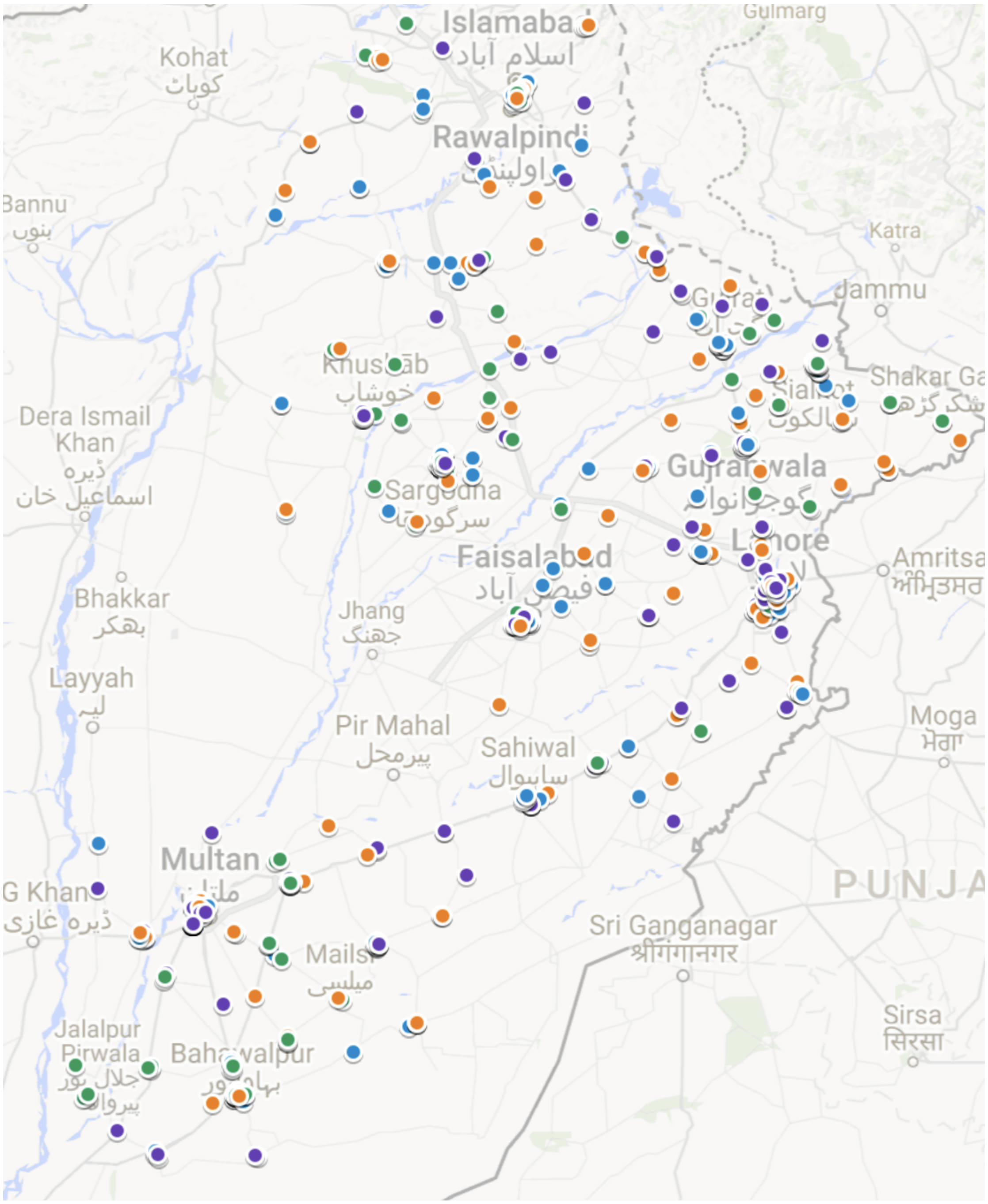
	(1)	(2)	(3)	(4)
Incentives	0.135	0.045	0.046	0.113
	(0.087)	(0.054)	(0.059)	(0.073)
Autonomy	0.075	0.049	0.039	0.073
	(0.084)	(0.060)	(0.063)	(0.071)
Both	0.013	-0.067	-0.074	-0.020
	(0.083)	(0.059)	(0.063)	(0.078)
Incentives × Price Dispersion Ratio	-0.060**	-0.029	-0.028	-0.053**
	(0.027)	(0.020)	(0.020)	(0.025)
Autonomy × Price Dispersion Ratio	-0.062**	-0.054**	-0.047*	-0.061**
	(0.029)	(0.023)	(0.024)	(0.026)
Both × Price Dispersion Ratio	-0.031	-0.007	0.001	-0.025
	(0.028)	(0.021)	(0.022)	(0.027)
Item Type Control	none	Attribs	Scalar	Coarse
Observations	11666	11666	11666	11666

Notes: The table shows heterogeneity of treatment effects by the degree of homogeneity of the good being purchased. Goods g are classified according to the ratio of the total variance in their log unit prices to the variance explained by the item's attributes $\phi_g = \mathsf{Var}(p) / \mathsf{Var}(\hat{p})$, where $\hat{p}$ is the price as predicted using the item's attributes and variances are across purchases of that good in the control group in year 1 of the project.

Web Appendix (Not For Publication)

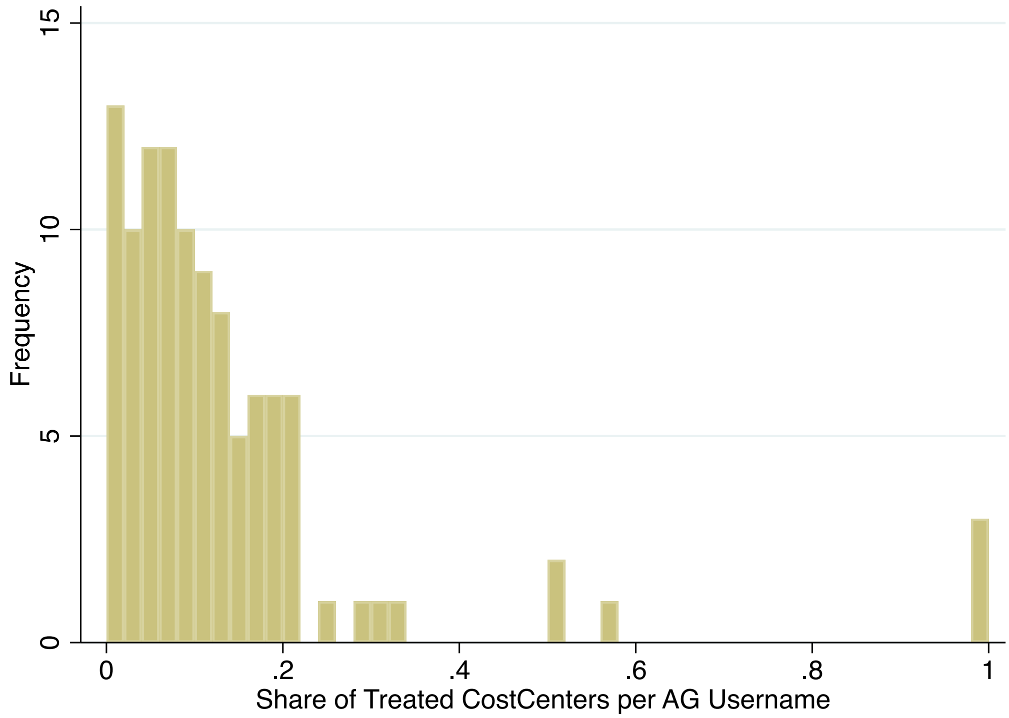
A Supplementary Figures and Tables

FIGURE A.1: LOCATION OF SAMPLE OFFICES



Notes: The figure shows the location of the offices in the study. The offices are located in 26 of the 36 districts in Punjab. Green dots denote control offices, orange dots the autonomy group, blue dots the performance pay group, and purple dots the combined treatment.

FIGURE A.2: SAMPLE OFFICES ARE A SMALL SHARE OF THE OFFICES OVERSEEN BY USERS AT THE ACCOUNTANT GENERAL'S OFFICE



*Notes:*Each transaction approved by the accountant general's office is associated with a particular officer's username. The figure shows the share of cost centers associated with each username that are in the treated groups of our experiment. The figure shows that for the vast majority of users at the accountant general's office, fewer than 20% of their offices are treated.

TABLE A.1: UNIVERSE OF GENERIC GOODS ACCOUNTING CODES

Code	Category	Description
		Panel A: A03 Operating Expenses
A03004	Other	Furnace Oil - Non Operational
A03070		Others
A03170	Fees	Others
A03204	Communication	Electronic Communication
A03205		Courier And Pilot Service
A03206		Photography Charges
A03270		Others
A03304	Utilities	Hot And Cold Weather
A03305		POL For Generator
A03370		Others
A03401	Occupancy Costs	Charges
A03405		Rent Other Than Building
A03408		Rent Of Machine & Equipment
A03410		Security
A03470		Others
A03501	Operating Leases	Machinery And Equipment
A03502		Buildings
A03503		Motor Vehicles
A03504		Computers
A03506		Medical Machinery And Technical Equipment
A03570		Others
A03901	General	Stationery
A03902		Printing And Publication
A03904		Hire Of Vehicles
A03905		Newspapers Periodicals And Books
A03907		Advertising & Publicity
A03919		Payments To Others For Service Rendered
A03921		Unforeseen Exp. For Disaster Preparedness
A03927		Purchase Of Drug And Medicines
A03933		Service Charges
A03940		Unforeseen Expenditure
A03942		Cost Of Other Stores
A03955		Computer Stationary
A03970		Others
A03971		Cost Of State Trading Medicines
A03972		Expenditure On Diet For Patient

Continued on next page

Table A.1 – *Continued from previous page*

Code	Category	Description
A03978		Free Text Books
	Panel B: A09 Physical Assets	
A09105	Purchase of Physical Assets	Transport
A09107		Furniture And Fixtures
A09108		Livestock
A09170		Others
A09204	Computer Accessories	License Fee For Software
A09302	Commodity Purchases	Fertilizer
A09303		Coal
A09370		Others
A09401	Other Stores and Stock	Medical Stores
A09402		Newsprint
A09403		Tractors
A09404		Medical And Laboratory Equipment
A09405		Workshop Equipment
A09406		Storage And Carrying Receptacles
A09407		Specific Consumables
A09408		Generic Consumables
A09409		Medical Stocks
A09410		Life Saving Medical Supplies
A09411		General Utility Chemicals
A09412		Specific Utility Chemicals
A09413		Drapery Fabrics Clothing And Allied Materials
A09414		Insecticides
A09470		Others
A09501	Transport	Transport
A09502		Diplomatic Cars
A09503		Others
A09601	Plant & Machinery	Plant And Machinery
A09602		Cold Storage Equipment
A09603		Signalling System
A09604		Railways Rolling Stock
A09701	Furniture & Fixtures	Furniture And Fixtures
A09702		Unkempt Furnishings
A09801	Livestock	Livestock
A09802		Purchase Of Other Assets - Others
A09803		Meters & Services Cables

Continued on next page

Table A.1 – *Continued from previous page*

Code	Category	Description
A09899		Others
	Panel C: A13 Repairs and Maintenance	
A13101	Machinery & Equipment	Machinery And Equipment
A13199		Others
A13201	Furniture & Fixture	Furniture And Fixture
A13370	Buildings & Structure	Others
A13470	Irrigation	Others
A13570	Embankment & Drainage	Others
A13701	Computer Equipment	Hardware
A13702		Software
A13703		I.T. Equipment
A13920	Telecommunication	Others