GR 6307
Public Economics and Development

4. The Personnel Economics of the Developing State: Delivering Services to the Poor

Michael Carlos Best

Spring 2022

Outline

Theory

Financial Incentives

Non-financial Incentives

Recruitment & Selection

Open Questions

Outline

Theory

- Aghion & Tirole (JPE 1997) *Formal and Real Authority in Organizations*
- Benabou & Tirole (AER 2006) *Incentives and Prosocial Behavior*
- Besley & Ghatak (AER 2005) *Competition and Incentives with Prosocial Agents*

Aghion & Tirole (1997): Model Setup

- Principal-agent framework: Agent is choosing among $n \geq 3$ a priori identical projects.
- Project k has profit B_k for the principal and private benefit b_k for the agent.
- They can also do nothing: $B_0 = b_0 = 0$
- Congruence:
  - Choosing the principal's preferred project gives her B and the agent βb.
  - Choosing the agent's preferred project gives him b and the principal αB.
  - $0 < \alpha, \beta \leq 1$ are exogenous parameters

Aghion & Tirole (1997): Model Setup

- Principal is risk neutral. Utility is
  $$B_k - w$$
  w is wage paid to the agent
- Agent is risk averse and has limited liability: $w \geq 0$. Utility is
  $$u\left(w\right) + b_k$$
  Agent is so risk averse that w can't depend on outcomes
- Initially, nobody knows projects' payoffs. Gathering information is costly.
- If agent pays cost $g_A\left(e\right)$ he learns the payoffs of all projects with probability e. With probability $1 - e$ he learns nothing.
- Principal can pay cost $g_P\left(E\right)$ to learn payoffs with probability E. With probability $1 - E$ she learns nothing.

Aghion & Tirole (1997): Authority

1. *P-formal authority:* The principal has formal authority. She may overrule the agent's recommendation.
2. *A-formal authority:* The agent picks his preferred project and cannot be overruled by the principal.

- Contracts specify an allocation of formal authority to either the principal or the agent.
- *Real authority:* Who actually gets to make the decision? Either because agent has formal authority or because P is just "rubber-stamping" agent's recommendation
- Timing:
    1. Prinicpal proposes a contract
    2. Parties gather information
    3. The party without formal authority communicates a subset of the projects' payoffs (s)he has learned
    4. The controlling party picks a project

Aghion & Tirole (1997): Utilities

- Under *P*-formal authority, the utilities are:

$$u_P = \underbrace{EB}_{\text{P picks her preferred project}} + \underbrace{(1-E)\,e\alpha B}_{\text{A suggests his preferred project}} - g_P(E)$$

$$u_A = \underbrace{E\beta b}_{\text{P picks her preferred project}} + \underbrace{(1-E)\,eb}_{\text{A suggests his preferred project}} - g_A(e)$$

- Under *A*-formal authority, the utilities are:

$$u_P^d = \underbrace{e\alpha B}_{\text{A picks his preferred project}} + \underbrace{(1-e)\,EB}_{\text{P suggests her preferred project}} - g_P(E)$$

$$u_A^d = \underbrace{eb}_{\text{A picks his preferred project}} + \underbrace{(1-e)\,E\beta b}_{\text{P suggests her preferred project}} - g_A(e)$$

Aghion & Tirole (1997): Basic Tradeoff

- In this model there is a basic tradeoff between loss of control and initiative.
- The reason is that efforts are *strategic substitutes:* The more effort the principal makes, the less the agent wants to (&vv).
- To see this, the FOCs for effort when the principal has formal authority are

$$(1-\alpha e)\,B = g_P'\,(E)$$
$$(1-E)\,b = g_A'\,(e)$$

- Both of these reaction curves slope *down*.
- Imagine the principal's effort became more costly: $g_P' \uparrow$
  - Probability of learning the best project goes down. The principal loses real authority (control)
  - The reduction in E will encourage initiative by the agent:, $e \uparrow$. The principal gains

Aghion & Tirole (1997): Delegation

- If the principal cedes formal authority to the agent the effort FOCs become

$$(1-e)\,B = g_P'\,(E)$$
$$(1-\beta E)\,b = g_A'\,(e)$$

- These yield an equilibrium $\left(E^d, e^d\right)$ where
  - $e < e^d$: Greater initiative by the agent
  - $E > E^d$: Loss of formal *and* real authority to the agent.
  - Less effort required from principal
  - Agent is better off $\rightarrow$ slackens participation constraint so could lower wage

Aghion & Tirole (1997): Span of Control

- Consider a principal with multiple agents where the principal doesn't want to delegate.
- How many agents to hire? How to encourage effort among many agents?
- m identical agents. Each one solving the problem above.
- Principal's disutility is $g_P\left(\sum_i E_i\right)$, agents' tasks are independent. Fixed cost f per agent.

$$u_P = \sum_i \left[E_i B + (1 - E_i)\, e_i \alpha B - f\right] - g_P\left(\sum_i E_i\right)$$

Aghion & Tirole (1997): Span of Control

- Assume a symmetric equilibrium, each agent gets the same effort E from the principal. FOCs are

$$(1-\alpha e)\,B = g_P'\,(mE)$$
$$(1-E)\,b = g_A'\,(e)$$

with solution $\{E\,(m)\,,e\,(m)\}$.

- Principal's utility from m agents is

$$u_P\,(m) \equiv mR\,(E\,(m)\,,e\,(m)) - g_P\,(mE\,(m))$$

where $R\,(E\,(m)\,,e\,(m)) \equiv E\,(m)\,B + [1-E\,(m)]\,e\,(m)\,\alpha B - f$ is revenue per agent.

Aghion & Tirole (1997): Span of Control

- The optimal team size m then satisfies

$$\frac{du_P}{dm} = \underbrace{\underbrace{R\left(E\left(m\right), e\left(m\right)\right)}_{\text{extra revenue}} - \underbrace{E\left(m\right) g_P'\left(mE\left(m\right)\right)}_{\text{overload cost}}}_{\text{Marginal profit}<0}$$

$$+ \quad \underbrace{m\frac{\partial R}{\partial e}\frac{\partial e}{\partial m}}_{\text{initiative effect }>0} \quad = 0$$

- Principal commits to overhiring, being overloaded and underinvesting in E in order to encourage initiative e

Aghion & Tirole (1997): Wages and Effort

- Now reintroduce wage effects in the model where the principal has formal authority.
- How do changes in wages affect real authority?
- Suppose that two of the projects are relevant and give the principal profits of B and 0. This implies $\alpha = \beta =$probability they have the same preferred project.
- The agent gets a wage $w \geq 0$ when the principal's profit is B
- Principal's net gain is now $B - w$
- If the agent has information and real authority, his average net payoff is

$$\tilde{b} = \begin{cases} \underbrace{b}_{\text{choose preferred proj}} + \underbrace{\alpha u(w)}_{\text{w/pr } \alpha \text{, congruence}} & \text{if } u(w) < b \\ \underbrace{u(w)}_{\text{choose principal's preferred proj}} + \underbrace{\alpha b}_{\text{w/pr } \alpha \text{, congruence}} & \text{if } u(w) \geq b \end{cases}$$

Aghion & Tirole (1997): Wages and Effort

- Now the FOCs are

$$(1-\alpha e)\,\tilde{B} = g_P'(E)$$
$$(1-E)\,\tilde{b} = g_A'(e)$$

- Denote solution to this as $\{E(w), e(w)\}$. Then by backward induction solve for w

$$\frac{du_P}{dw} = \underbrace{(1-E)\,\alpha\,(B-w)\,\frac{de}{dw}}_{\text{additional effort}} - \underbrace{[E+(1-E)\,e\alpha]}_{\text{higher wage bill}}$$

- Higher wages increase real authority:
  1. Stronger incentives $\rightarrow$ agent more likely to make a recommendation
  2. Principal monitors less $\rightarrow$ less likely to overrule the agent

Outline

Theory

- Aghion & Tirole (JPE 1997) *Formal and Real Authority in Organizations*
- Benabou & Tirole (AER 2006) *Incentives and Prosocial Behavior*
- Besley & Ghatak (AER 2005) *Competition and Incentives with Prosocial Agents*

Benabou & Tirole 2006: Introduction

- People often do things that are costly to themselves and primarily benefit others. Why?

1. Rewards and punishments for prosocial behavior sometimes backfire.
2. Social pressure and norms successfully use honor and shame to direct behavior
3. People care about their *self-image*. People want to think they are prosocial.

- Develop a theory of prosocial behavior.
  - Heterogeneity in degree of altruism/greed
  - desire for social reputation/self-respect
- People's behavior has 3 motivations *intrinsic, extrinsic,* and *reputational.*

Benabou & Tirole 2006: Model

- Agents are choosing how much to participate in a pro-social activity.
- Choose a from choice set $A \subset \mathbb{R}$ incurring cost $C(a)$
- Monetary reward is ya, $y \lesseqgtr 0$
- Agents' types are
  - v_a: intrinsic valuation
  - v_y: extrinsic valuation
  - $\mathbf{v} \equiv (v_a, v_y) \in \mathbb{R}^2$. continuous density $f(\mathbf{v})$ and mean $(\bar{v}_a, \bar{v}_y)$
- Direct benefit of participating is

$$(v_a + v_y y)\, a - C(a)$$

Benabou & Tirole 2006: Model

- Participation decisions also create reputational costs/benefits.
- Assume these depend linearly on observers' posterior expectations of the agent's type $\mathbf{v}$

$$R(a,y) \equiv x\left(\gamma_a \mathbb{E}\left[v_a|a,y\right] - \gamma_y \mathbb{E}\left[v_y|a,y\right]\right),\ \gamma_a \geq 0,\ \gamma_y \geq 0$$

- $\Rightarrow$ people want to be seen as *prosocial* $\gamma_a \geq 0$ and *disinterested* $\gamma_y \geq 0$
- $x > 0$ measures the visibility/salience of actions. Defining $\mu_a = x\gamma_a$ and $\mu_y = x\gamma_y$, agents solve

$$\max_{a \in A} \left(v_a + v_y y\right) a - C(a) + \mu_a \mathbb{E}\left[v_a|a,y\right] + \mu_y \mathbb{E}\left[v_y|a,y\right]$$

Benabou & Tirole 2006: Choice of a

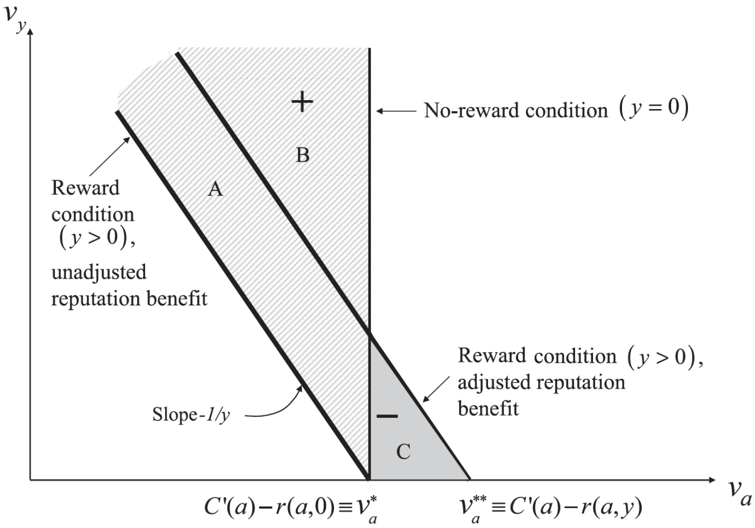
- The agent's optimal choice satisfies the FOC

$$C'(a) = v_a + v_y y + r(a, y; \boldsymbol{\mu})$$
$$r(a, y; \boldsymbol{\mu}) \equiv \mu_a \frac{\partial \mathbb{E}[v_a | a, y]}{\partial a} - \mu_y \frac{\partial \mathbb{E}[v_y | a, y]}{\partial a}$$

1. Observing a reveals the *sum* of intrinsic, extrinsic & reputational concerns $\rightarrow$ signal extraction problem
2. A higher incentive y makes a more informative about v_y but less about v_a
3. $\boldsymbol{\mu}$ makes inference about v_a and v_y noisier. This gets worse when actions are more visible (higher x)

Benabou & Tirole 2006: Analysis

- Start with the case where μ_a and μ_y are fixed.



Benabou & Tirole 2006: Analysis

- Add a few assumptions: $A = \mathbb{R}$, $C(a) = ka^2/2$,

$$\mathbf{v} \equiv \begin{pmatrix} v_a \\ v_y \end{pmatrix} \sim \mathcal{N}\left(\begin{matrix} \bar{v}_a \\ \bar{v}_y \end{matrix}, \begin{matrix} \sigma_a^2 & \sigma_{ay} \\ \sigma_{ay} & \sigma_y^2 \end{matrix} \right), \quad \bar{v}_a \lesseqgtr 0,\ \bar{v}_y > 0$$

- Start with case where $\boldsymbol{\mu}$ is fixed. Implies that

$$\bar{r}(a, y) \equiv \bar{\mu}_a \frac{\partial \mathbb{E}[v_a|a, y]}{\partial a} - \bar{\mu}_y \frac{\partial \mathbb{E}[v_y|a, y]}{\partial a}$$

- With normal $\mathbf{v}$, the posteriors are

$$\mathbb{E}[v_a|a, y] = \bar{v}_a + \rho(y)[ka - \bar{v}_a - \bar{v}_y y - \bar{r}(a, y)]$$
$$\mathbb{E}[v_y|a, y] = \bar{v}_y + \chi(y)[ka - \bar{v}_a - \bar{v}_y y - \bar{r}(a, y)]$$

  where $\rho(y) = \frac{\sigma_a^2 + y\sigma_{ay}}{\sigma_a^2 + 2y\sigma_{ay} + y^2\sigma_y^2}$ and $y\chi(y) \equiv 1 - \rho(y)$

- Equilibrium solves these two differential equations.

Benabou & Tirole 2006: Signal Extraction

PROPOSITION 1: *Let all agents have the same image concern* $(\bar{\mu}_a, \bar{\mu}_y)$. *There is a unique (differentiable-reputation) equilibrium, in which an agent with preferences* (v_a, v_y) *contributes at the level*

$$a = \frac{v_a + v_y y}{k} + \bar{\mu}_a \rho(y) - \bar{\mu}_y \chi(y)$$

The reputational returns are $\partial \mathbb{E}[v_a | a, y] / \partial a = \rho(y) k$ *and* $\partial \mathbb{E}[v_y | a, y] / \partial a = \chi(y) k$, *resulting in a net value* $\bar{r}(y) = k(\bar{\mu}_a \rho(y) - \bar{\mu}_y \chi(y))$, *independent of* a.

- How do extrinsic incentives affect inference and behavior? higher y increases direct payoff, but decreases both dimensions of signaling. e.g. when $\sigma_{ay} = 0$

$$\rho(y) = \frac{1}{1 + y^2 \sigma_y^2 / \sigma_a^2} \quad \chi(y) = \frac{y \sigma_y^2 / \sigma_a^2}{1 + y^2 \sigma_y^2 / \sigma_a^2}$$

- $\Rightarrow$ Higher y is like increasing the noise to signal ratio σ_y / σ_a
- When $\sigma_{ay} \neq 0$, a positive correlation amplifies this.

Benabou & Tirole 2006: Crowd-out

- Aggregate supply of the public good $\bar{a}(y) = \int_i a_i di$ has slope

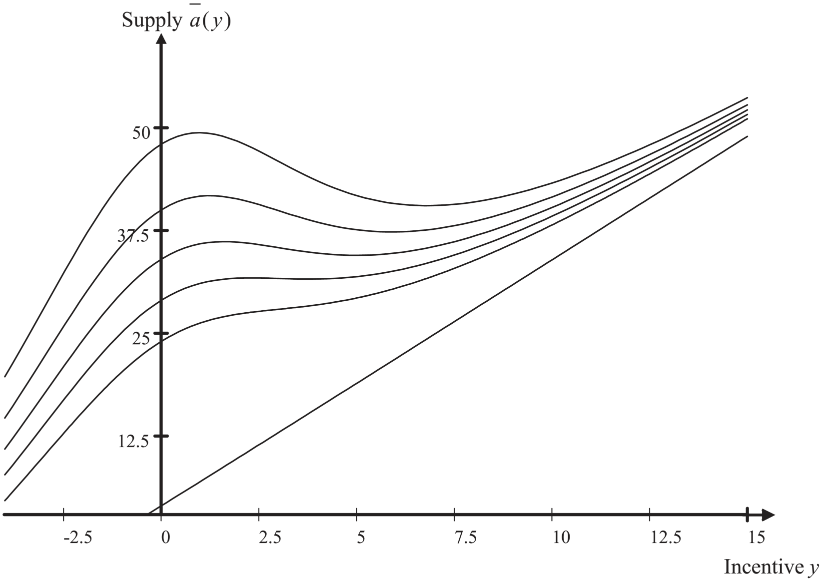
$$\bar{a}'(y) = \frac{\bar{v}_y}{k} + \bar{\mu}_a \rho'(y) - \bar{\mu}_x \chi'(y)$$

PROPOSITION 2 (Overjustification and crowding out): *Let $\sigma_{ay} = 0$ and define $\theta \equiv \sigma_y / \sigma_a$. Incentives are counterproductive, $\bar{a}'(y) < 0$, at all levels such that*

$$\frac{\bar{v}_y}{k} < \bar{\mu}_a \frac{2y\theta^2}{(1 + y^2\theta^2)^2} + \bar{\mu}_y \frac{\theta^2 \left(1 - y^2\theta^2\right)}{(1 + y^2\theta^2)^2}$$

Consequently, for all $\bar{\mu}_a$ above some threshold $\mu_a^* \geq 0$, there exists a range $[y_1, y_2]$ such that $\bar{a}(y)$ is decreasing on $[y_1, y_2]$ and increasing everywhere else on $\mathbb{R}$. If $\bar{\mu}_y < \bar{v}_y / k\theta^2$, then $\mu_a^* > 0$ and $0 < y_1 < y_2$; as $\bar{\mu}_a$ increases, y_1 falls and y_2 rises, so $[y_1, y_2]$ widens. If $\bar{\mu}_y > \bar{v}_y / k\theta^2$, then $\mu_a^* = 0$ and $y_1 < 0 < y_2$; as $\bar{\mu}_a$ increases both y_1 and y_2 rise and, for $\bar{\mu}_a$ large enough, $[y_1, y_2]$ again widens.

Benabou & Tirole 2006: Crowd-out



Benabou & Tirole 2006: Image Rewards

- We have studied how extrinsic incentives (y) affect participation. Can providing visibility to contributions (x) do a better job of encouraging participation?
- Yes, but: When we have a homothetic increase in μ_a, μ_y this works, but with heterogeneity people may suspect that contributors are just doing it to look good: That they are *image-motivated*. This dampens incentives to participate.
- Allow image concerns also to be heterogeneous:

$$\begin{pmatrix} \mu_a \\ \mu_y \end{pmatrix} \sim \mathcal{N}\left(\begin{array}{c} \bar{\mu}_a \\ \bar{\mu}_y \end{array}, \begin{bmatrix} \omega_a^2 & \omega_{ay} \\ \omega_{ay} & \omega_y^2 \end{bmatrix} \right), \ \bar{\mu}_a, \bar{\mu}_y \geq 0$$

and $\mathbf{v}$ and $\boldsymbol{\mu}$ are independent.

Benabou & Tirole 2006: Image Rewards

- The first order condition for the choice of a is still

$$C'(a) = v_a + v_y y + r(a, y; \boldsymbol{\mu})$$

- Now the reputational concern term in the first order condition $r(a, y; \boldsymbol{\mu})$ is also normally distributed, with mean $\bar{r}(a, y; \boldsymbol{\mu})$ and variance

$$\Omega(a, y)^2 \equiv \left(\frac{\partial \mathbb{E}[v_a | a, y]}{\partial a} - \frac{\partial \mathbb{E}[v_y | a, y]}{\partial a} \right) \times \begin{pmatrix} \omega_a^2 & \omega_{ay} \\ \omega_{ay} & \omega_y^2 \end{pmatrix} \times \begin{pmatrix} \frac{\partial \mathbb{E}[v_a | a, y]}{\partial a} \\ -\frac{\partial \mathbb{E}[v_y | a, y]}{\partial a} \end{pmatrix}$$

Benabou & Tirole 2006: Image Rewards

- Updating still satisfies

$$\mathbb{E}\left[v_a|a,y\right] = \bar{v}_a + \rho\left(a,y\right)\left[ka - \bar{v}_a - \bar{v}_y y - \bar{r}\left(a,y\right)\right]$$
$$\mathbb{E}\left[v_y|a,y\right] = \bar{v}_y + \chi\left(a,y\right)\left[ka - \bar{v}_a - \bar{v}_y y - \bar{r}\left(a,y\right)\right]$$

  but now

$$\rho\left(a,y\right) \equiv \frac{\sigma^2 + y\sigma_{ay}}{\sigma_a^2 + 2y\sigma_{ay} + y^2\sigma_y^2 + \Omega\left(a,y\right)^2}$$
$$\chi\left(a,y\right) \equiv \frac{y\sigma^2 + \sigma_{ay}}{\sigma_a^2 + 2y\sigma_{ay} + y^2\sigma_y^2 + \Omega\left(a,y\right)^2}$$

- Equilibrium solves these differential equations.
  - But note they are now nonlinear because of the Ω^2 term.
  - Restrict attention to equilibria in the class where $\Omega \perp a$. This keeps reputations linear in a

Benabou & Tirole 2006: Image Rewards

PROPOSITION 4: (1) *A linear-reputation equilibrium corresponds to a fixed point* $\Omega(y)$, *solution to*

$$\frac{\Omega(y)^2}{k^2} = \omega_a^2 \rho(y)^2 - 2\omega_{ay}\rho(y)\chi(y) + \omega_y^2 \chi(y)^2$$

The optimal action chosen by an agent with type $(\mathbf{v}, \boldsymbol{\mu})$ *is then*

$$a = \frac{v_a + v_y y}{k} + \mu_a \rho(y) - \mu_y \chi(y)$$

and the marginal reputations are $\partial\mathbb{E}[v_a|a,y]/\partial a = \rho(y)k$ *and* $\partial\mathbb{E}[v_y|a,y]/\partial a = \chi(y)k$, *with a net value of* $r(y;\boldsymbol{\mu}) = [\mu_a\rho(y) - \mu_y\chi(y)]k$ *for the agent.*
(2) There always exists such an equilibrium, and if $\omega_{ay} = 0$ *it is unique (in the linear reputation class)*

- Fixed point intuition:
  - The more variable image motives are, the noisier behavior is as a signal of v_a, v_y, reducing $\rho(y)$ and $\chi(y)$.
  - But the variance is endogenous to behavior which takes into account its effect on signal-extraction.

Benabou & Tirole 2006: Image Rewards

- Image rewards give rise to an offsetting *overjustification effect*. To see this, consider scaling all the reputational weights $\boldsymbol{\mu} = (\mu_a, \mu_y)$ up by a prominence factor x holding the material incentive y constant.
- Aggregate supply is

$$\bar{a}(y, x) = \frac{\bar{v}_a + \bar{v}_y y}{k} + x\left[\bar{\mu}_a \rho(y, x) - \bar{\mu}\chi(y, x)\right]$$

- Increasing x has 2 effects:

1. Direct *amplifying* effect with sign $sign(\mu_a \rho(y, x) - \mu_y \chi(y, x))$
   1.1 For socially minded people with $\mu_a \gg \mu_y$ this increases incentives to contribute
   1.2 For people worried not to look greedy $\mu_a \ll \mu_y$ this decreases incentives.
2. Indirect *dampening* effect. Increasing x increases the noise $\Omega \to$ people attribute behavior more to image-seeking $\rho(y, x)$ and $\chi(y, x)$ shrink $\to$ people respond less to image rewards.

Outline

Theory

- Aghion & Tirole (JPE 1997) *Formal and Real Authority in Organizations*
- Benabou & Tirole (AER 2006) *Incentives and Prosocial Behavior*
- Besley & Ghatak (AER 2005) *Competition and Incentives with Prosocial Agents*

Besley & Ghatak 2005: Introduction

- Money is not the only way that workers are motivated
- Many organizations, especially in the non-profit & public sectors have a "mission"
- (some) workers too care about the mission of the organization they work with.
- Build a model to study this.
  - Matching on mission $\rightarrow$ less need for explicit incentives
  - But, entrenches conservatism/resistance to innovation.

Besley & Ghatak 2005: Principal-Agent Setup

- A firm = a risk-neutral principal, and a risk-neutral agent.
- Principal needs agent to do a project.
- Project outcome is high $\rightarrow Y_H$ or low $\rightarrow Y_L < Y_H$
- Probability of high outcome is effort by agent e.
- Effort is non-contractible and costs agent $e^2/2$
- Agent has limited liability so requires wage $\underline{w} \geq 0$ every period.

Besley & Ghatak 2005: Organizational Mission

- 3 types of principals $i \in \{0, 1, 2\}$
- If project succeeds, principal gets $\pi_i > 0$.
- Type 0 principals are "standard": π_0 is purely monetary. Think of them as the private sector, the "**Profit-oriented sector**"
- Types 1 and 2: Part of π_1, π_2 are nonpecuniary payoffs: Think of them as non-profits/govt, the "**Mission-oriented sector**"
- Assume $\pi_1 = \pi_2 = \hat{\pi} \rightarrow$ this is a model of horizontal matching: no productivity differences across orgs when there is efficient matching.

Besley & Ghatak 2005: Intrinsic Motivation

- 3 types of agents $j \in \{0, 1, 2\}$
- Agents get a nonpecuniary benefit θ_{ij} from working at a type i organization
- Type 0s don't care: $\theta_{i0} = 0$,
- Types 1 and 2 are "Motivated Agents": Get $\bar{\theta}$ from working at "their" type, $\underline{\theta}$ from working at the other type. $\bar{\theta} > \underline{\theta} \geq 0$

$$\theta_{ij} = \begin{cases} 0 & \text{if } i = 0 \text{ and/or } j = 0 \\ \underline{\theta} & \text{if } i \in \{1,2\}, j \in \{1,2\}, i \neq j \\ \bar{\theta} & \text{if } i \in \{1,2\}, j \in \{1,2\}, i = j \end{cases}$$

- Assume: $\max\left\{\pi_0, \hat{\pi} + \bar{\theta}\right\} < 1$ to guarantee interior solutions for effort in all matches

Besley & Ghatak 2005: Optimal Contracts

- Contracts have 2 terms
  1. A fixed wage w_{ij} paid regardless of the project outcome
  2. A bonus b_{ij} if the outcome is Y_H
- Consider the first-best as a benchmark. Effort is contractible and solves

$$\max_e e\left[\pi_i + \theta_{ij}\right] + (1 - e)\left[0\right] - e^2/2$$

- First-best optimal effort:

$$e = \pi_i + \theta_{ij}$$

- Generates total surplus

$$\frac{\left(\pi_i + \theta_{ij}\right)^2}{2}$$

Besley & Ghatak 2005: Optimal Contracts

- In the second best, effort is not contractible. Principal solves

$$\max_{[b_{ij},w_{ij}]} u_{ij}^{P} = (\pi_i - b_{ij})\, e_{ij} - w_{ij}$$

- Subject to 3 constraints:
  - limited liability: Agent gets at least $\underline{w}$:

$$b_{ij} + w_{ij} \geq \underline{w} \quad w_{ij} \geq \underline{w}$$

  - participation: Agent prefers this to outside option

$$u_{ij}^{a} = e_{ij}\,(b_{ij} + \theta_{ij}) + w_{ij} - \frac{1}{2} e_{ij}^{2} \geq \bar{u}_j$$

  - Incentive compatibility: Agent picks e_{ij}

$$e_{ij} \in \arg \max_{e_{ij} \in [0,1]} \left\{ e_{ij}\,(b_{ij} + \theta_{ij}) + w_{ij} - \frac{1}{2} e_{ij}^{2} \right\}$$

    which simplifies to $e_{ij} = b_{ij} + \theta_{ij}$ as long as this is $\in [0,1]$

Besley & Ghatak 2005: Optimal Contracts

- Assume the project is always worth trying:

$$\frac{1}{4}\left[\min\left\{\pi_0, \hat{\pi}\right\}\right]^2 - \underline{w} > 0$$

- Define $\bar{v}_{ij}$ as the value of the reservation payoff to an agent of type j such that a principal of type i makes zero expected profits under the optimal contract. And define $\underline{v}_{ij}$ as the lowest $\bar{u}_j$ for which the participation constraint binds.

Besley & Ghatak 2005: Optimal Contracts

PROPOSITION 1: *Suppose Assumptions 1 and 2 hold. An optimal contract* $\left(b_{ij}^{*}, w_{ij}^{*}\right)$ *between a principal of type* i *and an agent of type* j *given a reservation payoff* $\bar{u}_j \in [0, \bar{v}_{ij}]$ *exists, and has the following features:*

1. *The fixed wage is set at the subsistence level:* $w_{ij}^{*} = \underline{w}$
2. *The bonus payment is characterized by*

$$b_{ij}^{*} = \begin{cases} \max\left\{0, \frac{\pi_i - \theta_{ij}}{2}\right\} & \textit{if } \bar{u}_j \in [0, \underline{v}_{ij}] \\ \sqrt{2\left(\bar{u}_j - \underline{w}\right)} - \theta_{ij} & \textit{if } \bar{u}_j \in [\underline{v}_{ij}, \bar{v}_{ij}] \end{cases}$$

3. *The optimal effort level solves:* $e_{ij}^{*} = b_{ij}^{*} + \theta_{ij}$

Besley & Ghatak 2005: Optimal Contracts

- Gives rise to 3 cases

1. If the agent is more motivated than the principal and the outside option is low, $b_{ij}^{*}=0$
2. If the principal is more motivated than the agent and the outside option is low, $b_{ij}^{*}=\frac{1}{2}\left(\pi_{i}-\theta_{ij}\right)$
3. If the outside option is high, then $b_{ij}^{*}=\sqrt{2\left(\bar{u}_{ij}-\underline{w}\right)}-\theta_{ij}$

Besley & Ghatak 2005: Optimal Contracts in the Profit-Oriented Sector

COROLLARY 1: *In the profit-oriented sector* $(i = 0)$, *the optimal contract is characterized by the following:*

(a) The fixed wage is set at the subsistence level, i.e., $w_{0j}^* = \underline{w}$ $(j = 0, 1, 2)$

(b) The bonus payment is characterized by

$$b_{0j}^* = \begin{cases} \frac{\pi_0}{2} & \text{if } \bar{u}_j \in \left[0, \underline{v}_{0j}\right] \\ \sqrt{2\left(\bar{u}_j - \underline{w}\right)} & \text{if } \bar{u}_j \in \left[\underline{v}_{0j}, \bar{v}_{0j}\right] \end{cases}$$

for $j = 0, 1, 2$

(c) The optimal effort level solves: $e_{0j}^* = b_{0j}^*$ $(j = 0, 1, 2)$

Besley & Ghatak 2005: Optimal Contracts in the Mission-oriented sector

COROLLARY 2: *Suppose that* $\bar{u}_0 = \bar{u}_1 = \bar{u}_2$. *Then, in the mission-oriented sector* $(i = 1, 2)$, *effort is higher and the bonus payment is lower if the agent's type is the same as that of the principal.*

- bonuses and intrinsic motivation are perfect substitutes

COROLLARY 3: *Suppose that* $\bar{u}_0 = \bar{u}_1 = \bar{u}_2$. *Then, in the mission-oriented sector* $(i = 1, 2)$ *bonus payments and effort are negatively correlated in a cross section of organizations*

- This is a selection effect: Places with better match will have lower bonuses because of corollary 2.

Besley & Ghatak 2005: Competing for Workers

- What happens when the different sectors are competing for workers?
- Define $\mathcal{A}_p = \{p_0, p_1, p_2\}$ as the set of types of the principals. $\mathcal{A}_a = \{a_0, a_1, a_2\}$ is the set of types of the agents.
- A matching process is a matching function $\mu : \mathcal{A}_p \cup \mathcal{A}_a \to \mathcal{A}_p \cup \mathcal{A}_a$ such that
  1. $\mu(p_i) \in \mathcal{A}_a \cup \{p_i\} \ \forall p_i \in \mathcal{A}_p$
  2. $\mu(a_j) \in \mathcal{A}_p \cup \{a_j\} \ \forall a_i \in \mathcal{A}_a$
  3. $\mu(p_i) = a_j \iff \mu(a_j) = p_i \ \forall (p_i, a_j) \in \mathcal{A}_p \times \mathcal{A}_j$
- n_i^p = number of principals of type i. Analogously n_j^a
- Assume $n_1^a = n_1^p$ and $n_2^a = n_2^p$.
- However, allow *unemployment* $(n_0^a > n_0^p)$ and *full employment* $(n_0^a < n_0^p)$

Besley & Ghatak 2005: Competing for Workers

- Assume that the individuals on the long side of the market gets none of the surplus.
- This pins down the outside options. For any set of outside options, proposition 1 tells us the optimal contracts.

PROPOSITION 2: *Consider a matching μ and associated optimal contracts $\left(w_{ij}^{*}, b_{ij}^{*}\right)$ for $i = 0, 1, 2$ and $j = 0, 1, 2$. Then this matching is stable only if $\mu\left(p_i\right) = a_i$ for $i = 0, 1, 2$*

- Assume that when the two sectors are competing it's still worth having mission-oriented production (surplus is high enough):

$$\bar{\theta} + \hat{\pi} \geq \pi_0$$

Besley & Ghatak 2005: Competing for Workers: Full Employment

PROPOSITION 3: *Suppose that $n_0^a < n_0^p$ (full employment in the profit-oriented sector). Then the following matching μ is stable: $\mu(a_j) = p_j$ for $j = 0, 1, 2$ and the associated optimal contracts have the following features:*

(a) *The fixed wage is set at the subsistence level, i.e. $w_{ij}^* = \underline{w}$ for $j = 0, 1, 2$*

(b) *The bonus payment in the mission-oriented sector is*

$$b_{11}^* = b_{22}^* = \tfrac{1}{2} \max \left\{ \max \left\{ \bar{\theta}, \hat{\pi} \right\}, \pi_0 + \sqrt{\pi_0^2 - 4\underline{w}} - \bar{\theta} \right\}$$

and the bonus payment in the profit-oriented sector is

$$b_{00}^* = \frac{\pi_0 + \sqrt{\pi_0^2 - 4\underline{w}}}{2}$$

(c) *The optimal effort level solves: $e_{jj}^* = b_{jj}^* + \bar{\theta}$ for $j = 1, 2$ and $e_{00}^* = b_{00}^*$.*

Besley & Ghatak 2005: Competing for Workers: Full Employment

- Competition for workers and incentives interact in important ways

1. *matching effect*. Less heterogeneity in contracts compared to a world in which principals and agents don't match assortatively. When the participation constraint doesn't bind, incentive pay is lower.
2. *outside option effect.* Full employment drives profit-oriented principals' payoff to zero. Motivated agent's reservation utility is what she'd get by switching to the profit-oriented sector.
   2.1 When productivity is high in the profit-oriented sector, the mission-oriented sector has to pay more and use incentive pay more.
   2.2 Even with a binding participation constraint, incentive pay is lower in the mission-oriented sector than in the profit-oriented sector

Besley & Ghatak 2005: Competing for Workers: Unemployment

PROPOSITION 4: *Suppose that* $n_0^a > n_0^p$ *(unemployment in the profit-oriented sector). Then the following matching* μ *is stable:* $\mu(a_j) = p_j$ *for* $j = 0, 1, 2$ *and the associated optimal contracts have the following features:*

(a) The fixed wage is set at the subsistence level $w_{ij}^* = \underline{w}$ for $j = 0, 1, 2$;

(b) *The bonus payment in the mission-oriented sector is:*

$$b_{11}^* = b_{22}^* = \frac{\max\{\bar{\theta}, \hat{\pi}\} - \bar{\theta}}{2}$$

and the bonus payment in the profit-oriented sector is

$$b_{00}^* = \frac{\pi_0}{2}$$

(c) *The optimal effort level solves:* $e_{ij}^* = b_{ij}^* + \bar{\theta}$ *for* $j = 1, 2$ *and* $e_{00}^* = b_{00}^*$

Besley & Ghatak 2005: Competing for Workers

- Now there's only a matching effect.
- Application of BG framework to public sector bureaucracy
  - Lower powered incentives due to mission-oriented production
  - If an election changes the mission, may reduce productivity of bureaucracy
  - If private-sector opportunities improve $\rightarrow$ more high-powered incentives in bureaucracy
  - Lack of innovation: In profit-oriented sector, any innovation that increases π_0 will be adopted. However, in a mission-oriented organization, only innovations that increase $\pi_i + \theta_{ij}$ will be adopted. If the innovation increases π_i but decreases θ_{ij} it may not be adopted.

Outline

Theory

Financial Incentives

Non-financial Incentives

Recruitment & Selection

Open Questions

Outline

Financial Incentives

- Muralidharan & Sundararaman (JPE 2011) *Teacher Performance Pay: Experimental Evidence from India*
- Duflo, Hanna & Ryan (AER 2012) *Incentives Work: Getting Teachers to Come to School*
- Khan, Khwaja & Olken (QJE 2016) *Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors*
- Duflo Greenstone Pande & Ryan (QJE 2013): *Truth-telling by Third-party Auditors and the Response of Polluting Firms: Experimental Evidence from India*

Muralidharan & Sundararaman 2011: Introduction

- Randomized evaluation of teacher performance pay
- Large scale experiment in Andhra Pradesh, India
- Can we really measure teacher performance?
- Teacher incentive programs can backfire (multitasking, teaching to the test etc.)
- How should bonus contracts be set up?
- Are bonuses a cost-effective way to increase performance?

Muralidharan & Sundararaman 2011: Model

- Teachers do 2 tasks
  1. T_1: teaching using curricular best practices
  2. T_2: activities to increase scores on exams (drills, teaching to the test, cheating)
- t_1 and t_2 denote time allocated to these tasks. Human capital gains are

$$H = f_1 t_1 + f_2 t_2 + \varepsilon$$

  where f_1, f_2 are marginal products and ε is noise outside teacher's control
- Planner cannot observe H, t_1 or t_2 but observes performance measure P (e.g. test scores)

$$P = g_1 t_1 + g_2 t_2 + \phi$$

Muralidharan & Sundararaman 2011: Model

- Principal offers a wage contract depending on P: e.g. $w = s + bP$
- Teacher's utility is

$$U = \mathbb{E}[w] - C(t_1, t_2; \bar{t})$$

  where $\bar{t}$ is an effort norm. Teachers suffer a psychic cost if $t_1 + t_2 < \bar{t}$
- Optimal bonus b^* depends on functional form of C, but when t_1 and t_2 are substitutes, easy to construct examples s.t. $b^* = 0$: better to accept the output generated by the norm $\bar{t}$ than to distort input allocation.
- But, if $\bar{t}$ is small, then gains from increasing effort can exceed costs of distorting effort. Plausible in India: Absenteeism is very high.
- Moreover, if f_1/f_2 is not too much greater than 1, less substitution. Plausible in India: Tests are central to the system so best practice may be to teach to the test.

Muralidharan & Sundararaman 2011: Experiment

A

	India	AP
Gross enrollment (Ages 6-11) (%)	95.9	95.3
Literacy (%)	64.8	60.5
Teacher absence (%)	25.2	25.3
Infant mortality (per 1,000)	63	62

Muralidharan & Sundararaman 2011: Experiment

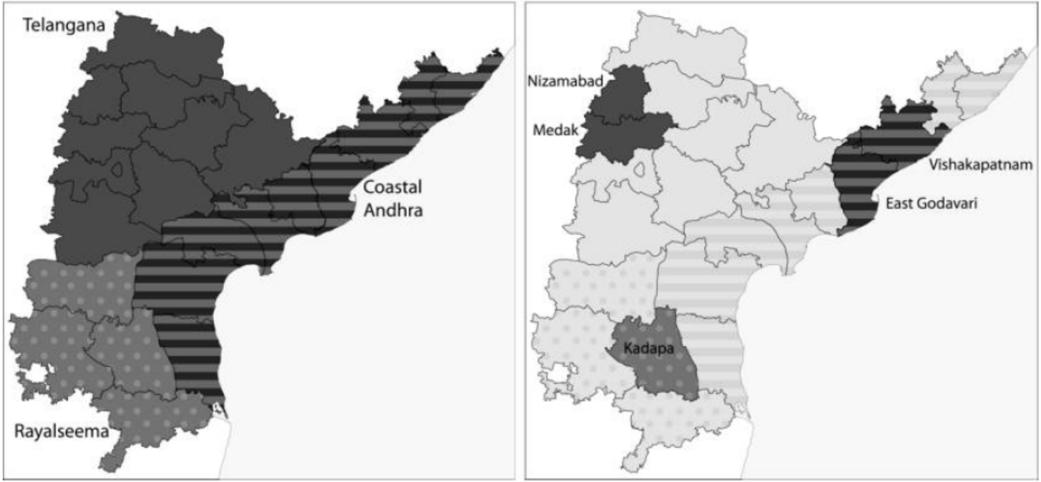


FIG. 1.—*A*, Andhra Pradesh (AP); *B*, district sampling (stratified by sociocultural regions of AP).

Muralidharan & Sundararaman 2011: Experiment

- Work in Andhra Pradesh: 5th most populous state in India. Population > 80 million
- State consists of 3 sociocultural regions and 23 districts
- Within each district randomly sample one division (out of 3-5 in each district)
- within each division randomly sample 10 mandals (out of 10-15 in each division)
- In each of the 50 mandals, randomly sample 10 schools with probabilities proportional to enrolment.
- Sample is now representative of a typical **child** attending a government-run primary school

Muralidharan & Sundararaman 2011: Experiment

TABLE 1
INCENTIVES

INPUTS	INCENTIVES (Conditional on Improvement in Student Learning)		
	None	Group Bonus	Individual Bonus
None	Control (100 schools)	100 schools	100 schools
Extra contract teacher	100 schools		
Extra block grant	100 schools		

	Control (1)	Group Incentive (2)	Individual Incentive (3)	p-Value (Equality of All Groups) (4)
	A. Means of Baseline Variables			
School-level variables:				
1. Total enrollment (baseline: grades 1–5)	113.2	111.3	112.6	.82
2. Total test takers (baseline: grades 2–5)	64.9	62.0	66.5	.89
3. Number of teachers	3.07	3.12	3.14	.58
4. Pupil-teacher ratio	39.5	40.6	37.5	.66
5. Infrastructure index (0–6)	3.19	3.14	3.26	.84
6. Proximity to facilities index (8–24)	14.65	14.66	14.72	.98
Baseline test performance:				
7. Math (raw %)	18.5	18.0	17.5	.69
8. Math (normalized; in SD)	.032	.001	−.032	.70
9. Telugu (raw %)	35.1	34.9	33.5	.52
10. Telugu (normalized; in SD)	.026	.021	−.046	.53

	B. Means of End Line Variables			
Teacher turnover and attrition:				
Year 1 (relative to year 0):				
11. Teacher attrition (%)	.30	.34	.30	.54
12. Teacher turnover (%)	.34	.34	.32	.82
Year 2 (relative to year 0):				
13. Teacher attrition (%)	.35	.38	.34	.57
14. Teacher turnover (%)	.34	.36	.33	.70
Student turnover and attrition:				
Year 1 (relative to year 0):				
15. Student attrition from baseline to end-of-year tests	.081	.065	.066	.15
16. Baseline math test score of attritors (equality of all groups)	−.17	−.13	−.22	.77
17. Baseline Telugu test score of attritors (equality of all groups)	−.26	−.17	−.25	.64
Year 2 (relative to year 0):				
18. Student attrition from baseline to end-of-year tests	.219	.192	.208	.23
19. Baseline math test score of attritors (equality of all groups)	−.13	−.05	−.14	.56
20. Baseline Telugu test score of attritors (equality of all groups)	−.18	−.11	−.21	.64

Muralidharan & Sundararaman 2011: Bonuses

- Bonuses based on average improvement in test scores

$$\text{Bonus} = \begin{cases} Rs.500 \times (\%\text{gain in avg test scores} - 5\%) & \text{if gain} > 5\% \\ 0 & \text{otherwise} \end{cases}$$

- In group incentive schools, all teachers got the same bonus based on school-level average improvement
- In individual incentive schools, based on average test score of the specific teacher.
- Slope (Rs.500) set so expected payment would equal additional spending in input treatments

Muralidharan & Sundararaman 2011: Tests

- To reduce cheating, tests conducted by external teams
- Baseline test (June-July 2005) tested math and language
- At the end of year one (March-April 2006), two rounds of tests separated by 2 weeks
  - round 1 (Lower endline, LEL) tested competencies up to previous school year
  - round 2 (higher endline, HEL) tested material from the current school year's syllabus
- Same procedure repeated at the end of year 2
- Scores in year 0 normalized to distribution across all schools
- Scores in years 1 and 2 normalized to distribution in control schools

Muralidharan & Sundararaman 2011: Results

- Teacher attrition: no significant difference in teacher attrition across schools (worried teachers try to select into incentive schools for e.g.)
- Student attrition: 7.1% attrition at year 1. 20.6% at year 2. Higher attrition for lower test score children. But balanced across treatments.

$$T_{ijkm}(Y_n) = \alpha + \gamma T_{ijkm}(Y_0) + \delta \text{Incentives} + \beta Z_m + \varepsilon_k + \varepsilon_{jk} + \varepsilon_{ijk}$$

where T_{ijkm} is score of student i in grade j at school k in mandal m. Y_0 denotes baseline tests and $Y_n$indicates test at end of n years. Z_m are mandal dummies

Muralidharan & Sundararaman 2011: Results

TABLE 3
IMPACT OF INCENTIVES ON STUDENT TEST SCORES
Dependent Variable: Normalized End-of-Year Test Score

	YEAR 1 ON YEAR 0		YEAR 2 ON YEAR 0	
	(1)	(2)	(3)	(4)
	A. Combined (Math and Language)			
Normalized lagged test score	.503***	.498***	.452***	.446***
	(.013)	(.013)	(.015)	(.015)
Incentive school	.149***	.165***	.219***	.224***
	(.042)	(.042)	(.047)	(.048)
School and household controls	No	Yes	No	Yes
Observations	42,145	37,617	29,760	24,665
R^2	.31	.34	.24	.28

Muralidharan & Sundararaman 2011: Results

	B. Math			
Normalized lagged test score	.492***	.491***	.414***	.408***
	(.016)	(.016)	(.022)	(.022)
Incentive school	.180***	.196***	.273***	.280***
	(.049)	(.049)	(.055)	(.056)
School and household controls	No	Yes	No	Yes
Observations	20,946	18,700	14,797	12,255
R^2	.30	.33	.25	.28
	C. Telugu (Language)			
Normalized lagged test score	.52***	.510***	.49***	.481***
	(.014)	(.014)	(.014)	(.014)
Incentive school	.118***	.134***	.166***	.168***
	(.040)	(.039)	(.045)	(.044)
School and household controls	No	Yes	No	Yes
Observations	21,199	18,917	14,963	12,410
R^2	.33	.36	.26	.30

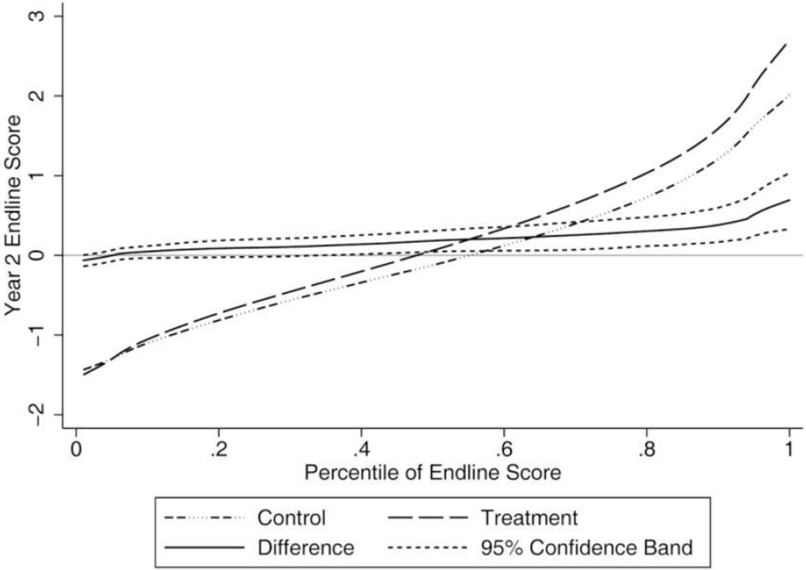
Muralidharan & Sundararaman 2011: Heterogeneity

- How does the distribution of test scores change:

$$\delta(\tau) = G_n^{-1}(\tau) - F_m^{-1}(\tau)$$

  where G_n is treatment distribution, F_m control
- NB This is a quantile treatment effect not a treatment effect at different quantiles
- Treatment effect at different quantiles: Estimate nonparametric reg of endline scores on baseline scores separately for treatment and control.
- Heterogeneity by observables:

$$\begin{aligned} T_{ijkm}(Y_n) = \alpha + \gamma T_{ijkm}(Y_0) + \delta_1 \text{Incentives} + \delta_2 \text{Characteristic} \\ + \delta_3 \text{Incentives} \times \text{Characteristic} + \beta Z_m + \varepsilon_k + \varepsilon_{jk} + \varepsilon_{ijk} \end{aligned}$$



3
2
1
0
−1
−2
Year 2 Endline Score
0
.2
.4
.6
.8
1
Percentile of Endline Score
Control
Treatment
Difference
95% Confidence Band

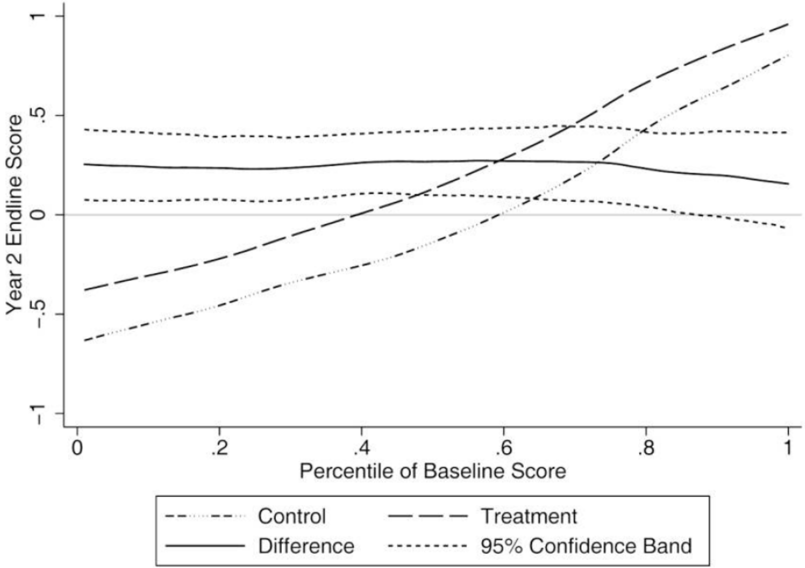

1
.5
0
-.5
-1
Year 2 Endline Score
0
.2
.4
.6
.8
1
Percentile of Baseline Score
Control
Treatment
Difference
95% Confidence Band

TABLE 6
HETEROGENOUS TREATMENT EFFECTS
A. HOUSEHOLD AND SCHOOL CHARACTERISTICS

	Log School Enrollment (1)	School Proximity (8–24) (2)	School Infrastructure (0–6) (3)	Household Affluence (0–7) (4)	Parental Literacy (0–4) (5)	Scheduled Caste/Tribe (6)	Male (7)	Normalized Baseline Score (8)
				Two-Year Effect				
Incentive	−.198	−.019	.28**	.09	.224***	.226***	.233***	.219***
	(.354)	(.199)	(.130)	(.073)	(.054)	(.049)	(.049)	(.047)
Covariate	−.065	−.005	.025	.017	.068***	−.066	.029	.448***
	(.058)	(.010)	(.038)	(.014)	(.015)	(.042)	(.027)	(.024)
Interaction	.083	.018	−.02	.038**	−.003	−.013	−.02	.006
	(.074)	(.014)	(.040)	(.019)	(.019)	(.056)	(.034)	(.031)
Observations	29,760	29,760	29,760	25,231	25,226	29,760	25,881	29,760
R^2	.244	.244	.243	.272	.273	.244	.266	.243
				One-Year Effect				
Incentive	−.36	−.076	.032	.004	.166***	.164***	.157***	.149***
	(.381)	(.161)	(.110)	(.060)	(.047)	(.045)	(.044)	(.042)
Covariate	−.128**	−.016*	−.001	.017	.08***	.007	.016	.502***
	(.061)	(.008)	(.025)	(.013)	(.012)	(.035)	(.020)	(.021)
Interaction	.103	.017	.041	.042**	−.013	−.06	.002	.000
	(.081)	(.011)	(.031)	(.017)	(.016)	(.048)	(.025)	(.026)
Observations	42,145	41,131	41,131	38,545	38,525	42,145	39,540	42,145
R^2	.31	.32	.32	.34	.34	.31	.33	.31

B. Teacher Characteristics (Pooled Regression Using Both Years of Data)

	Education (1)	Training (2)	Years of Experience (3)	Salary (Log) (4)	Male (5)	Teacher Absence (6)	Active Teaching (7)	Active or Passive Teaching (8)
Incentive	−.113	−.224	.258***	1.775**	.031	.15***	.084	.118
	(.163)	(.176)	(.059)	(.828)	(.091)	(.050)	(.054)	(.074)
Covariate	.003	−.051	−.001	−.034	−.084	−.149	.055	.131
	(.032)	(.041)	(.003)	(.066)	(.057)	(.137)	(.078)	(.093)
Interaction	.086*	.138**	−.009**	−.179*	.09	.013	.164*	.064
	(.050)	(.061)	(.004)	(.091)	(.069)	(.171)	(.098)	(.111)
Observations	53,737	53,890	54,142	53,122	54,142	53,609	53,383	53,383
R^2	.29	.29	.29	.29	.29	.29	.29	.29

Impact of Incentives on Nonincentive Subjects
Dependent Variable: Normalized End Line Score

	Year 1		Year 2	
	Science	Social Studies	Science	Social Studies
	A. Reduced-Form Impact			
Normalized baseline math score	.215***	.224***	.156***	.167***
	(.019)	(.018)	(.023)	(.024)
Normalized baseline language score	.209***	.289***	.212***	.189***
	(.019)	(.019)	(.023)	(.024)
Incentive school	.112**	.141***	.113**	.18***
	(.052)	(.048)	(.044)	(.050)
Observations	11,786	11,786	9,143	9,143
R^2	.26	.31	.19	.18

Group versus Individual Incentives
Dependent Variable: Normalized End-of-Year Test Score

	Year 1 on Year 0			Year 2 on Year 0		
	Combined (1)	Math (2)	Telugu (3)	Combined (4)	Math (5)	Telugu (6)
Individual incentive school	.156***	.184***	.130***	.283***	.329***	.239***
	(.050)	(.059)	(.045)	(.058)	(.067)	(.054)
Group incentive school	.141***	.175***	.107**	.154***	.216***	.092*
	(.050)	(.057)	(.047)	(.057)	(.068)	(.052)
F-statistic p-value (testing group incentive school = individual incentive school)	.765	.889	.610	.057	.160	.016
Observations	42,145	20,946	21,199	29,760	14,797	14,963
R^2	.31	.299	.332	.25	.25	.26

IMPACT OF INPUTS VERSUS INCENTIVES ON LEARNING OUTCOMES
Dependent Variable: Normalized End-of-Year Test Score

	YEAR 1 ON YEAR 0			YEAR 2 ON YEAR 0		
	Combined (1)	Math (2)	Language (3)	Combined (4)	Math (5)	Language (6)
Normalized lagged score	.512***	.494***	.536***	.458***	.416***	.499***
	(.010)	(.012)	(.011)	(.012)	(.016)	(.012)
Incentives	.15***	.179***	.121***	.218***	.272***	.164***
	(.041)	(.048)	(.039)	(.049)	(.057)	(.046)
Inputs	.102***	.117***	.086**	.085*	.089*	.08*
	(.038)	(.042)	(.037)	(.046)	(.052)	(.044)
F-statistic *p*-value (inputs = incentives)	.178	.135	.298	.003	.000	.044
Observations	69,157	34,376	34,781	49,503	24,628	24,875
R^2	.30	.29	.32	.225	.226	.239

Outline

Financial Incentives

- Muralidharan & Sundararaman (JPE 2011) *Teacher Performance Pay: Experimental Evidence from India*
- Duflo, Hanna & Ryan (AER 2012) *Incentives Work: Getting Teachers to Come to School*
- Khan, Khwaja & Olken (QJE 2016) *Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors*
- Duflo Greenstone Pande & Ryan (QJE 2013): *Truth-telling by Third-party Auditors and the Response of Polluting Firms: Experimental Evidence from India*

Duflo et al. 2012: Introduction

- Access to primary school has increased dramatically in low-income countries, but school quality hasn't
  - 65% of children in grades 2-5 in Indian government schools in 2006 couldn't read a simple paragraph (Pratham 2006)
  - 24% of teachers in India are absent in unannounced visits (Kremer et al 2005)
- This paper: Experiment and structural model of direct monitoring of para-teachers' attendance in India.
- Ambiguous effect on presence.
  - Standard labor supply model predicts more effort, but only if strong enough incentives
  - Incentives could crowd out intrinsic motivation (Benabou & Tirole 2006).
  - Teachers may stop working after reaching target income (Fehr & Goette 2007)

Duflo et al. 2012: Introduction

- Will presence increase learning?
  - Multitasking means incentives for presence could crowd out other dimensions of effort (Holmstrom & Milgrom 1991).
  - Incentives may demoralize teacher or reduce their intrinsic motivation to teach.
- But if the main reason people don't show up is the opportunity cost of being at the school and the marginal cost of teaching once you're at the school is low, this might just work.

Duflo et al. 2012: Setting & Experiment

- The setting are rural nonformal education centers (NFEs) in Udaipur, Rajasthan, India.
- In september 2003 Seva Mandir, the operator chose 120 schools for the experiment.
  - In 60 schools the teachers got a camera and were told that one of the students had to take a photograph of the teacher with the children at the start and the end of the day. Cameras had a tamper-proof date & time function.
  - The other 60 schools are controls
- Teachers' base salary was Rs. 1,000 for at least 20 days of work per month.
- Treatment teachers got a Rs. 50 bonus for every day in excess of 20 days and a Rs. 50 fine for each day of the 20 that they skip. Fines capped at Rs. 500

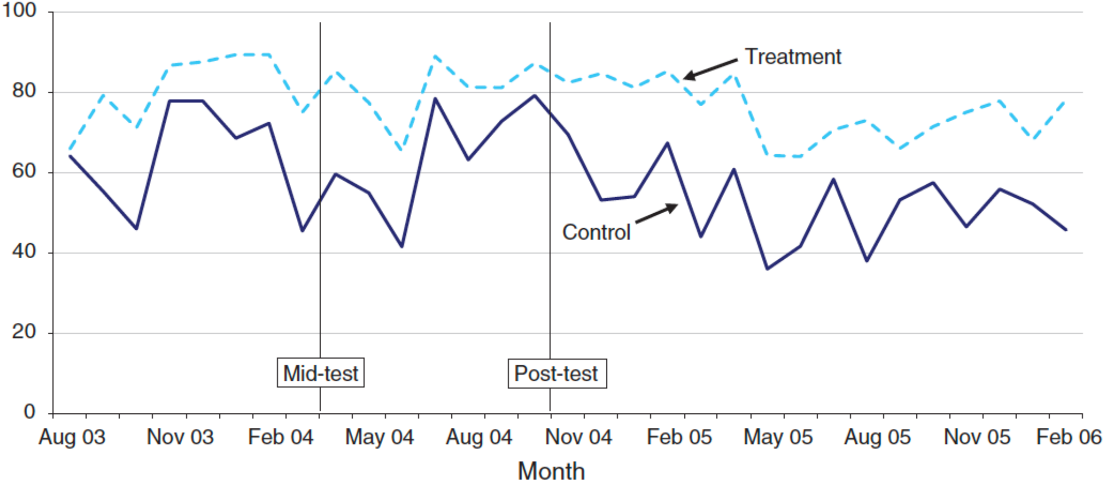
Duflo et al. 2012: Data

- Attendance data
  1. 1 random, unannounced visit to each school each month.
  2. Camera and payment data for treatment schools
- Additional data from random checks. How many children, whether anything on the board, whether the teacher was talking to the children, and roll call.
- 3 basic competency exams. Oral exams testing simple math, basic Hindi vocabulary. Written exam testing addition, multiplication, ability to construct sentences, and reading comprehension.
  1. a pretest in August 2003
  2. a mid-test in April 2004
  3. a post-test in September 2004

TABLE 1—BASELINE DATA

	Treatment (1)	Control (2)	Difference (3)
Panel A. Teacher attendance			
School open	0.66	0.64	0.02
			(0.11)
	41	39	80
Panel B. Student participation (random check)			
Number of students present	17.71	15.92	1.78
			(2.31)
	27	25	52
Panel C. Teacher qualifications			
Teacher test scores	34.99	33.54	1.44
			(2.02)
	53	54	107
Panel D. Teacher performance measures (random check)			
Percentage of children sitting within classroom	0.83	0.84	0.00
			(0.09)
	27	25	52
Percent of teachers interacting with students	0.78	0.72	0.06
			(0.12)
	27	25	52
Blackboards utilized	0.85	0.89	−0.04
			(0.11)
	20	19	39
F-stat (1,110)			1.21
p-value			(0.27)
Panel E. Baseline test scores			
Took written exam	0.17	0.19	−0.02
			(0.04)
	1,136	1,094	2,230
Total score on oral exam	−0.08	0.00	−0.08
			(0.07)
	940	888	1,828
Total score on written exam	0.16	0.00	0.16
			(0.19)
	196	206	402

Duflo et al. 2012: Attendance Increased

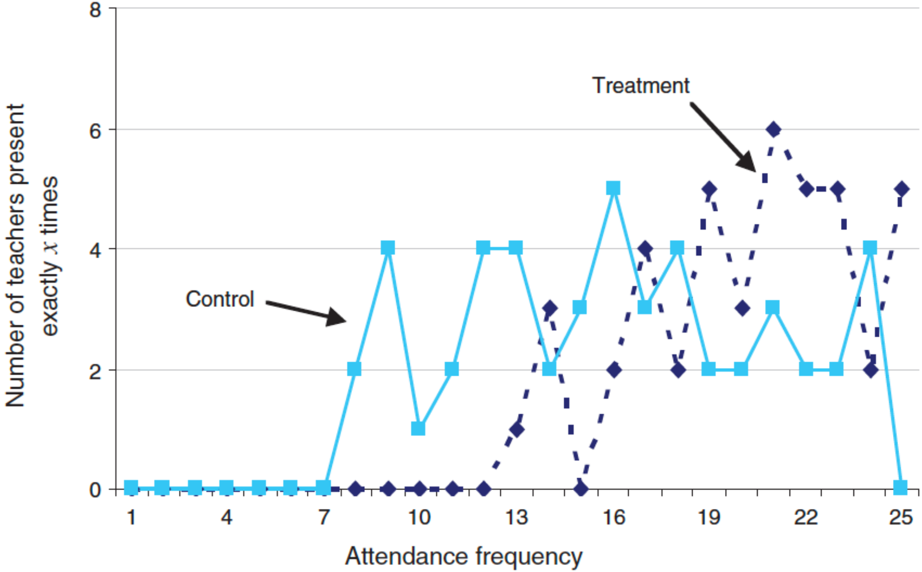


Duflo et al. 2012: Attendance Increased

Table 2—Teacher Attendance

September 2003–February 2006			Difference between treatment and control schools		
Treatment	Control	Diff	Until mid-test	Mid- to post-test	After post-test
(1)	(2)	(3)	(4)	(5)	(6)
Panel A. All teachers					
0.79	0.58	0.21	0.20	0.17	0.23
		(0.03)	(0.04)	(0.04)	(0.04)
1,575	1,496	3,071	882	660	1,529
Panel B. Teachers with above median test scores					
0.78	0.63	0.15	0.15	0.15	0.14
		(0.04)	(0.05)	(0.05)	(0.06)
843	702	1,545	423	327	795
Panel C. Teachers with below median test scores					
0.78	0.53	0.24	0.21	0.14	0.32
		(0.04)	(0.05)	(0.06)	(0.06)
625	757	1,382	412	300	670

Duflo et al. 2012: Attendance Increased



Duflo et al. 2012: Financial Incentives

- People in the treatment group got both financial incentives and monitoring, so difficult to disentangle the two
- The cap of Rs. 500 on the fine makes the incentive scheme non-linear permiting an assessment of the financial incentives independent of the monitoring as follows:
  - Imagine a teacher who was sick a lot one month and missed most of the first 20 days of school. Assume on day 21 he has worked 5 days and has 5 days to go. Even if he works all 5 days, he will earn Rs. 500, the same as if he works none. At the start of the next month the clock resets, so he has an incentive to start working again.
  - Now imagine a teacher who has worked 10 days by the 21st day of the month. She earns Rs. 50 for every day she works. No different before or after the end of the month.

Duflo et al. 2012: Financial Incentives

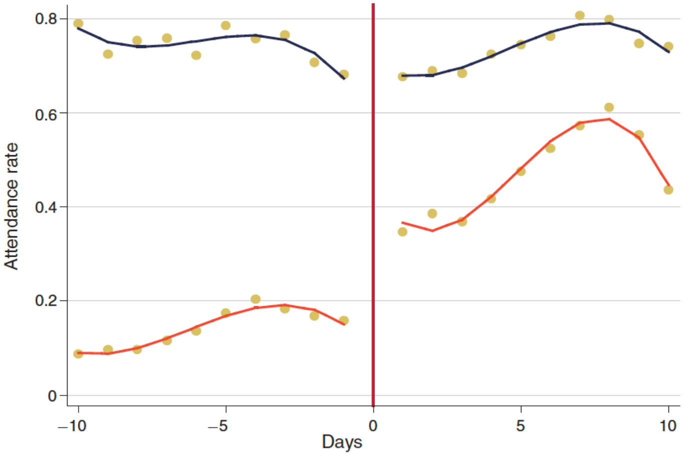


FIGURE 3. RDD REPRESENTATION OF TEACHER ATTENDANCE AT THE START AND END OF THE MONTH

Notes: The top lines represent the months in which the teacher is in the money, while the bottom lines represent the months in which the teacher is not in the money. The estimation includes a third-order polynomial of days on the left and right side of the change of month.

Duflo et al. 2012: Financial Incentives

- For teachers in the treatment group. create a dataset of attendance records around the end of the month. The last day of a month and the first day of the next month form a pair m
- $Work_{itm}$ is a dummy for working in day t of pair m

$$\begin{aligned} Work_{itm} = \alpha + \beta \mathbf{1}_{im}\left(d > 10\right) + \gamma Firstday_t \\ + \lambda \mathbf{1}_{im}\left(d > 10\right) \times Firstday_t + v_i + \mu_m + \epsilon_{itm} \end{aligned}$$

where $\mathbf{1}_{im}\left(d > 10\right)$ is a dummy =1 in both t if the teacher is “in the money” and $Firstday_t$ indicates the first day of the month

Duflo et al. 2012: Financial Incentives

TABLE 3—DO TEACHERS WORK MORE WHEN THEY ARE "IN THE MONEY"?

	(1)	(2)	(3)	(4)
Beginning of month	0.19	0.12	0.46	0.39
	(0.05)	(0.06)	(0.04)	(0.03)
In the money	0.52	0.37	0.6	0.48
	(0.04)	(0.05)	(0.03)	(0.01)
Beginning of the month × in the money	−0.19	−0.12	−0.34	−0.3
	(0.06)	(0.06)	(0.04)	(0.02)
Observations	2,813	2,813	27,501	27,501
R^2	0.06	0.22	0.08	0.16
Sample	First and last day of month	First and last day of month	First ten and last ten days of month	First ten and last ten days of month
Third-order polynomial on days on each side			X	X
Teacher fixed effects		X		X
Month fixed effects		X		X
Clustered standard errors	X		X	

Duflo et al. 2012: Dynamic Labor Supply Model

- Teachers on day $t = \{1, \ldots, T_m\}$ of month m value consumption C_{tm} and leisure L_{tm}

  $$U_{tm} = U\left(C_{tm}, L_{tm}\right) = \beta C_{tm}\left(\pi_m\right) + \left(\mu_{tm} - P\right) L_{tm}$$

  where P is nonpecuniary cost of missing work.
- Consumption depends on earned income π_m, β turns rupees of consumption into utility.
- L_{tm} is 1 if the teacher doesn't attend work, and zero otherwise.
- Leisure coefficient has deterministic and stochastic parts

  $$\mu_{tm} = \mu + \epsilon_{tm}$$

  where ϵ_{tm} is assumed to be normal.

Duflo et al. 2012: Dynamic Labor Supply Model

- Not attending school has two costs: P and a probability $p_m(t,d)$ of being fired that depends on the number of days worked d by time t in month m. If they are fired, teachers get F, their outside option.
- Income in the treatment group is

$$\pi_m = 500 + 50\max\{0, d_{m-1} - 10\}$$

while in the control group π_m is Rs. 1000.

- Control group has simple binary choice. Bellman equation on every day except last day of the month:

$$\begin{aligned} V_m(t,d;\epsilon_{tm}) = p_m(t,d)\,F + [1 - p_m(t,d)] \times \\ \max\{\mu - P + \epsilon_{tm} + EV_m(t+1,d;\epsilon_{t,m+1}) \\ , EV_m(t+1,d+1;\epsilon_{t,m+1})\} \end{aligned}$$

Duflo et al. 2012: Dynamic Labor Supply Model

- Treatment group have a very different problem to solve since they face incentives for attendance.
- In periods $t < T_m$

$$\begin{aligned} V_m\left(t, d; \epsilon_{tm}\right) = p_m\left(t, d\right) + \left(1 - p_m\left(t, d\right)\right) \times \\ \max\left\{\mu - P + \epsilon_{tm} + EV_m\left(t+1, d; \epsilon_{t,m+1}\right)\right. \\ \left., EV_m\left(t+1, d+1; \epsilon_{t,m+1}\right)\right\} \end{aligned}$$

- In period T_m

$$\begin{aligned} V_m\left(T_m, d; \epsilon_{T_m,m}\right) = p_m\left(T_m, d\right) F + \left[1 - p_m\left(T_m, d\right)\right] \times \\ \max\left\{\mu - \bar{P} + \epsilon_{T_m,m} + \beta\pi\left(d\right)\right. \\ + EV_{m+1}\left(1, 0; \epsilon_{t,m+1}\right) \\ \left., \beta\pi\left(d+1\right) + EV_{m+1}\left(1, 0; \epsilon_{t,m+1}\right)\right\} \end{aligned}$$

Duflo et al. 2012: Estimation

- In period T_m, EV_{m+1} doesn't depend on action in T_m so we can solve the model backwards.
- Need to make some assumptions about μ and the distribution of ϵ
- In the data noone ever gets fired, so assume that teachers perceive $p_m(t,d) = 0$
- Model 1: iid errors. Simplest case. In period $t < T$

$$\begin{aligned}
\mathbb{P}(work; t, d, \theta) &= \mathbb{P}(\mu + \epsilon_{tm} + EV(t+1, d) < EV(t+1, d+1)) \\
&= \mathbb{P}(\epsilon_{tm} < EV(t+1, d+1) - EV(t+1, d) - \mu) \\
&= \Phi(EV(t+1, d+1) - EV(t+1, d) - \mu)
\end{aligned}$$

Duflo et al. 2012: Estimation

- Each value function can be computed using backword recursion.
- Let w_{imt} be an indicator for working on day t in month m. Then the log likelihood is

$$LLH(\theta) = \sum_{i=1}^{N} \sum_{m=1}^{M_i} \sum_{t=1}^{T_m} \left[w_{imt} \mathbb{P}(work; t, d, \theta) \right. \\ \left. + (1 - w_{imt})(1 - \mathbb{P}(work, t, d, \theta)) \right]$$

- This likelihood is concave and can be evaluated quickly, no numerical integration is needed. Just need to evaluate it at many points.

Duflo et al. 2012: Estimation

- Now introduce some serial correlation in two ways.
- Approach 1: Serially correlated preference shocks

$$\mu_{mt} = \mu + w_{m,t-1}\gamma$$

- Now the likelihood is

$$LLH(\theta) = \sum_{i=1}^{N}\sum_{m=1}^{M_i}\sum_{t=1}^{T_m}\left[w_{imt}\mathbb{P}\left(work; t, d, \theta, w_{m,t-1}\right)\right.$$
$$\left. + \left(1 - w_{imt}\right)\left(1 - \mathbb{P}\left(work, t, d, \theta, w_{m,t-1}\right)\right)\right]$$

- Approach 2: Serially correlated cost shocks:

$$\epsilon_{mt} = \rho\epsilon_{m,t-1} + \nu_{mt}$$

- Can't estimate this by ML, need to use Method of Simulated Moments. Match sequences of attendance of length 5.

Duflo et al. 2012: Estimation

- Extend the above in 2 ways

1. Incorporate observables into μ. Use attendance in control group in same geographic block and teacher's score on the admission exam to shift μ
2. Relax assumption that the outside option is the same for everyone. Estimate fixed effects μ_i or random coefficients μ_{im} drawn from normal distribution or from a mixture of two normally distributed types.

Parameter	Model I (1)	Model II (2)	Model III (3)	Model IV (4)	Model V (5)	Model VI (6)	Model VII (7)	Model VIII (8)
β	0.049	0.027	0.055	0.057	0.013	0.017	0.017	0.016
	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
μ_1	1.564		1.777	1.778	−0.428	−0.304	−0.160	−0.108
	(0.013)		(0.013)	(0.021)	(0.045)	(0.042)	(0.092)	(0.057)
ρ			0.422	0.412	0.449			
			(0.030)	(0.021)	(0.043)			
σ_1^2				0.043	0.007	0.252	0.418	0.235
				(0.012)	(0.019)	(0.015)	(0.052)	(0.028)
μ_2					1.781			
					(0.345)			
σ_2^2					0.050			
					(0.545)			
ρ					0.024			
					(0.007)			
Yesterday shifter						0.094	0.024	0.095
						(0.010)	(0.009)	(0.014)
Attendance								−0.132
								(0.095)
Test score								−0.005
								(0.002)
Heterogeneity	None	FE	None	RC	RC	RC	RC	RC
Three-day window	No	No	No	No	No	No	Yes	No

Duflo et al. 2012: Counterfactual Policies

- With the model we can do counterfactuals. Authors use model V.
- Find the cost minimizing combination of the bonus size and the threshold to get a bonus that yield a particular number of expected work days.

Expected days worked (1)	Bonus cutoff (2)	Bonus (3)	Expected cost (4)	Test score gain over control group (13 days) (5)
14	0	0	500	0.04
15	21	25	521	0.07
16	22	75	664	0.11
17	21	75	672	0.15
18	20	75	755	0.18
19	20	100	921	0.22
20	20	125	1,112	0.26
21	16	225	2,642	0.29
22	11	275	4,604	0.33

Duflo et al. 2012: Teacher performance

TABLE 6—TEACHER PERFORMANCE

	September 2003–February 2006			Difference between treatment and control schools		
	Treatment (1)	Control (2)	Diff. (3)	Until mid-test (4)	Mid- to post-test (5)	After post-test (6)
Percent of children sitting within classroom	0.72	0.73	−0.01	0.01	0.04	−0.01
			(0.01)	(0.89)	(0.03)	(0.02)
	1,239	867	2,106	643	408	983
Percent of teachers interacting with students	0.55	0.57	−0.02	−0.02	0.05	−0.04
			(0.02)	(0.04)	(0.05)	(0.03)
	1,239	867	2,106	643	480	983
Blackboards utilized	0.92	0.93	−0.01	−0.03	0.01	−0.01
			(0.01)	(0.02)	(0.02)	(0.02)
	990	708	1,698	613	472	613

Notes: Teacher Performance Measures from Random Checks include only schools that were open during the random check. Standard errors are clustered by school.

Table 7—Child Attendance

	September 2003–February 2006			Difference between treatment and control schools		
	Treatment (1)	Control (2)	Diff (3)	Until mid-test (4)	Mid- to post-test (5)	After post-test (6)
Panel A. Attendance conditional on school open						
Attendance of students present at pretest exam	0.46	0.46	0.01	0.02	0.03	0.00
			(0.03)	(0.03)	(0.04)	(0.03)
	23,495	16,280	39,775			
Attendance for children who did not leave NFE	0.62	0.58	0.04	0.02	0.04	0.05
			(0.03)	(0.03)	(0.04)	(0.03)
	12,956	10,737	23,693			
Panel B. Total instruction time (presence)						
Presence for students present at pretest exam	0.37	0.28	0.09	0.10	0.10	0.08
			(0.03)	(0.03)	(0.04)	(0.03)
	29,489	26,695	56,184			
Presence for student who did not leave NFE	0.50	0.36	0.13	0.10	0.13	0.15
			(0.03)	(0.04)	(0.05)	(0.04)
	16,274	17,247	33,521			

Duflo et al. 2012: Student Achievement

- Run treatment regressions of scores in mid- and end-term exams. Test scores are highly autocorrelated so gain lots of precision by controlling for pre-scores.

$$
\begin{aligned}
Score_{ijk} = \beta_1 + \beta_2 Treat_j + \beta_3 Pre\_Writ_{ij} + \beta_r Oral\_Score_{ij} \\
+ \beta_5 Written\_Score_{ij} + \varepsilon_{ijk}
\end{aligned}
$$

  where Pre_Writ_{ij} is a dummy for taking the written test at baseline (they did either the written or oral test), $Oral\_Score_{ij}$ is the score on the oral exam (or 0 if did the written exam) and $Written\_Score_{ij}$ is the score on the written exam (or 0 if did the oral exam).

TABLE 9—ESTIMATION OF TREATMENT EFFECTS FOR THE MID- AND POST-TEST

	Mid-test				Post-test			
	Took written (1)	Math (2)	Lang. (3)	Total (4)	Took written (5)	Math (6)	Lang. (7)	Total (8)
Panel A. All children								
	0.04	0.15	0.16	0.17	0.06	0.21	0.16	0.17
	(0.03)	(0.07)	(0.06)	(0.06)	(0.04)	(0.12)	(0.08)	(0.09)
	1,893	1,893	1,893	1,893	1,760	1,760	1,760	1,760
Panel B. With controls								
	0.04	0.13	0.14	0.14	0.06	0.18	0.14	0.15
	(0.03)	(0.07)	(0.06)	(0.06)	(0.04)	0.13	0.08	0.09
	1,752	1,752	1,752	1,752	1,760	1,760	1,760	1,760
Panel C. Took pretest oral								
		0.14	0.13	0.15		0.2	0.13	0.16
		(0.08)	(0.06)	(0.07)		(0.14)	(0.09)	(0.10)
		1,550	1,550	1,550		1,454	1,454	1,454
Panel D. Took pretest written								
		0.19	0.28	0.25		0.28	0.28	0.25
		(0.12)	(0.11)	(0.11)		(0.18)	(0.11)	(0.12)
		343	343	343		306	306	306

Outline

Financial Incentives

Muralidharan & Sundararaman (JPE 2011) *Teacher Performance Pay: Experimental Evidence from India*

Duflo, Hanna & Ryan (AER 2012) *Incentives Work: Getting Teachers to Come to School*

Khan, Khwaja & Olken (QJE 2016) *Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors*

Duflo Greenstone Pande & Ryan (QJE 2013): *Truth-telling by Third-party Auditors and the Response of Polluting Firms: Experimental Evidence from India*

Khan et al 2016: Introduction

- Focus on a bureaucrat (tax inspector) who has to assess, enforce, and audit property taxes.
- One possibility: sell the right to collect taxes to the bureaucrat and then let the bureaucrat keep most/all of the revenue: “Tax Farming” common throughout history (Rome, France etc.)
- But this is a setting where the bureaucrat and the citizen might collude.
    - performance incentives affect division of rents without necessarily increasing revenue.
    - e.g. if collusion is costless, then performance pay increases bribes without increasing revenue since taxpayer now has to compensate the inspector for the foregone bonus.
- Conduct an experiment in Punjab, Pakistan to investigate these issues.

Khan et al 2016: Setting

- Punjab population >80 million. Very low property tax collection.
- Tax levied on the Gross Annual Rental Value (GARV) of the property. Determined by a formula with size of land and buildings as inputs.
- Different rates for owner-occupied and rental properties. And for residential and commercial. This is main way tax evasion happens.
- Tax administered in “circles”: geographic areas covering 2-10K houses.
  - Each circle has 3 officers: An “inspector” (boss) a “clerk” (records) and a “constable” (assists inspector)
  - Each year they send the taxpayer a bill.

Khan et al 2016: Model

- Taxpayer has true tax liability τ_i^*. Inspector knows τ_i^* but can choose to report $\tau_i < \tau_i^*$
- Inspector receives incentive payment $r\tau_i$
- Colluding to underreport is costly
  - taxpayer cost: $\alpha\left(\tau_i^* - \tau_i\right)$
  - inspector cost: $\beta\left(\tau_i^* - \tau_i\right)$
- Taxpayer and inspector Nash bargain over bribe b_i
- If they don't agree a bribe, the taxpayer gets $-\tau_i^*$ and the inspector gets $r\tau_i^*$
- If they agree, the taxpayer gets $-\tau_i - \alpha_i\left(\tau_i^* - \tau_i\right) - b_i$ and the inspector gets $r\tau_i - \beta\left(\tau_i^* - \tau_i\right) + b_i$

Khan et al 2016: Model

- Joint surplus is

$$\underbrace{\tau_i^* - \tau_i - \alpha_i\left(\tau_i^* - \tau_i\right) - b_i}_{\text{taxpayer}} + \underbrace{r\left(\tau_i - \tau_i^*\right) - \beta_i\left(\tau_i^* - \tau_i\right) + b_i}_{\text{inspector}}$$

$$= \quad -\tau_i\left(1 - \alpha_i - \beta_i - r\right) + \tau_i^*\left(1 - \alpha_i - \beta_i - r\right)$$

- Two cases

$$\tau_i = \begin{cases} 0 & \text{if } r + \alpha_i + \beta_i < 1 \\ \tau_i^* & \text{if } r + \alpha_i + \beta_i > 1 \end{cases}$$

- With bargaining weight γ_i on the taxpayer

$$b_i = \left[\left(\beta_i + r\right)\gamma_i + \left(1 - \gamma_i\right)\left(1 - \alpha_i\right)\right]\tau_i^*$$

Khan et al 2016: Model

- What happens to total revenue T? Denote $f(\alpha, \beta, \tau^*)$ joint dist of α, β, τ^*

$$\begin{aligned}\frac{dT}{dr} &= \iiint\limits_{r+\alpha+\beta=1, \tau^* \in (0,\infty)} \tau^* f(\alpha, \beta, \tau^*)\, d\alpha d\beta d\tau^* \\ &= \iint\limits_{\alpha, \tau^*} \tau^* f(\alpha, 1-r-\alpha, \tau^*)\, d\alpha d\tau^*\end{aligned}$$

- If we increase r the effect depends on how many people are on the margin between the two cases
- Simplifying assumptions to note
  - linearity
  - τ^* known, no effort by bureaucrat
  - outside option is τ^* so no extortion/overtaxation

Khan et al 2016: Experiment

- Treatment 1: *Revenue* based performance pay:

$$\text{Bonus}_c = \alpha_c \max\{\text{Revenue}_c - \text{Benchmark}_c, 0\}$$

- Benchmark calculated from 3-year averages of historical collection plus normal rate of increase.
- α_c was 40% for lowest 50% of circles, 30% between 50th and 75th percentiles, 20% at the top.
- Note incentive depends only on revenue, not whether it's more money from existing houses or finding new houses to bring into the tax net.
- Bonus divided 40%-30%-30% between inspector, constable & clerk.

Khan et al 2016: Experiment

- Treatment 2. *Revenue Plus* performance pay:
- Same as above, but also incentives to address multitasking
- Pay adjusted according to taxpayer satisfaction (survey of 21,000 households) and assessment accuracy $(1 - |\text{survey } GARV/\text{official } GARV|)$
- Circles ranked and divided into three equal-sized groups
- Top group got another bonus of 0.75$\times$base pay
- Bottom group lost 0.75$\times$base pay (subject to total experimental payments >0)

Khan et al 2016: Experiment

- Treatment 3. *Flexible bonus*:
- Bonuses analogous to private sector.
- Performance Evaluation Committee ranks the circles and divided into 3 groups.
- Same adjustments to payout as in Revenue Plus treatment
- Treatment 4: *Information only* in year 2. Same treatment (monitoring, reports, meetings etc) except no payments
- Treatment 5: *Supervisors performance pay* in year 2. Similar to Revenue treatment, but for the supervisors (only 26 treatments, 25 controls though)

Khan et al 2016: Experiment

Experimental Design

	Randomization		Implementation	
	Year 1	Year 2	Year 1	Year 2
Revenue	53	72	47	68
Revenue plus	54	74	48	68
Flexible bonus	54	73	49	67
Information	0	70	0	66
Control	322	194	338	213

Khan et al 2016: Measurement

- 2 main data sources
  1. Administrative data at the circle-level
  2. Survey data at property/taxpayer-level to measure accuracy of tax assessment, customer satisfaction and corruption.
- Create a measure of under/overtaxation

$$TaxGap = \frac{GARV_{\text{Inspector}} - GARV_{\text{Survey}}}{GARV_{\text{Inspector}} + GARV_{\text{Survey}}}$$

  sample mean is -0.10 suggesting undertaxation
- Use absolute value of $TaxGap$ as measure of inaccuracy (mean 0.34)
- Bribes are common: Respondents report average bribes of Rs.2000 (US$20) which $\sim$ 50% of reported property taxes paid

Khan et al 2016: Estimation

- Small amount of attrition (refuse to participate/transferred out) so use 2SLS: Instrument actual treatment status with randomization outcome

$$\ln Y_{cst} = \alpha_s + \beta Treatment_{cst} + \gamma \ln Y_{cs0} + \epsilon_{cst}$$

  where Y_{cst} is outcome in circle c in stratum s at time t. $Treatment_{cst}$ is continuous between 0 & 1 representing fraction of treated circle staff present in circle c in the last quarter of year t.
- For survey-based outcomes estimate

$$Y_{ics} = \alpha_s + \beta Treatment_{cs} + \epsilon_{ics}$$

	(1)	(2)	(3)	(4)	(5)	(6)
	Year 1			Year 2		
	Total	Current	Arrears	Total	Current	Arrears
Panel A: Main treatment						
Any treatment	0.091***	0.073***	0.152**	0.094***	0.091***	0.113
	(0.028)	(0.027)	(0.069)	(0.031)	(0.032)	(0.083)
Panel B: Subtreatments						
Revenue	0.118***	0.109***	0.134	0.129***	0.152***	0.005
	(0.035)	(0.034)	(0.099)	(0.043)	(0.044)	(0.133)
Revenue plus	0.080	0.086*	0.072	0.093**	0.081*	0.175
	(0.053)	(0.052)	(0.110)	(0.045)	(0.049)	(0.114)
Flexible bonus	0.071*	0.024	0.243**	0.056	0.035	0.148
	(0.038)	(0.035)	(0.098)	(0.041)	(0.042)	(0.108)
N	481	481	481	482	482	479
Mean of control group	15.671	15.379	14.030	15.745	15.518	13.915
Rev. vs. multitasking p	0.323	0.193	0.830	0.233	0.049	0.262
Objective vs. subjective p	0.530	0.090	0.212	0.220	0.084	0.634
Equality of schemes p	0.562	0.143	0.433	0.359	0.086	0.527
Joint significance p	0.004	0.010	0.073	0.012	0.005	0.305

	(1) Quality	(2) Satisfaction	(3) Inaccuracy	(4) Tax gap
Panel A: Main treatment				
Any treatment	−0.006	−0.011	0.004	0.007
	(0.022)	(0.022)	(0.012)	(0.022)
Panel B: Subtreatments				
Revenue	0.006	−0.006	0.002	−0.022
	(0.036)	(0.037)	(0.017)	(0.029)
Revenue plus	0.040	0.029	0.028*	0.015
	(0.026)	(0.027)	(0.016)	(0.032)
Flexible bonus	−0.060*	−0.053*	−0.016	0.029
	(0.031)	(0.032)	(0.018)	(0.031)
N	6050	6050	9870	9870
Sample	Phase 1	Phase 1	Full	Full
Mean of control group	0.538	0.555	0.339	−0.103
Rev. vs. multitasking p	0.683	0.876	0.813	0.159
Objective vs. subjective p	0.015	0.064	0.099	0.315
Equality of schemes p	0.014	0.059	0.090	0.344
Joint significance p	0.035	0.129	0.160	0.533

Khan et al 2016: Collusion

- The model suggests that the incentive payments will alter the nature of collusion

1. How many properties get revalued? i.e. how often does collusion break down?
2. How does collusion change? if collusion breaks down, tax should go up and bribes should go down. If collusion doesn't break down bribes + tax should go up. How shared depends on γ
3. When does collusion break down? Which properties are the ones the inspectors reassess?

Impacts on Number of Reassessed Properties

	(1) Total number of section 9 properties added to tax rolls in treatment period	(2) Number of new properties added to tax rolls in treatment period	(3) Number of reassessed properties added to tax rolls in treatment period
Treatment	83.0*	74.0**	9.0
	(45.27)	(34.39)	(22.35)
N	234	234	234
Mean of control group	96.7	36.7	60.0

Impacts on Tax Payments and Corruption, by Reassessed Status

	(1) Self-reported tax payment	(2) Bribe payment	(3) Frequency of bribe payment	(4) Perception of corruption
Panel A: General population sample only				
Treatment	−62.81	594.1*	0.2021**	0.0113
	(264.7)	(341.7)	(0.0951)	(0.0254)
N	11,586	5,993	4,802	6,050
Mean of control group	4,069.425	1,874.542	0.683	0.644
Panel B: Reassessed and general population sample				
Reassessed * treatment	1,884*	−557.4	−0.1592*	−0.0031
	(1,083)	(380.1)	(0.0942)	(0.0221)
Reassessed	2,763***	−66.38	0.0137	−0.0191*
	(572.9)	(177.5)	(0.0403)	(0.0107)
N	16,353	8,207	6,993	8,268
Sample	Full	Phase 1	Phase 1	Phase 1
Mean of control group in gen. pop. sample	3928.252	1874.542	0.683	0.644

SELECTION EFFECTS ON REASSESSMENTS

	(1) GARV	(2) Number of floors	(3) Last renovation was ≤ 2 years ago	(4) Land area (sq. feet)	(5) Total covered area (sq. feet)	(6) Main Road	(7) Tax Category	(8) Percent of property commercial	(9) Percent of property commercial and rented	(10) Tax Liability
Reassess * treatment	20,137.796	0.002	−0.005	−32.599	852.092	−0.002	−0.226**	0.018	0.075**	3,897.980
	(16,187.550)	(0.050)	(0.020)	(82.473)	(771.516)	(0.048)	(0.088)	(0.037)	(0.029)	(3,539.474)
Reassess	24,683.609***	0.078***	0.094***	37.396	−156.619	0.064***	0.212***	0.217***	0.176***	5,503.481***
	(7,944.915)	(0.026)	(0.011)	(57.199)	(379.299)	(0.024)	(0.044)	(0.019)	(0.015)	(1,754.013)
N	15,090	16,352	16,354	16,352	16,352	16,352	15,090	16,226	16,227	15,090
Mean of control group in gen. pop. sample	36,808.77	1.57	0.02	301.13	2,779.82	0.46	3.78	0.35	0.17	6,642.00

Notes. Property-level 2SLS regressions. Specifications follow equation (12) of the main text and includes a control for whether the response came from the short version of the questionnaire. This table looks at selection effects on property characteristics. The characteristics labeled components of GARV are those that directly enter into the formula used to calculate GARV. Tax category (column (7)) is seven-tiered categorical variable with 7 being the most expensive tax bracket and 1 being the cheapest. Standard errors are clustered by robust partition, the partition of circles such that all circles that merged or split with each other are included within the same partition. * $p<.10$, ** $p<.05$, *** $p<0.01$.

Selection Effects on Reassessments

	(1) Approximate age of owner	(2) Owner's level of education	(3) Per capita wages	(4) Predicted expenditure given assets	(5) Connected to politician	(6) Connected to politician, government, police
Reassess * treatment	–0.348	–0.523*	–821.749	110.798	0.021*	0.005
	(0.799)	(0.317)	(1,078.191)	(213.234)	(0.012)	(0.027)
Reassess	–0.656*	0.303*	13.126	–94.529	–0.013**	0.005
	(0.398)	(0.157)	(510.006)	(122.380)	(0.006)	(0.014)
N	13,406	16,254	13,765	13,954	16,354	16,354
Mean of control group in gen. pop. sample	50.70	9.19	16,281.55	6,292.58	0.05	0.36

Khan et al 2016: Results

- Evidence broadly consistent with model:
  - Inspectors do reassess houses and bring in new ones
  - where reassessed, more likely to pay more tax, pay less bribes
  - Suggestive evidence it's richer, more powerful people who get reassessed
- Was this worthwhile for the benefit? We can do a cost-benefit assessment

Khan et al 2016: Cost-Benefit

Cost-Effectiveness of Incentives

	(1) Additional revenue	(2) Cost of incentives	(3) ROI
Panel A: Information in controls			
Any treatment	124,961,461	108,387,160	15.29
Revenue	50,578,024	37,349,784	35.42
Revenue plus	40,671,290	35,549,342	14.41
Flexible bonus	30,555,313	35,488,035	−13.90
Panel B: Information out of controls			
Any treatment	140,973,016	108,387,160	30.06
Revenue	56,269,064	37,349,784	50.65
Revenue plus	45,539,845	35,549,342	28.10
Flexible bonus	35,571,720	35,488,035	0.24

Outline

Financial Incentives

- Muralidharan & Sundararaman (JPE 2011) *Teacher Performance Pay: Experimental Evidence from India*
- Duflo, Hanna & Ryan (AER 2012) *Incentives Work: Getting Teachers to Come to School*
- Khan, Khwaja & Olken (QJE 2016) *Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors*
- Duflo Greenstone Pande & Ryan (QJE 2013): *Truth-telling by Third-party Auditors and the Response of Polluting Firms: Experimental Evidence from India*

Duflo et al 2013: Introduction

- Third-party audits are a common mechanism to monitor compliance with regulations
  - financial accounting of firms
  - food safety, health care, flowers etc in consumer & commodity markets
  - environmental regulation
- Usually, the auditor is chosen by, paid by, and reports to the audited firm.
- Creates a conflict of interest: Report the truth or report what is beneficial for your client?
- Conduct an experiment in India to investigate whether this can be overcome.

Duflo et al 2013: Setting

- Study conducted in Gujarat: 5% of India's population but 9% of manufacturing employment and 19% of output
- 3 of the 5 most polluted rivers in India are in Gujarat. Air pollution dangerous for human health in the cities.
- Gujarat Pollution Control Board (GPCB) has a stringent regulatory framework for pollution control.
- 2 Main instruments to enforce regulations
  1. Inspections (Duflo et al 2018 next week)
  2. third-party environmental audits.

Duflo et al 2013: Audits

- Plants with high pollution potential must submit a yearly environmental audit
- Auditors
  - visit each plant for about one day in each of three seasons of the year.
  - Observe environmental management practices and measure pollution.
  - Compile findings into a standardized format.
  - Submit report to the plant and to GPCB by February 15 of the following year.
- Auditors:
  - audit at most 15 plants a year and audit a plant at most 3 years in a row.
  - Auditors with inaccurate reports are liable to be decertified.
- Firms
  - no audit report $\rightarrow$ can have water & electricity cut off.
  - Reports showing noncompliance with regs $\rightarrow$ can be closed or fined (fines happen often)

Duflo et al 2013: Experiment

- Work with GPCB to implement a modified auditing system
- 663 small/medium scale plants in Ahmedabad & Surat selected
- Just before 2009 audits, randomly assigned half to audit treatment
- Then collect detailed data on each plant to determine eligibility $\rightarrow$ 473 plants, of which 49.2% in the treatment group.
- Treatment plants assigned to the audit treatment in year 1 (2009) and year 2 (2010)
- Notified by letter from GPCB that their audit rules had changed.

Duflo et al 2013: Treatment

- Treatment with 3 components

1. *Assignment and Fixed Pay*. Auditors randomly assigned to treatment plants and paid a flat fee (Rs. 45,000)
2. *Backchecks*. Randomly selected 20% of readings to be redone by technical staff from engineering colleges roughly 10 days after auditor's visit.
3. *Incentive Pay.* In year 2 added explicit incentive pay for auditor accuracy.
   3.1 Calculate δ_p: difference between audit and backcheck for pollutant p. Average it: $\Delta_{\text{Water}} = \sum_{p \in \text{Water}} \delta_p$, $\Delta_{\text{Air}} = \sum_{p \in \text{Air}} \delta_p$, $\Delta_{\text{All}} = (\Delta_{\text{Water}} + \Delta_{\text{Air}})/2$
   3.2 Rank auditors: Least accurate quartile get Rs. 35,000 per audit. Next least accurate quartile get Rs. 40,000. Most accurate half get Rs. 52,500 (NB average pay unchanged)

Submission of Audit Reports

	(1) Treatment	(2) Control	(3) Difference
Panel A: 2009			
Audit submitted	163	177	
Total plants	233	240	
Share submitted	0.70	0.74	−0.038
			(0.041)
Panel B: 2010			
Audit submitted	164	153	
Total plants	233	240	
Share submitted	0.70	0.64	0.066
			(0.043)

	(1) Treatment	(2) Control	(3) Difference
Panel A: Plant characteristics			
Capital investment INR 50 m to 100 m (= 1)	0.092	0.14	−0.051
	[0.29]	[0.35]	(0.033)
Located in industrial estate (= 1)	0.57	0.53	0.042
	[0.50]	[0.50]	(0.051)
Textiles (= 1)	0.88	0.93	−0.030
	[0.33]	[0.26]	(0.025)
Effluent to common treatment (= 1)	0.41	0.35	0.078
	[0.49]	[0.48]	(0.049)
Wastewater generated (kl/day)	420.5	394.6	35.4
	[315.9]	[323.4]	(31.6)
Lignite used as fuel (= 1)	0.71	0.77	−0.024
	[0.45]	[0.42]	(0.029)
Diesel used as fuel (= 1)	0.29	0.25	0.038
	[0.45]	[0.43]	(0.046)
Air emissions from flue gas (= 1)	0.85	0.87	−0.0095
	[0.35]	[0.33]	(0.016)
Air emissions from boiler (= 1)	0.93	0.92	0.026
	[0.26]	[0.27]	(0.027)
Bag filter installed (= 1)	0.24	0.34	−0.10**
	[0.43]	[0.47]	(0.046)
Cyclone installed (= 1)	0.087	0.079	0.0010
	[0.28]	[0.27]	(0.027)
Scrubber installed (= 1)	0.41	0.41	−0.018
	[0.49]	[0.49]	(0.050)

Panel B: Regulatory interactions in year prior to study			
Whether audit submitted (= 1)	0.82	0.81	0.022
	[0.38]	[0.39]	(0.038)
Any equipment mandated (= 1)	0.42	0.49	–0.047
	[0.50]	[0.50]	(0.047)
Any inspection conducted (= 1)	0.79	0.78	0.016
	[0.41]	[0.42]	(0.042)
Any citation issued (= 1)	0.28	0.24	0.035
	[0.45]	[0.43]	(0.045)
Any water citation issued (= 1)	0.12	0.12	–0.0031
	[0.33]	[0.33]	(0.034)
Any air citation issued (= 1)	0.027	0.0052	0.021*
	[0.16]	[0.072]	(0.013)
Any utility disconnection (= 1)	0.098	0.094	0.0029
	[0.30]	[0.29]	(0.031)
Any bank guarantee posted (= 1)	0.033	0.026	0.0045
	[0.18]	[0.16]	(0.017)
Observations	184	191	

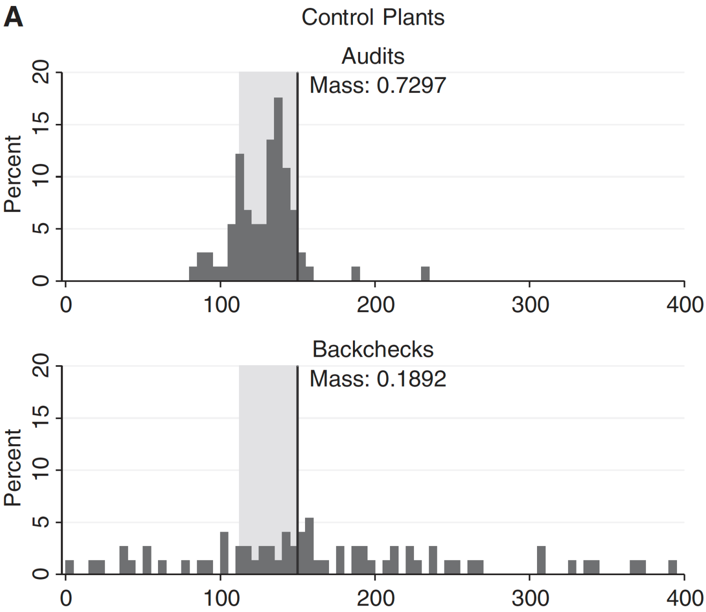
Duflo et al 2013: Misreporting in Control Group

- Do auditors misreport?

1. Compare distribution of pollution scores in audit reports to backcheck data
2. Regression version. Stack data from audit and backchecks and run

$$\mathbf{1}\left\{\text{Compliant}\right\}_{ij} = \beta_1 \mathbf{1}\left\{\text{AuditReport}\right\} + \alpha_r + \epsilon_{ij}$$

where $\mathbf{1}\left\{\text{Compliant}\right\}_{ij}$ denotes pollutant i at plant j being between 75% and 100% of the regulatory standard.


A
Control Plants
Audits
Mass: 0.7297
Percent
0
5
10
15
20
0
100
200
300
400
Backchecks
Mass: 0.1892
Percent
0
5
10
15
20
0
100
200
300
400

Compliance in Audits Relative to Backchecks, Control Group Only

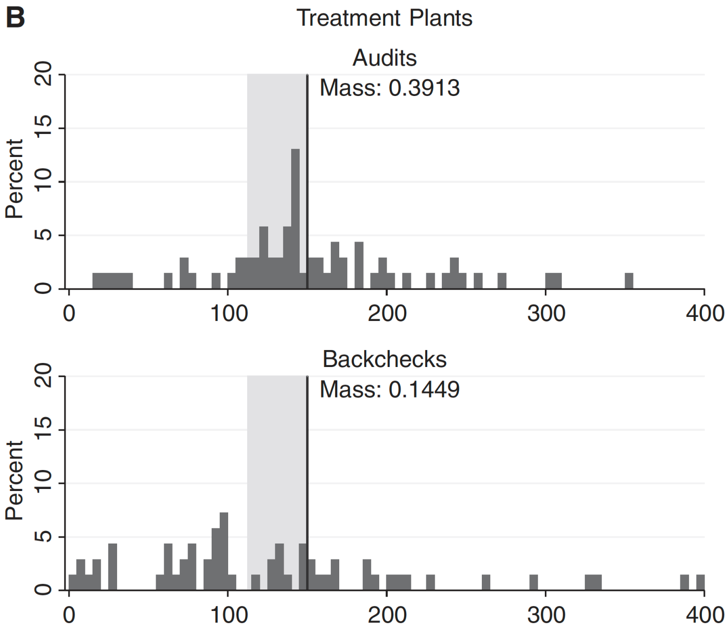
	(1) All pollutants	(2) Water pollutants	(3) Air pollutants
Panel A: Dependent variable: Narrow compliance (dummy for pollutant between 75% and 100% of regulatory standard)			
Audit report (= 1)	0.270***	0.297***	0.230***
	(0.025)	(0.034)	(0.033)
Control mean in backchecks	0.097	0.110	0.077
Panel B: Dependent variable: Compliance (dummy for pollutant at or below regulatory standard)			
Audit report (= 1)	0.288***	0.273***	0.311***
	(0.023)	(0.033)	(0.032)
Control mean in backchecks	0.557	0.538	0.586
Observations	1132	688	444

Duflo et al 2013: Treatment Effects on Misreporting

- Does the treatment improve auditors' reporting?

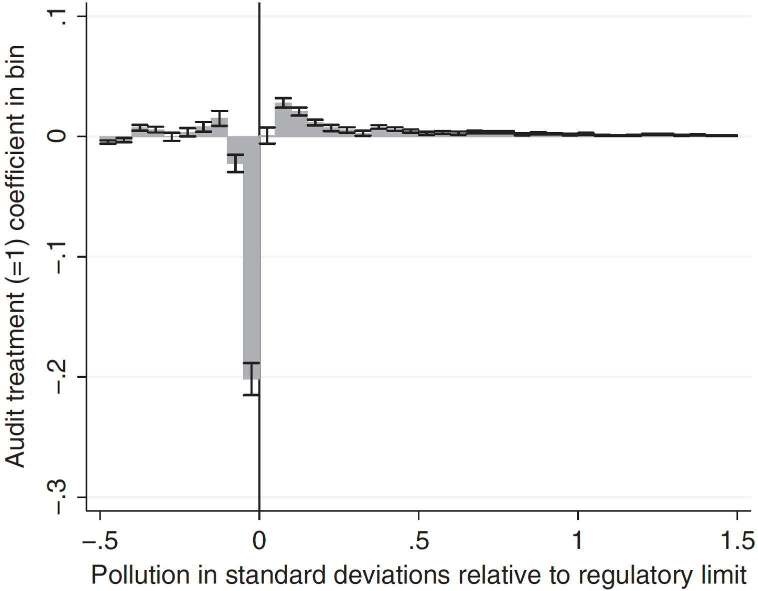
1. Compare distribution of pollution scores in audit reports and backcheck data
2. Run difference in difference regression

$$\mathbf{1}\left\{\text{Compliant}\right\}_{ij} = \beta_1 \mathbf{1}\left\{\text{AuditReport}\right\} \times T_j + \beta_2 \mathbf{1}\left\{\text{AuditReport}\right\} + \beta_3 T_j + \alpha_r + \epsilon_{ij}$$


B
Treatment Plants
Audits
Mass: 0.3913
Percent
0
5
10
15
20
0
100
200
300
400
Backchecks
Mass: 0.1449
Percent
0
5
10
15
20
0
100
200
300
400

Compliance in Audits Relative to Backchecks by Treatment Status

	(1) All pollutants	(2) Water pollutants	(3) Air pollutants
Panel A: Dependent variable: Narrow compliance (dummy for pollutant between 75% and 100% of regulatory standard)			
Audit report × Treatment group	−0.185***	−0.212***	−0.143***
	(0.034)	(0.044)	(0.046)
Audit report (= 1)	0.270***	0.297***	0.230***
	(0.025)	(0.034)	(0.033)
Treatment group (= 1)	−0.0034	−0.013	0.011
	(0.0176)	(0.025)	(0.024)
Control mean in backchecks	0.097	0.110	0.077
Panel B: Dependent variable: Compliance (dummy for pollutant at or below regulatory standard)			
Audit report × Treatment group	−0.234***	−0.166***	−0.345***
	(0.039)	(0.050)	(0.056)
Audit report (= 1)	0.288***	0.273***	0.311***
	(0.023)	(0.033)	(0.032)
Treatment group (= 1)	0.058*	0.0075	0.145***
	(0.034)	(0.0477)	(0.041)
Control mean in backchecks	0.557	0.538	0.586
Observations	2236	1378	858


Audit treatment (=1) coefficient in bin
.1
0
−.1
−.2
−.3
−.5
0
.5
1
1.5
Pollution in standard deviations relative to regulatory limit

Duflo et al 2013: Plant Responses

- Do plants reduce pollution now that the auditors tell the truth?
- In endline survey (after year 2) gather new pollution measurements from all plants, not just those who submit audit reports
- Estimate effects on pollution

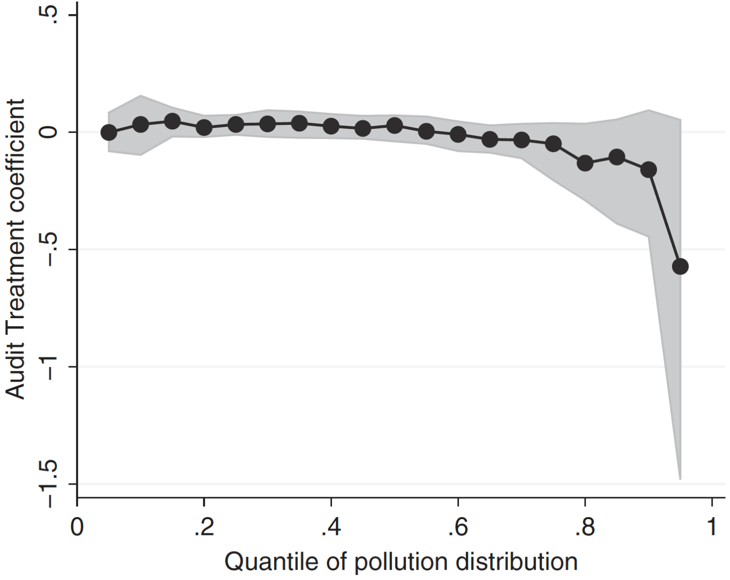
$$y_{ij} = \alpha_r + \beta T_j + \epsilon_{ij}$$

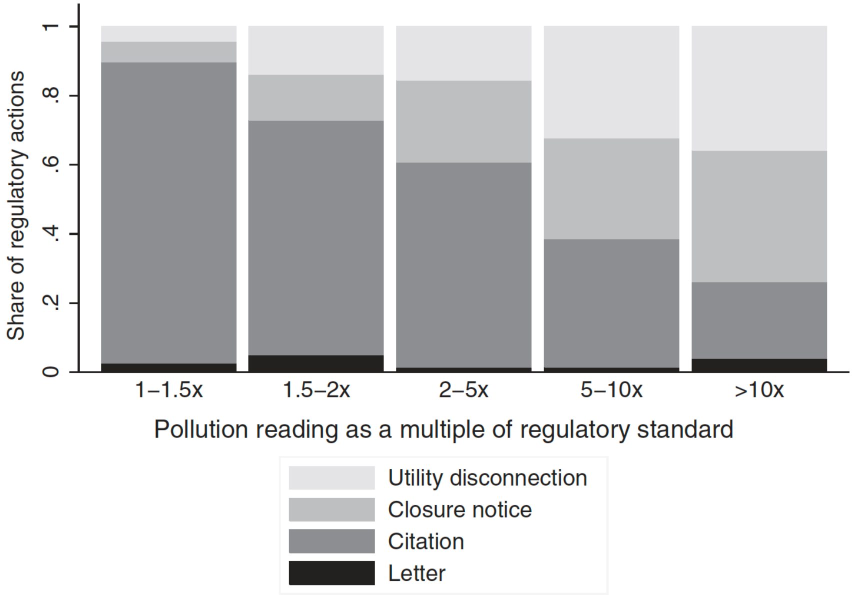
- Also estimate quantile regressions:

$$Q_{y_{ij}|X_j}(\tau) = \beta T_j + \alpha_r + \epsilon_{ij}$$

ENDLINE POLLUTANT CONCENTRATIONS ON TREATMENT STATUS

	(1) All pollutants	(2) Water pollutants	(3) Air pollutants
Panel A: Dependent variable: Level of pollutant in endline survey, all pollutants (standard deviations relative to backcheck mean)			
Audit treatment assigned (= 1)	−0.211**	−0.300*	−0.053
	(0.099)	(0.159)	(0.057)
Control mean	0.076	0.114	0.022
Observations	1439	860	579
Panel B: Dependent variable: Compliance (dummy for pollutant in endline survey at or below regulatory standard)			
Audit treatment assigned (=1)	0.027	0.039	0.002
	(0.027)	(0.039)	(0.028)
Control mean	0.573	0.516	0.656
Observations	1,439	860	579


Audit Treatment coefficient
.5
0
−.5
−1
−1.5
0
.2
.4
.6
.8
1
Quantile of pollution distribution


Share of regulatory actions
0
.2
.4
.6
.8
1
1–1.5x
1.5–2x
2–5x
5–10x
>10x
Pollution reading as a multiple of regulatory standard
Utility disconnection
Closure notice
Citation
Letter

Outline

Theory

Financial Incentives

Non-financial Incentives

Recruitment & Selection

Open Questions

Outline

Non-financial Incentives

- Bandiera, Best, Khan & Prat (WP 2020): *The Allocation of Authority in Organizations: A Field Experiment with Bureaucrats*
- Duflo, Greenstone, Pande & Ryan (WP 2017): *The Value of Regulatory Discretion: Estimates from Environmental Inspections in India*
- Khan, Khwaja & Olken (WP 2018) *Making Moves Matter: Experimental Evidence on Incentivizing Bureaucrats through Performance-Based Postings*

Agency and authority

- Organizations bring together people with different interests, information and skills to work towards a common goal.
- Incentive contracts designed to meet that goal are at the core of organisational economics
- Mostly this takes as given the allocation of authority, that is the actions over which the agents has decision rights

- But the allocation of authority is itself a choice variable
- Observed solutions range from total control to total freedom

Costs and benefits of autonomy

- Giving decision rights to agents incentivises them to acquire information and use it to the benefit of the organisation *but also to their private benefit*
- Red tape/rules that limit decision rights limit the agents' capture *but stifle initiative and creates compliance costs (time to tick boxes)*
- Shifting agency
  - in most organisations rules require an additional layer of agents to enforce them
  - stricter rules effectively shift authority from one set of agents to another, whose incentives need not be better aligned

What we know

- A number of studies show that cross-sectional variation is consistent with the predictions of agency theory (Acemoglu et al 07, Bloom et al 12)
- Yet, there is a positive correlation between autonomy and performance
  - across organisations (Bandiera, Prat and Valletti 09, Rasul&Rogger 18, Rasul et al 19)
  - within organisations (Duflo et al 17, Wu 18)

This paper

- *Design a field experiment to exogenously shift authority to lower tier bureaucrats*
- Cross-randomise with performance pay to study the interaction between the two

Setting: Public Procurement

- Textbook example of moral hazard:
  - Agent buys goods she won't use with money she doesn't own
  - Misalignment of interests $\Rightarrow$ low effort and/or corruption

Setting: Public Procurement

- Textbook example of moral hazard:
  - Agent buys goods she won't use with money she doesn't own
  - Misalignment of interests $\Rightarrow$ low effort and/or corruption
- Stakes are high
  - Spending on public procurement as GDP share in 2015 (OECD):
    - United States: 9.35%
    - Average OECD country: 13.18%
  - Potential for large savings (Olken and Pande 2012)

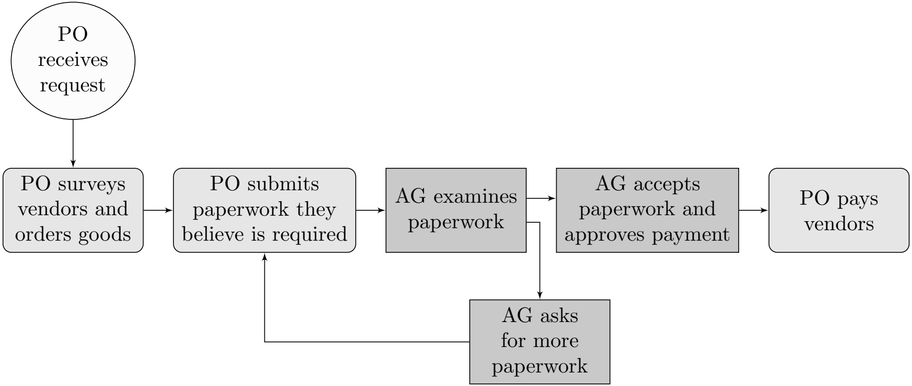
Policy Debate

- Strict auditing and monitoring is necessary to avoid private capture..
  - OECD Third Principle for Integrity in Public Procurement: “The management of public funds should be monitored by internal control and internal audit bodies.”
- ..but red tape stifles inititative and creates compliance costs.
  - Kelman (1990), Procurement and Public Management: The Fear of Discretion and the Quality of Government Performance

Procurement in Punjab, Pakistan: Agents and Supervisors

- Legal authority for public procurement is vested in **Procurement Officers** (POs)
- POs are allocated budget under different accounting heads (salary, repairs, etc.), including procurement, by the Finance Department
- An independent federal agency - office of the **Accountant General** (AG) needs to approve POs purchases *before* payment can be made
- We focus on procurement of *generic goods*

The Procurement Process for Generic Goods



Model: Agents and Supervisors

- define c the minimal price at which the supplier is willing to sell
- under autonomy, the price paid depends only on the PO's alignment $p_M > p_A = c$.
- under monitoring, the outcome depends on the type of both agents: better types reduce the purchase price: $p_{MM} > p_{AM} > p_{AA}$
- denote the share of misaligned POs (AGs) by θ_{PO} (θ_{AG})

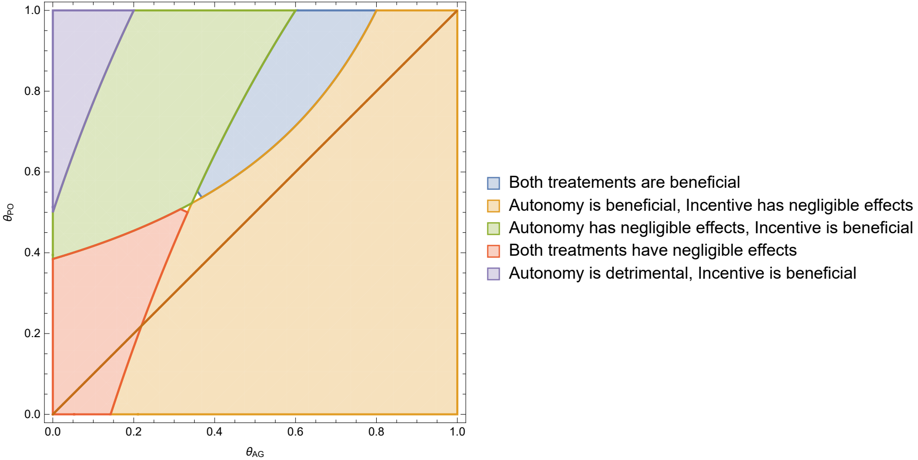
The effect of autonomy

- The effect of the autonomy treatment is

$$\Delta_A = \theta_{PO}\theta_{AG}\left(p_M - p_{MM}\right) + \theta_{PO}\left(1-\theta_{AG}\right)\left(p_M - p_{MA}\right) \\ + \left(1-\theta_{PO}\right)\theta_{AG}\left(c - p_{AM}\right)$$

- for a share $\theta_{PO}\theta_{AG}$ of PO -AG pairs autonomy eliminates “double bandit” ($p_M - p_{MM} < 0$)
- for $\left(1-\theta_{PO}\right)\theta_{AG}$ autonomy gives decision making rights to the most aligned agent ($c - p_{AM} < 0$).
- for $\theta_{PO}\left(1-\theta_{AG}\right)$ autonomy gives decision making rights to the least aligned agent ($p_M - p_{MA} > 0$),

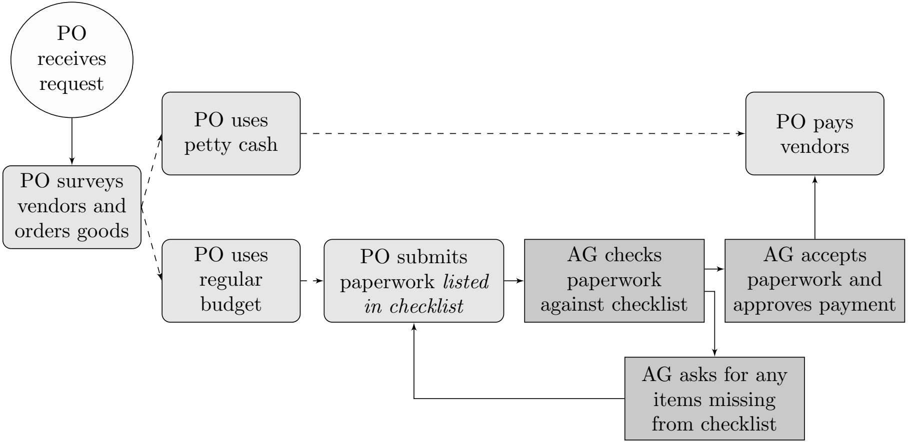
Treatment effects, in theory



Design: Autonomy Treatment

- Focus groups with POs in 2013/14 highlights two key issues
  - many vendors cannot wait through the process of approval
  - finding out appropriate documentation is difficult and time consuming
- Autonomy treatment addresses both by
  - Limit AG discretion on documents to require for audit
  - Cash in hand allow POs to pay vendors without AG pre-approval
    - Rs 100,000 (USD 1k) , 10% of budget; median (mean) purchase value 1,000 (3,062)

Procurement Process Under Autonomy Treatment



Design: Incentive Treatment

- 3 prizes based on value for money in all generics
- awarded by a commission made of: senior private sector auditor & head PPRA (co-chair), representatives of all departments (10 members)
- data on quality adjusted prices provided by us

1. "gold": 2 months wages, to the top 7.5%
2. "silver": 1 month wages, to the next 22.5%
3. "bronze": 0.5 month wages, to the next 45%
4. nothing to remaining 25%

Sample

- 688 Procurement Officers in charge of procurement of 778 Public Bodies (88% in charge of 1 PB, 10% 2, 2% 3 or more)
  - take-up 85% –> sample contains 587 POs
- 4 Departments (chosen by number of POs with generics budget
- 26 Districts (out of 36) - cover over 80% of the population (110 million)
- AG has offices in each district, responsible for POs in that district

Design: Randomization

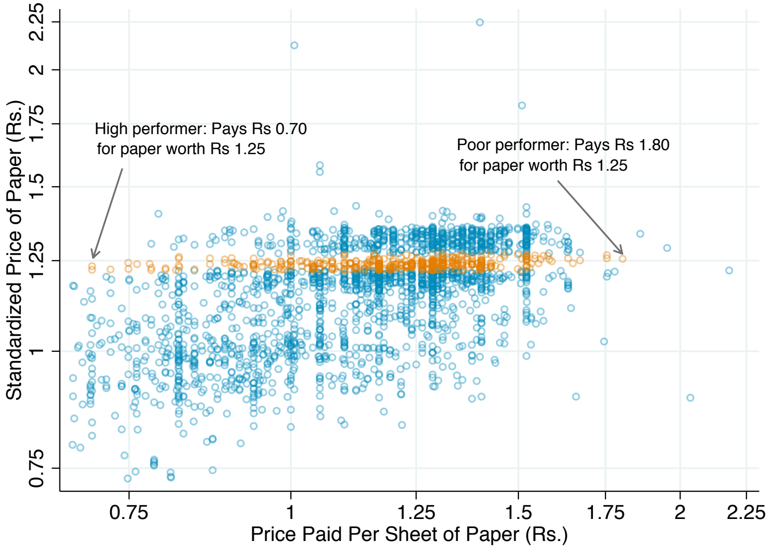
- Randomise POs in four groups, stratified by district and department
- Follow over two years, stagger the introduction

	year 1	year 2
A	none	autonomy
B	p4p	p4p+autonomy
I	p4p	p4p
C	none	none

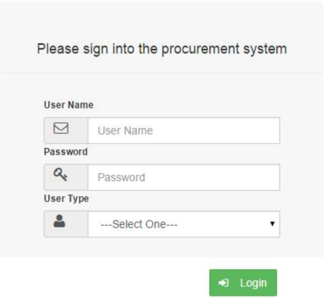
Measuring Value for Money

- Our sample covers all the generic (off-the-shelf) goods that POs buy (e.g. printer paper, soap..)
- Bought by many consumers and produced by several suppliers
- Measurable (with some effort!) and comparable performance
- Most are sold in competitive markets, so everybody should pay the same price. And yet. . . .

POs Pay Different Prices for Exactly the Same Good

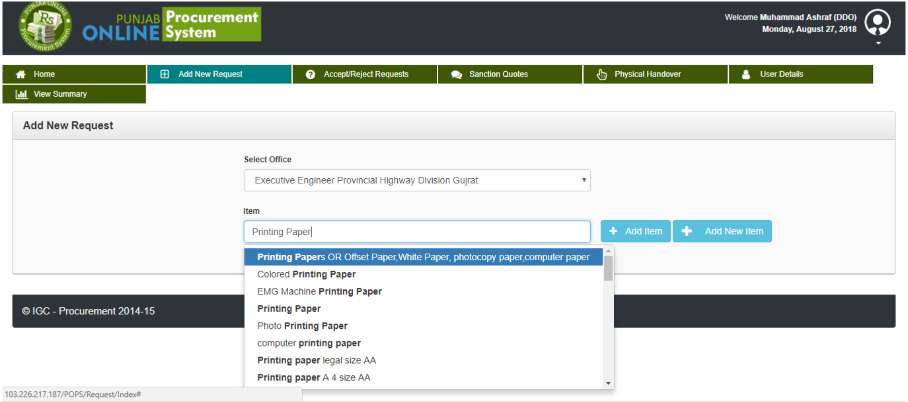


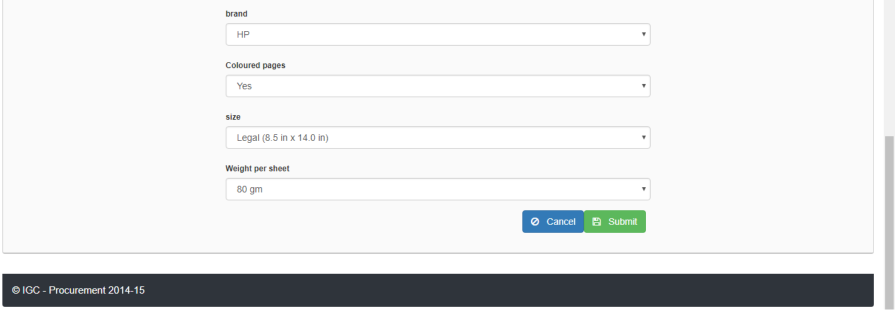
Together with Punjab Procurement Regulatory Authority (PPRA) and Punjab Information Technology Board (PITB), we set up an E-Governance platform: Punjab Online Procurement System (POPS)



POPS Collects Detailed Spending Data

- Through POPS, office staff enter detailed data on what they are buying

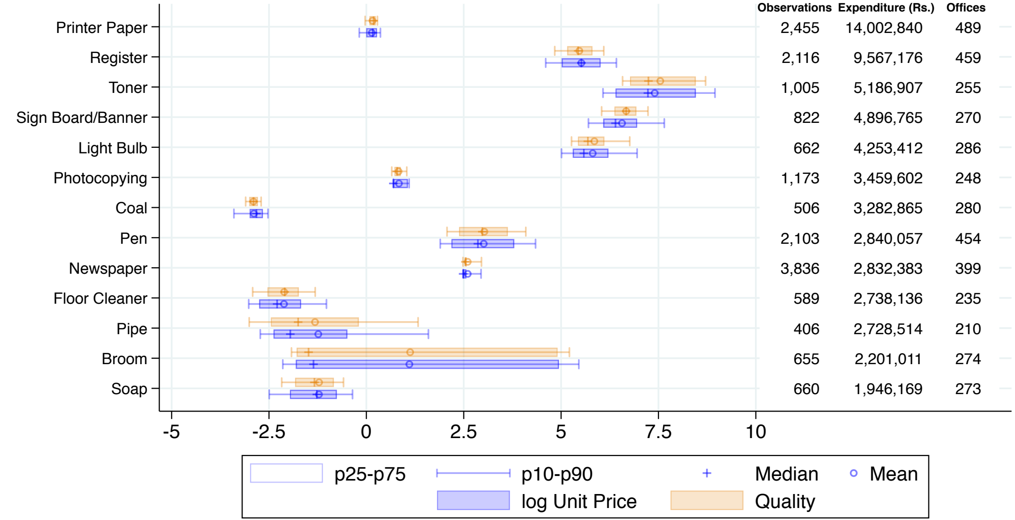



brand
HP
Coloured pages
Yes
size
Legal (8.5 in x 14.0 in)
Weight per sheet
80 gm
Cancel
Submit
© IGC - Procurement 2014-15

Item Selection

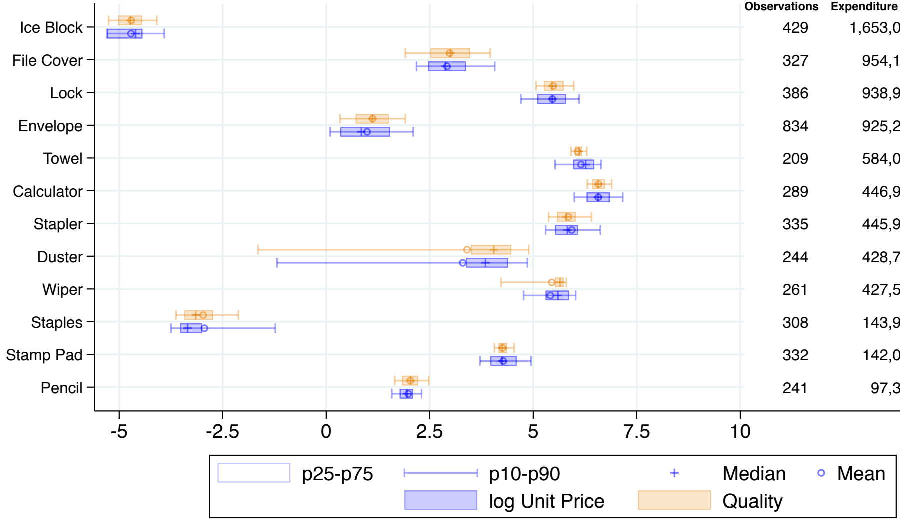
- Generics account for 52% of the non wage budget for the average office
  - mean \$94k, median \$27k
- We rank 100+ distinct goods by unique number of POs buying them –> choose top 25 goods
- 21,503 purchases (total value \$7.7m) that account for 50% of generics budget
- the least common item (pencil) is bought by 25% of the sample POs
- the most common (paper) by 85%

25 goods, 21,503 purchases



	Observations	Expenditure (Rs.)	Offices
Printer Paper	2,455	14,002,840	489
Register	2,116	9,567,176	459
Toner	1,005	5,186,907	255
Sign Board/Banner	822	4,896,765	270
Light Bulb	662	4,253,412	286
Photocopying	1,173	3,459,602	248
Coal	506	3,282,865	280
Pen	2,103	2,840,057	454
Newspaper	3,836	2,832,383	399
Floor Cleaner	589	2,738,136	235
Pipe	406	2,728,514	210
Broom	655	2,201,011	274
Soap	660	1,946,169	273

25 goods, 21,503 purchases



	Observations	Expenditure (Rs.)	Offices
Ice Block	429	1,653,095	219
File Cover	327	954,161	184
Lock	386	938,951	191
Envelope	834	925,219	276
Towel	209	584,031	157
Calculator	289	446,954	201
Stapler	335	445,942	230
Duster	244	428,761	168
Wiper	261	427,569	173
Staples	308	143,918	191
Stamp Pad	332	142,028	216
Pencil	241	97,378	141

Treatment Effects I

- We estimate

$$p_{igto} = \alpha + \sum_{k=1}^{3} \eta_k \text{Treatment}_o^k + \beta v_{igto} + \rho_g q_{igto}$$
$$+ \delta_s \text{Department}_o \times \text{District}_o + \delta_g + \varepsilon_{igto}$$

  - p_{igto} is log unit price,
  - q_{igto} (log) quantity purchased,
  - v_{igto} is good "variety" i.e. all the attributes that might affect price,
  - δ_s, δ_g stratum, good FEs.
  - Weight by control expenditure shares,
  - cluster ε_{igto} by public body.

Identification

$$p_{igto} = \alpha + \sum_{k=1}^{3} \eta_k \text{Treatment}_o^k + \beta v_{igto} + \rho_g q_{igto}$$
$$+ \delta_s \text{Department}_o \times \text{District}_o + \delta_g + \varepsilon_{igto}$$

- η_k: causal effect of treatment k on quality-adjusted prices if

1. treatment does not affect control POs e.g. through AG.
   - Experimental POs are a small fraction of total POs supervised
   - Effect on prices paid by control DDOs does not depend on number of treated in same office
2. quality not affected by treatment.
   - no effect on quality, very similar Diff in Diff results
3. v_{igto} adequately captures item variety

Measuring Item Variety

3. v_{igto} adequately captures item variety

- Method 1: control for all goods' attributes in price regression (*fine* measure)
- Method 1b: aggregate by "pricing" attributes in control group. (*scalar* measure)

$$p_{igto} = \mathbf{X}_{igto}\lambda_g + \rho_g q_{igto} + \delta_g + \varepsilon_{igto}$$

q_{igto} is the quantity purchased, $\mathbf{X}_{igto}$ are attributes of the item

  - Use $\hat{\lambda}$s to control for quality $\Rightarrow v_{igto} = \sum_{j \in A(g)} \hat{\lambda}_j X_j$ where $A(g)$ is the set of attributes of good g
- Method 2: control for simpler measure of quality (*coarse* measure)
  - Use attributes with large $\hat{\lambda}$s to classify purchases into "high" or "low" quality
- Method 3: Use ML to a) find optimal coarseness; b) allow more nonlinearity

No Effects on Item Variety Purchased

	(1)	(2)
Autonomy	0.016	0.010
	(0.030)	(0.023)
	[0.646]	[0.705]
Item Type Control	Scalar	Coarse
p(All = 0)	0.660	0.079
p(Autonomy = Incentives)	0.749	0.537
p(Both = Autonomy + Incentives)	0.722	0.510
Observations	11,771	11,771

No Effects on Item Variety Purchased

	(1)	(2)
Autonomy	0.016	0.010
	(0.030)	(0.023)
	[0.646]	[0.705]
Incentives	0.006	0.025
	(0.030)	(0.023)
	[0.846]	[0.325]
Item Type Control	Scalar	Coarse
p(All = 0)	0.660	0.079
p(Autonomy = Incentives)	0.749	0.537
p(Both = Autonomy + Incentives)	0.722	0.510
Observations	11,771	11,771

No Effects on Item Variety Purchased

	(1)	(2)
Autonomy	0.016	0.010
	(0.030)	(0.023)
	[0.646]	[0.705]
Incentives	0.006	0.025
	(0.030)	(0.023)
	[0.846]	[0.325]
Autonomy & Incentives	0.037	0.059
	(0.030)	(0.023)
	[0.265]	[0.021]
Item Type Control	Scalar	Coarse
p(All = 0)	0.660	0.079
p(Autonomy = Incentives)	0.749	0.537
p(Both = Autonomy + Incentives)	0.722	0.510
Observations	11,771	11,771

No Effects on Composition

Item	Treatment Effect Autonomy	Incentives	Both	Joint Test All = 0
Toner	335.1 (4181.24)	1322.1 (4205.91)	243.0 (4201.91)	0.04 [0.989]
Ice Block	-71.5 (294.11)	-562.3* (295.85)	-160.3 (295.57)	1.49 [0.215]
Towel	-178.0 (172.48)	42.6 (173.50)	-204.2 (173.34)	1.08 [0.359]
Soap/Detergent	-4509.9 (9670.41)	92.5 (9727.48)	5128.5 (9718.23)	0.35 [0.788]
Duster	-190.8 (176.66)	233.1 (177.71)	-228.0 (177.54)	3.01 [0.030]
Wiper	-13.4 (136.22)	288.1** (137.02)	-100.2 (136.89)	3.22 [0.022]
Lock	90.7 (278.17)	129.0 (279.81)	-234.7 (279.55)	0.73 [0.537]
Pen	790.6 (700.16)	997.3 (704.29)	281.6 (703.62)	0.86 [0.463]
Envelope	214.7 (206.51)	-73.0 (207.73)	-105.8 (207.53)	1.03 [0.377]
Printer Paper	2432.1 (2778.08)	3227.7 (2794.47)	-1890.5 (2791.81)	1.48 [0.218]
Register	-2915.2 (5939.89)	-966.6 (5974.94)	976.1 (5969.26)	0.17 [0.920]
Stapler	-159.9 (109.74)	-131.2 (110.39)	-183.7* (110.29)	1.08 [0.356]
Staples	-15.6 (39.76)	-1.3 (39.99)	18.1 (39.95)	0.26 [0.857]

Item	Treatment Effect Autonomy	Incentives	Both	Joint Test All = 0
Calculator	-131.1 (104.43)	-158.2 (105.05)	-177.8* (104.95)	1.14 [0.333]
File Cover	389.7 (391.77)	-423.0 (394.08)	145.2 (393.71)	1.60 [0.189]
Stamp Pad	81.8 (57.32)	75.7 (57.66)	-17.7 (57.61)	1.66 [0.175]
Photocopying	337.1 (975.53)	693.7 (981.29)	943.2 (980.36)	0.36 [0.785]
Broom	655.0 (825.63)	1158.4 (830.50)	477.6 (829.71)	0.67 [0.572]
Coal	-283.8 (833.45)	745.7 (838.37)	972.5 (837.57)	1.07 [0.362]
Newspaper	329.4 (551.96)	-69.2 (555.22)	39.3 (554.69)	0.21 [0.887]
Pipe	626.5 (591.92)	1130.2* (595.42)	238.5 (594.85)	1.41 [0.240]
Light Bulb	959.2 (1472.59)	-585.2 (1481.28)	-25.0 (1479.87)	0.40 [0.752]
Pencil	90.4 (62.45)	0.4 (62.82)	-37.7 (62.76)	1.61 [0.185]
Floor Cleaner	-242.3 (660.41)	-47.6 (664.31)	216.6 (663.68)	0.17 [0.915]
Sign Board/Banner	1757.6 (2236.96)	292.4 (2250.16)	461.7 (2248.02)	0.25 [0.861]
Joint F-Test	0.89 [0.619]	1.28 [0.168]	0.86 [0.656]	1.06 [0.357]

Autonomy reduces prices up to 9%

	(1)	(2)	(3)	(4)
Autonomy	-0.085	-0.086	-0.080	-0.082
	(0.038)	(0.032)	(0.031)	(0.034)
	[0.046]	[0.018]	[0.023]	[0.030]
Item Type Control	None	Attribs	Scalar	Coarse
p(All = 0)	0.168	0.053	0.092	0.086
p(Autonomy = Incentives)	0.146	0.077	0.119	0.119
p(Both = Autonomy + Incentives)	0.620	0.552	0.562	0.799
Observations	11,771	11,771	11,771	11,771

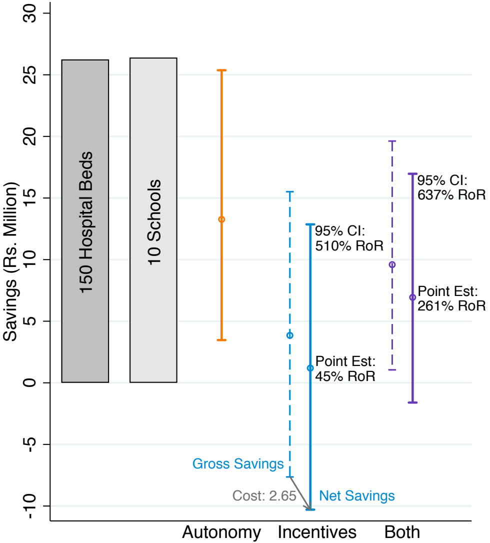
Autonomy reduces prices up to 9%

	(1)	(2)	(3)	(4)
Autonomy	-0.085	-0.086	-0.080	-0.082
	(0.038)	(0.032)	(0.031)	(0.034)
	[0.046]	[0.018]	[0.023]	[0.030]
Incentives	-0.016	-0.026	-0.022	-0.020
	(0.038)	(0.030)	(0.033)	(0.034)
	[0.723]	[0.476]	[0.571]	[0.625]
Item Type Control	None	Attribs	Scalar	Coarse
p(All = 0)	0.168	0.053	0.092	0.086
p(Autonomy = Incentives)	0.146	0.077	0.119	0.119
p(Both = Autonomy + Incentives)	0.620	0.552	0.562	0.799
Observations	11,771	11,771	11,771	11,771

Autonomy reduces prices up to 9%

	(1)	(2)	(3)	(4)
Autonomy	-0.085	-0.086	-0.080	-0.082
	(0.038)	(0.032)	(0.031)	(0.034)
	[0.046]	[0.018]	[0.023]	[0.030]
Incentives	-0.016	-0.026	-0.022	-0.020
	(0.038)	(0.030)	(0.033)	(0.034)
	[0.723]	[0.476]	[0.571]	[0.625]
Autonomy & Incentives	-0.070	-0.083	-0.072	-0.086
	(0.041)	(0.032)	(0.033)	(0.039)
	[0.130]	[0.025]	[0.053]	[0.043]
Item Type Control	None	Attribs	Scalar	Coarse
p(All = 0)	0.168	0.053	0.092	0.086
p(Autonomy = Incentives)	0.146	0.077	0.119	0.119
p(Both = Autonomy + Incentives)	0.620	0.552	0.562	0.799
Observations	11,771	11,771	11,771	11,771

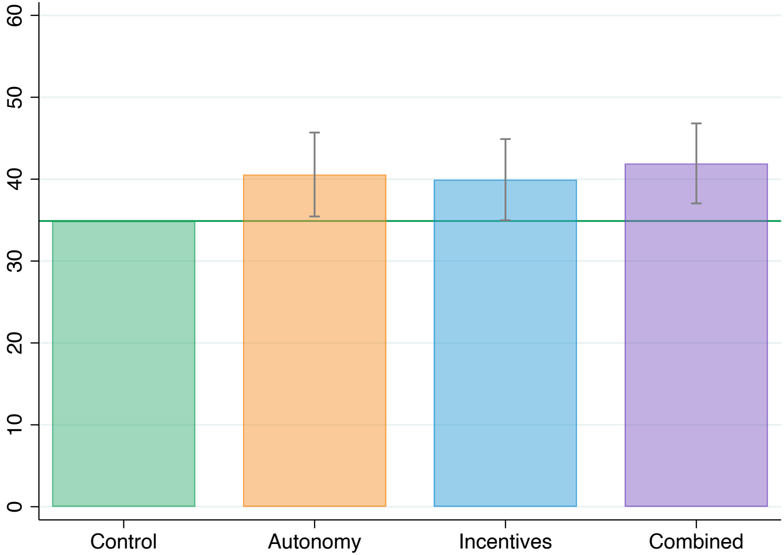
Benchmarking



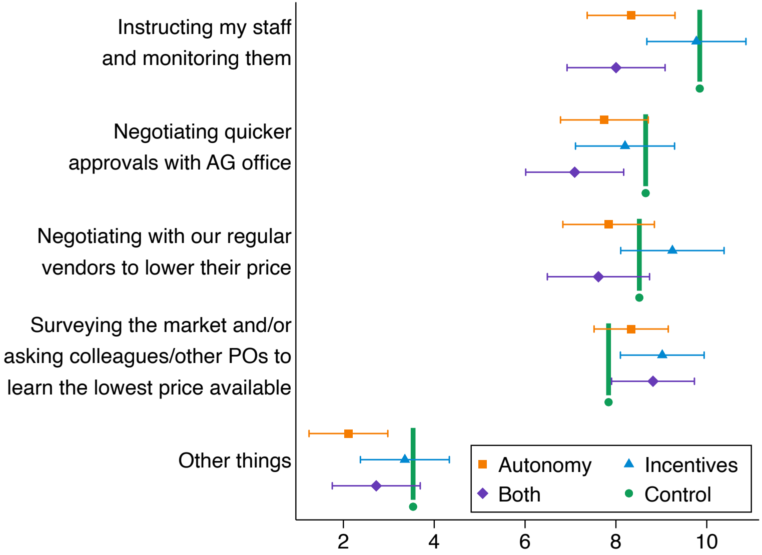
Open Questions

- We found that autonomy reduce prices, leave quantity, quality and the composition of purchases unchanged
- Effect at least as large as incentives' on their own, at no cost
- No additional effect of incentives
- Why?

Treated POs devote more time to procurement



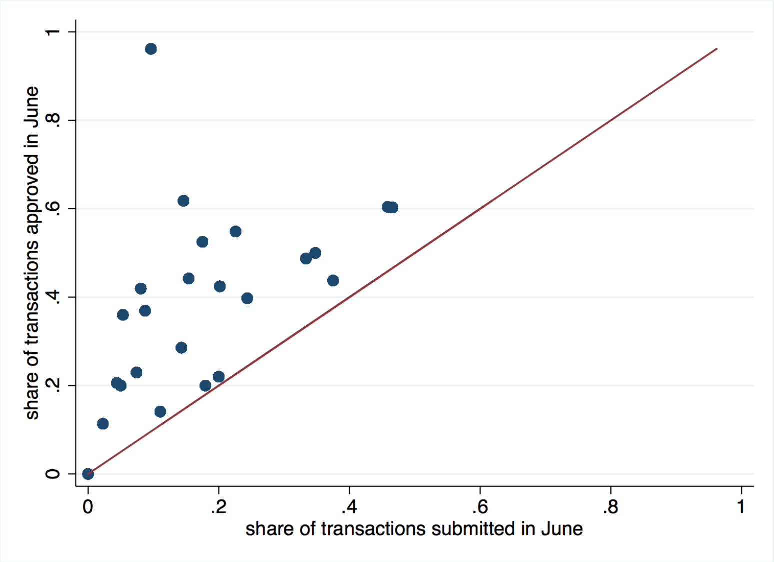
They spend more time searching deals, less time with the AG



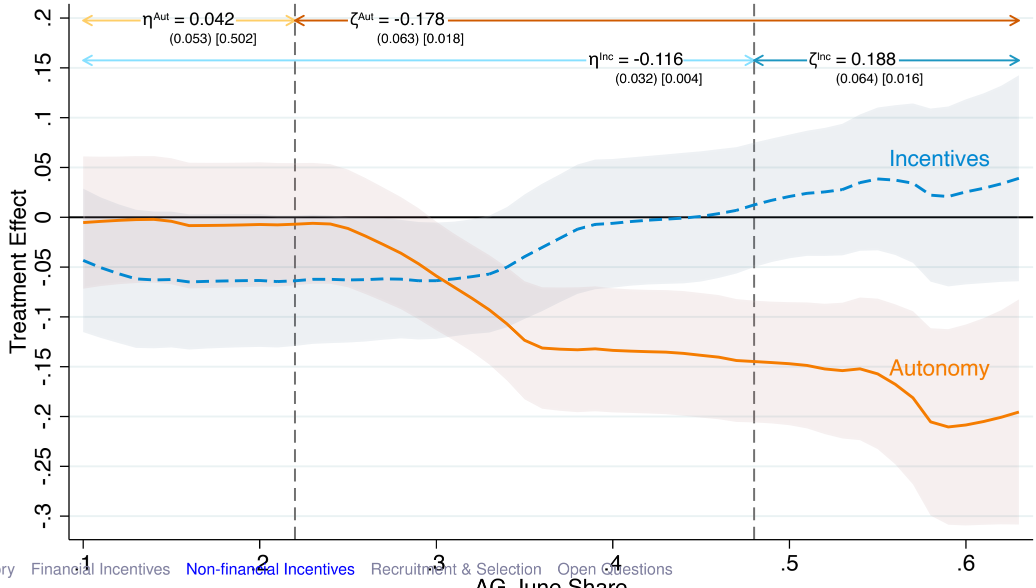
The accountant general

- Autonomy to PO= less power to AG
- Theory shows that shifting authority will lower prices if the AG is relatively more inefficient or corrupt whilst incentives will only lower prices in the opposite case
- AG are district specific –> use control group before the experiment to compute share of transactions approved at the FYE (Liebman and Mahoney 16)

Identifying variation



Heterogeneity of Treatment Effects by Monitor Alignment



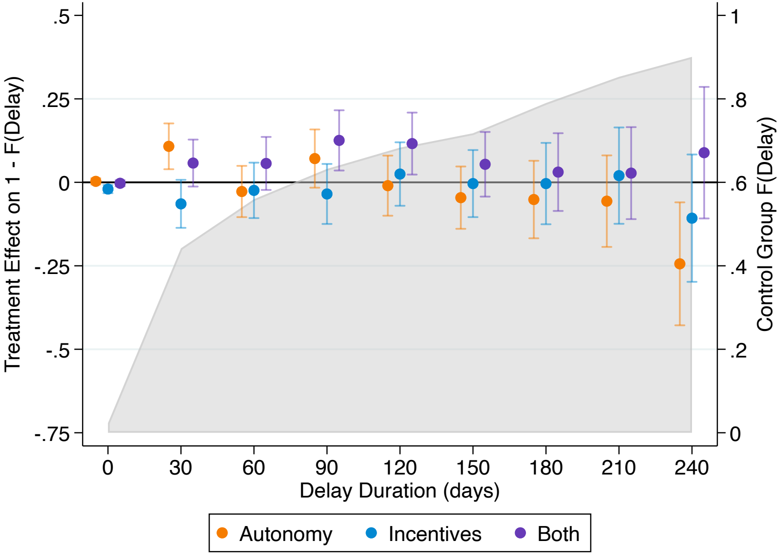
POs respond to incentives when the AG is good

	Overall			Good AG			Bad AG		
Panel B: Incentives									
	(1) OLS	(2) IV	(3) 1st Stage	(4) OLS	(5) IV	(6) 1st Stage	(7) OLS	(8) IV	(9) 1st Stage
Time Spent on Procurement	0.001	-0.008		0.001	-0.010		-0.000	-0.161	
	(0.001)	(0.005)		(0.001)	(0.005)		(0.002)	(2.053)	
		[0.011]			[0.001]			[0.903]	
Incentives			6.132			8.335			-0.330
			(2.165)			(2.464)			(4.338)
			[0.013]			[0.004]			[0.950]
Observations	8,556	8,556	8,556	6,355	6,355	6,355	2,201	2,201	2,201

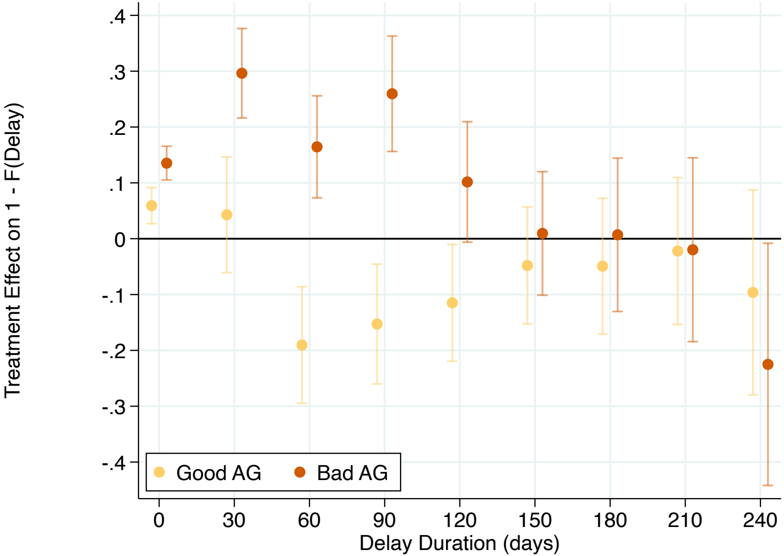
POs respond to autonomy when the AG is bad

Panel A: Autonomy									
	Overall			Good AG			Bad AG		
	(1) OLS	(2) IV	(3) 1st Stage	(4) OLS	(5) IV	(6) 1st Stage	(7) OLS	(8) IV	(9) 1st Stage
Time Spent on Procurement	-0.001	-0.006		-0.001	0.014		-0.001	-0.012	
	(0.001)	(0.005)		(0.002)	(0.022)		(0.001)	(0.006)	
		[0.022]			[0.304]			[0.001]	
Autonomy			5.856			3.785			6.975
			(2.490)			(4.930)			(2.867)
			[0.043]			[0.510]			[0.037]
Observations	9,727	9,727	9,727	3,454	3,454	3,454	6,273	6,273	6,273

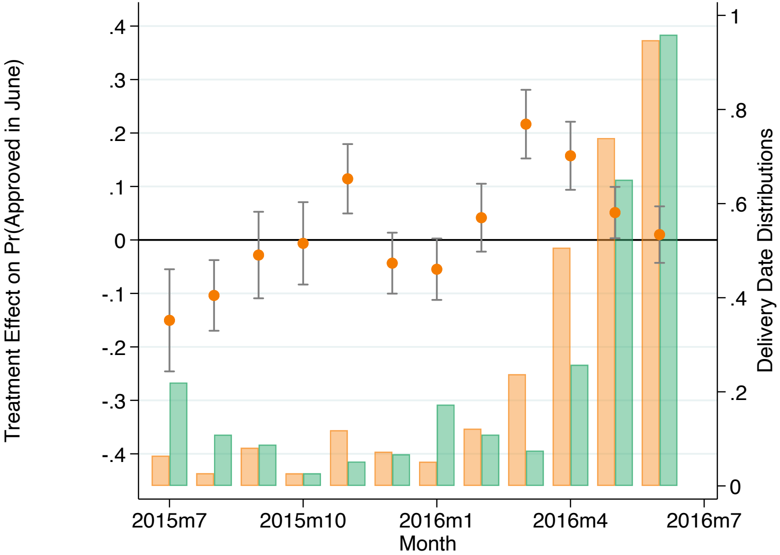
Autonomy reduces delays



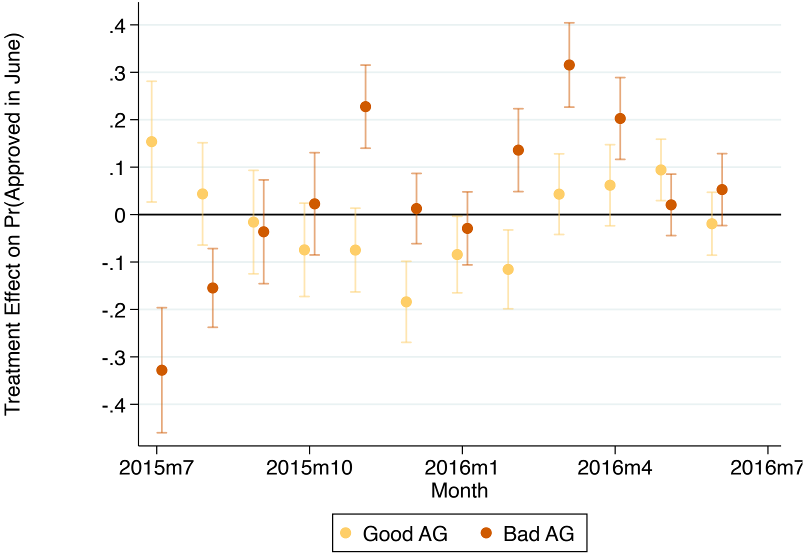
Especially when AG is inefficient



Autonomy Reduces Hold Up By the AG



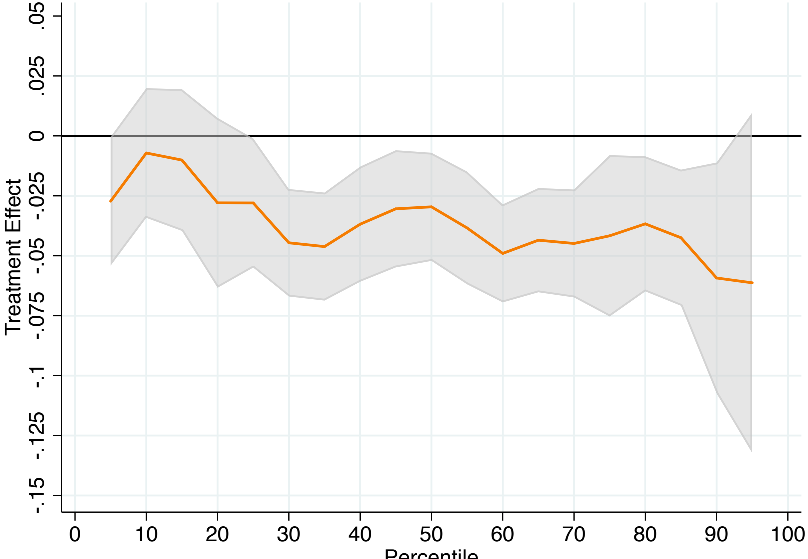
Especially when AG is inefficient



Beyond the mean

- Autonomy might allow extreme corruption
- One extreme event might be more damaging than the sum of many inefficient purchases
  - big scandals are visible, day-to-day inefficiency is not
- Look at QTEs on purchases

QTEs show no “tail”



Conclusion

- Organizations often limit autonomy to deal with agency issues
  - Limits capture but also initiative
  - Creates two sets of agents: implementing agents and monitoring agents.

Conclusion

- Organizations often limit autonomy to deal with agency issues
  - Limits capture but also initiative
  - Creates two sets of agents: implementing agents and monitoring agents.
- Experimental results from procurement bureaucrats in Punjab, Pakistan show
  - Granting authority to implementing agents helps by reducing overall capture
  - Standard p4p ineffective when the monitor's inefficiency drives prices

Conclusion

- Organizations often limit autonomy to deal with agency issues
  - Limits capture but also initiative
  - Creates two sets of agents: implementing agents and monitoring agents.
- Experimental results from procurement bureaucrats in Punjab, Pakistan show
  - Granting authority to implementing agents helps by reducing overall capture
  - Standard p4p ineffective when the monitor's inefficiency drives prices
  - Autonomy works when monitors are *ineffective*; Incentives when they are *effective.*
  - Results driven by reduction in delays and bureaucrats exerting greater effort

Conclusion

- Organizations often limit autonomy to deal with agency issues
  - Limits capture but also initiative
  - Creates two sets of agents: implementing agents and monitoring agents.
- Experimental results from procurement bureaucrats in Punjab, Pakistan show
  - Granting authority to implementing agents helps by reducing overall capture
  - Standard p4p ineffective when the monitor's inefficiency drives prices
  - Autonomy works when monitors are *ineffective*; Incentives when they are *effective.*
  - Results driven by reduction in delays and bureaucrats exerting greater effort
- In general: allocation of authority and incentives to different sets of agents hinges on precise nature of agency problems. Provide diagnostic tool

Conclusion

- Organizations often limit autonomy to deal with agency issues
  - Limits capture but also initiative
  - Creates two sets of agents: implementing agents and monitoring agents.
- Experimental results from procurement bureaucrats in Punjab, Pakistan show
  - Granting authority to implementing agents helps by reducing overall capture
  - Standard p4p ineffective when the monitor's inefficiency drives prices
  - Autonomy works when monitors are *ineffective*; Incentives when they are *effective.*
  - Results driven by reduction in delays and bureaucrats exerting greater effort
- In general: allocation of authority and incentives to different sets of agents hinges on precise nature of agency problems. Provide diagnostic tool
- In the long run: autonomy might attract talent

Outline

Non-financial Incentives

- Bandiera, Best, Khan & Prat (WP 2020): *The Allocation of Authority in Organizations: A Field Experiment with Bureaucrats*
- Duflo, Greenstone, Pande & Ryan (WP 2017): *The Value of Regulatory Discretion: Estimates from Environmental Inspections in India*
- Khan, Khwaja & Olken (WP 2018) *Making Moves Matter: Experimental Evidence on Incentivizing Bureaucrats through Performance-Based Postings*

Duflo et al 2017: Introduction

- Pollution is high despite strict standards for emissions.
- Why aren't the standards enforced?
- Typical explanations: Lack of resources, corruption
- Gujarat Pollution Control Board (GPCB) has limited resources, so uses discretion to pick which plants to inspect.
- How does the regulator choose who to inspect? Would removing that discretion help or hurt enforcement?
- Experiment and structural model in Gujarat to investigate these issues

Duflo et al 2017: Context

- Industrial plants are subject to stringent air and water emissions standards
- GPCB inspects plants to observe its condition, environmental management, and often collects pollution samples.
- Regulation mandates routine inspection. Every 90 days for large- or medium-scale plants, and once a year for small-scale plants.
- At baseline, 42% of control plants inspected less than prescribed.
- Penalties can be harsh. Regulator can mandate installation of abatement equipment, require a bond be posted against future performance, order utilities disconnected.

Duflo et al 2017: Experiment

- Experiment increased inspection frequency between August 2009 and May 2011
- Identified population of 3,455 red-category (high pollution) small- and medium-scale plants in 3 regions (Ahmedabad, Surat & Valsad).
- Drew sample of 960 plants:
  - all 473 audit-eligible plants in Ahmedabad and Surat (see Duflo et al 2013)
  - random sample of 488 plants from the remaining audit-ineligible population.
- Randomly assign inspection treatment stratifying by region $\times$ audit-treatment-status.
- 481 plants assigned to inspection guaranteed at least one annual inspection and up to 4 a year.
  - In Q1 assigned an initial inspection, and then each quarter inspected again w/pr 0.66,
  - Additional inspections done by 3 recently retired GPCB scientists rehired for the project

Table A3: Inspection Treatment Covariate Balance

	Control (1)	Treatment (2)	Difference (3)
Panel A. Plant Characteristics			
Capital investment Rs. 50m to Rs. 100m (=1)	0.087	0.071	-0.017
	[0.28]	[0.26]	(0.017)
Located in industrial estate (=1)	0.33	0.37	0.032
	[0.47]	[0.48]	(0.027)
Textiles (=1)	0.45	0.45	-0.0092
	[0.50]	[0.50]	(0.020)
Dyes and Intermediates (=1)	0.13	0.16	0.027
	[0.34]	[0.36]	(0.022)
Effluent to common treatment (=1)	0.37	0.35	-0.021
	[0.48]	[0.48]	(0.031)
Waste water generated (kl / day)	192.1	196.8	4.30
	[310.9]	[316.4]	(16.2)
Air emissions from boiler (=1)	0.50	0.52	0.019
	[0.50]	[0.50]	(0.020)

Panel B. Regulatory Interactions in Year Prior to Study

Number of inspections	1.22	1.25	0.026
	[1.32]	[1.32]	(0.079)
Inspections below prescribed (=1)	0.42	0.39	-0.031
	[0.49]	[0.49]	(0.029)
Number of pollution readings	3.64	3.92	0.28
	[5.65]	[5.58]	(0.31)
Pollution reading ever collected (=1)	0.40	0.44	0.048*
	[0.49]	[0.50]	(0.027)
Any pollution reading above limit (=1)	0.34	0.38	0.031
	[0.48]	[0.48]	(0.026)
Citations	0.22	0.20	-0.023
	[0.51]	[0.55]	(0.034)
Closure warnings	0.056	0.052	-0.0044
	[0.31]	[0.32]	(0.020)
Closure directions	0.075	0.077	0.0019
	[0.31]	[0.34]	(0.021)
Bank guarantees posted	0.019	0.029	0.010
	[0.15]	[0.21]	(0.012)
Equipment mandates	0.24	0.25	0.0082
	[0.54]	[0.53]	(0.029)
Any utility disconnection (=1)	0.010	0.0021	-0.0083
	[0.10]	[0.046]	(0.0051)
Observations	480	480	

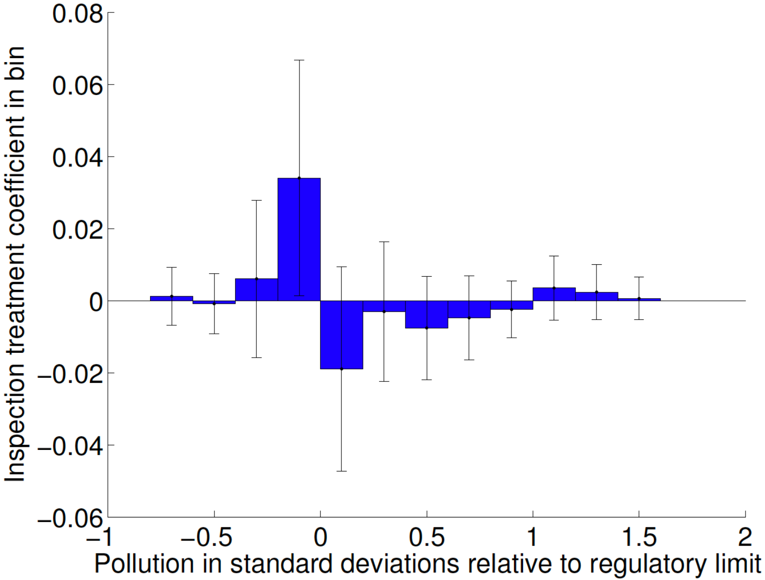
	Control	Treatment	Difference
Panel A. Inspections by Treatment Status			
Number inspections assigned in treatment, annual	0	2.12	2.12***
	[0]	[0.57]	(0.026)
Total inspections, annual over treatment	1.40	3.11	1.71***
	[1.59]	[1.77]	(0.11)
Initial inspections, annual over treatment	1.28	2.79	1.50***
	[1.38]	[1.52]	(0.094)
Observations	480	480	
Panel B. Perceived Inspections by Treatment Status			
Perceived Inspections, 2008	2.53	2.66	0.13
	[1.42]	[1.40]	(0.10)
Perceived Inspections, 2009	2.78	3.16	0.38***
	[1.44]	[1.37]	(0.100)
Perceived Inspections, 2010	2.92	3.62	0.71***
	[1.58]	[1.46]	(0.11)
Total perceived notices and closures received, 2010	0.27	0.30	0.025
	[0.64]	[0.70]	(0.048)
Observations	388	403	

Panel C. Regulatory Actions by Treatment Status

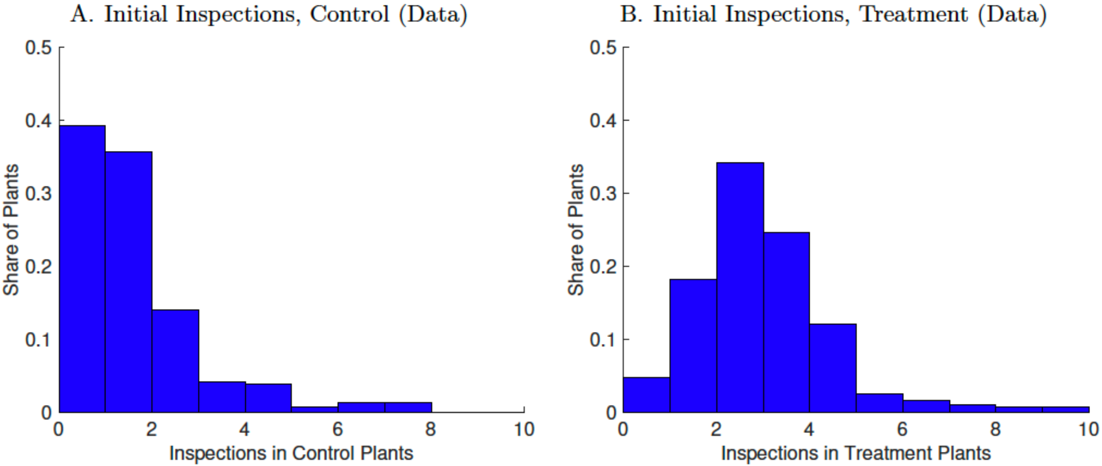
Pollution reading ever collected at plant (=1)	0.60	0.38	0.21***
	[0.49]	[0.49]	(0.032)
Any pollution reading above limit at plant (=1)	0.55	0.34	0.22***
	[0.50]	[0.47]	(0.031)
Number of pollution readings above limit at plant	2.84	1.17	1.67***
	[3.67]	[2.58]	(0.20)
Total citations	0.35	0.15	0.20***
	[0.69]	[0.42]	(0.037)
Total water citations	0.12	0.046	0.071***
	[0.37]	[0.22]	(0.020)
Total air citations	0.042	0.021	0.021*
	[0.20]	[0.14]	(0.011)
Total closure warnings	0.17	0.094	0.077***
	[0.48]	[0.34]	(0.027)
Total closure directions	0.20	0.16	0.042
	[0.54]	[0.48]	(0.033)
Total bank guarantees	0.065	0.060	0.0042
	[0.25]	[0.27]	(0.017)
Total equipment mandates	0.040	0.027	0.013
	[0.23]	[0.19]	(0.014)
Total utility disconnections	0.042	0.040	0.0021
	[0.20]	[0.22]	(0.013)
Observations	480	480	

Table 2: Endline Pollution and Compliance on Treatments

	(1)	(2)	(3)	(4)
	Panel A. Plant-level Costs			
	Capital costs		Maintenance costs	
	(USD '000s)	Any (=1)	(USD '000s)	Any (=1)
Inspection treatment (=1)	-0.221	0.0213	0.838*	0.00974
	(0.453)	(0.0344)	(0.499)	(0.0224)
Plant characteristics	Yes	Yes	Yes	Yes
Audit experiment	Yes	Yes	Yes	Yes
Control Mean	2.050	0.567	0.264	0.108
Observations	791	791	791	791

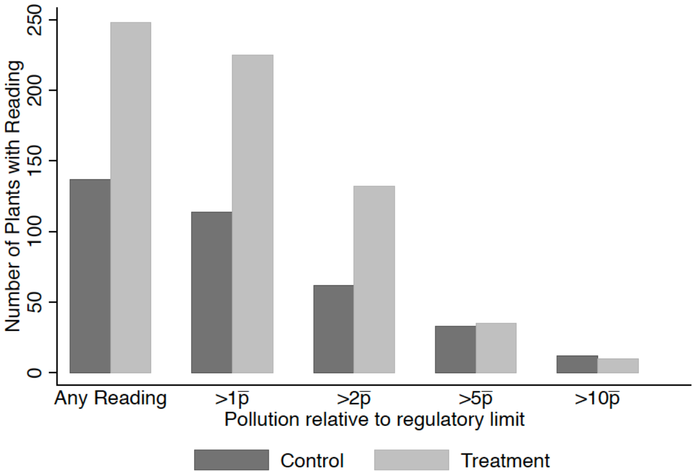

0.08
0.06
0.04
0.02
0
−0.02
−0.04
−0.06
Inspection treatment coefficient in bin
−1
−0.5
0
0.5
1
1.5
2
Pollution in standard deviations relative to regulatory limit

Treatment doesn't increase inspection of severe violators



Treatment doesn't increase inspection of severe violators

Figure 2: Regulatory Targeting of Extreme Polluters



Duflo et al 2017: Model

- Build a model of regulator-plant interactions with 2 stages

1. Targeting stage
    1.1 regulator chooses an inspection targeting rule
    1.2 plants choose whether to run abatement equipment
    1.3 regulator observes a signal of plant pollution and inspects plants by applying the targeting rule
2. Penalty stage
    2.1 Regulator acts as a regulatory machine, following exogenous rules for follow-up
    2.2 plants play a dynamic game against the machine, deciding whether to comply or risk future penalties

Duflo et al 2017: Targeting Stage

- Plant j has latent pollution in period m of

$$\log \tilde{P}_{jm} = \phi_0 + \phi_1 X_j + u_{1j} + u_{2jm}$$

  where X_j are plant observables, and the shocks are normal: $u_{1j} \sim \mathcal{N}\left(0, \sigma_1^2\right)$ and $u_{2jm} \sim \mathcal{N}\left(0, \sigma_2^2\right)$.
- Regulator sets a targeting rule $\mathcal{I}\left(u_{1j} | X_j, T_j, \theta_T\right)$ that assigns annual number of routine inspections as a function of the shock u_{1j}, plant characteristics, treatment status, and targeting parameters θ_T.
- Plants, know $\mathcal{I}(\cdot|\cdot)$, their X_j,T_j and u_{1j} so can calculate how often they will be inspected.
- Plants also know their abatement cost c_j, $\log c_j \sim \mathcal{N}\left(\mu_c, \sigma_c^2\right)$ and decide whether to Run their equipment (which the regulator doesn't observe), which reduces their pollution to

$$\log P_{jm} = \log \tilde{P}_{jm} + \phi_2 Run, \quad \phi_2 < 0$$

Duflo et al 2017: Targeting Stage

- Summarize the costs of regulation to plants by a penalty value function $V_0\left(P_{jm}\right)$ giving the money value to the plant of an initial inspection that finds pollution P_{jm} (expected value of all regulatory actions in penalty stage)
- If a plant expects to be inspected I_j times, the abatement equipment will be run if benefit exceeds maintenance costs

$$Run^* = \mathbf{1}\left\{I_j \times \left(V_0\left(P_{jm}\right) - V_0\left(\tilde{P}_{jm}\right)\right) > c_j\right\}$$

Duflo et al 2017: Targeting Stage

- The regulator wants to set a rule that maximizes total abatement.
- The rule depends on parameters $\lambda, \beta, \rho \in \theta_T$ but we assume β & ρ are exogenous, so regulator picks (vector) λ to solve

$$\lambda^* \in \arg\max_{\lambda} \sum_{j=1}^{N} \int \int \mathbb{P}\left(\mathcal{I}\left(u_{1j}|X_j, T_j, \theta_T\right)\left[V_0\left(P_{jm}\right) - V_0\left(\tilde{P}_{jm}\right)\right]\right)$$
$$\times \tilde{P}_{jm}\left(1 - e^{\phi_2}\right) dF\left(U_2\right) dF\left(U_1\right)$$

such that

$$\sum_{j=1}^{N} \int \mathcal{I}\left(u_{1j}, X_j, T_j, \theta_T\right) dF\left(U_1\right) = N\bar{I}$$

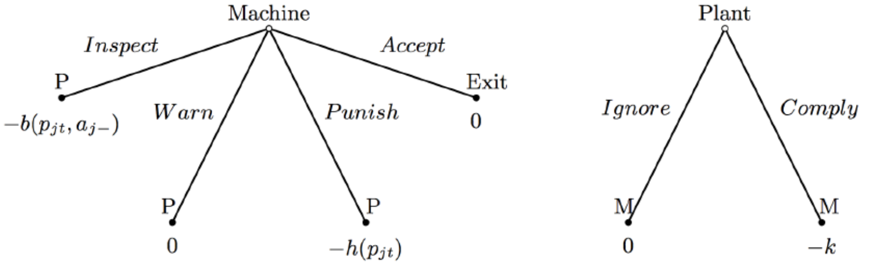
- For estimation assume

$$\mathcal{I}\left(u_{1j}|X_j, T_j, \theta_T\right) = \lambda_2 \Phi\left(\frac{\lambda_1 + X_j'\beta_1 + T_j'\beta_2 + u_1}{\rho}\right)$$

Duflo et al 2017: Penalty Stage

- We want to estimate the value function $V_0(P_{jm})$ by modeling the game after an initial inspection as a dynamic discrete choice problem.
- Game starts in round 1 with an initial inspection.
- In subsequent rounds, plant j and the regulatory machine R alternate moves.
- When the plant moves,
  - it can $Comply$ or $Ignore$ the regulatory machine.
  - Complying means paying to install abatement equipment.
  - Acts to minimize regulatory cost
- When the regulatory machine moves,
  - it has four actions $a_{R_t} : Inspect, Warn, Punish, Accept$
  - Regulatory machine: fixed probabilities of any action conditional on state (assume known to the plant)

Duflo et al 2017: Penalty Stage



A. Actions of the Regulatory Machine

B. Actions of the Plant

The figure gives the actions of the regulatory machine and plant at each node and the terminal nodes give the payoffs in each round for the plant. The penalty stage begins with an inspection where the Regulatory Machine (M) observes p_{j1}. The machine can take four actions. If M *Inspects*, M gets a new signal of pollution and the plant may have to offer a bribe with payoff $-b(p_{jt}, a_{j-})$. If M *Warns*, there is no cost to the plant. If *M Punishes*, the plant faces a cost $-h(p_{jt})$. After each of these moves the plant *Ignores* or *Complies* and *M* moves again. If M *Accepts*, the stage ends.

Duflo et al 2017: Estimating Penalty Stage

- Scanned and coded up 9,624 documents on interactions with plants.
- Use maximum likelihood to estimate the game
- The state space is the pollution reading, the last two actions by the regulator and the plant, and the game round t.
- Estimate transition probabilities using count estimator
- Estimate action probabilities for the regulatory machine conditional on the state using multinomial logit.

Duflo et al. 2017: Estimating Penalty Stage

- Have to specify
- the cost of inspections $b\left(p_{jt}, a_{j-}\right) = \left(1 - \mathbf{1}\left\{a_{j-} = Comply\right\}\right) \times \left(\nu_1 \mathbf{1}\left\{\bar{p} < p_{jt} < 2\bar{p}\right\} + \nu_2 \mathbf{1}\left\{2\bar{p} \leq p_{jt} < 5\bar{p}\right\} + \nu_3 \mathbf{1}\left(5\bar{p} \leq p_{jt}\right)\right)$
  - $h\left(p_{jt}\right)$ cost if regulator picks punish:
    - $-\tau_0$
    - $-\tau_1 \mathbf{1}\left\{\bar{p} < p_{jt} < 2\bar{p}\right\} - \tau_2 \mathbf{1}\left\{2\bar{p} \leq p_{jt} < 5\bar{p}\right\} - \tau_3 \mathbf{1}\left\{5\bar{p} \leq p_{jt}\right\}$
  - k, cost of abatement equipment: $k = \$17,000$, average cost observed in data
- Now for each action a_{jt} we know the within-round payoff $\pi_j\left(a_{jt} | st\right)$

Duflo et al 2017: Estimating Penalty Stage

- The plant's utility of taking action a_{jt} in state s_t is now

$$v_j(a_{jt}) = \underbrace{\pi_j(a_{jt}|s_t) + e_j(a_{jt}|s_t)}_{\text{payoff today}}$$
$$+ \delta \sum_{s_{t+1}} \underbrace{f(s_{t+1}|a_{jt}, s_t)}_{\text{state transition prob}} \sum_{a_{R,t+1}} \underbrace{\mathbb{P}(a_{R,t+1}|s_{t+1})}_{\text{machine's action prob}} \times$$
$$\left\{ \underbrace{\pi_j(a_{R,t+1}|s_{t+1})}_{\text{payoff from } a_{R,t+1}} + \delta \sum_{s_{t+2}} \underbrace{f(s_{t+2}|a_{R,t+1}, s_{t+1})}_{\text{state transition prob}} V_j(s_{t+2}) \right\}$$

  where $e_j(a_{jt}|s_t)$ are utility shocks with type-I EV distribution and δ is the discount factor
- The machine's action probabilities and the state transition probabilities are known to the plant

Duflo et al 2017: Estimating Penalty Stage

- Yields value of the state

$$V_j\left(s_t\right) = \max_{a \in A_p} v_j\left(a_{jt} | s_t\right)$$

- By backward induction, solve for the state values given $\theta_p = \{\tau, \nu\}$

Duflo et al 2017: Estimating Penalty Stage

- Now we have the actions the plant will take for any parameters and probabilities.
- Estimate the state transitions with

$$\hat{f}\left(s'|a_{jt}, s_t\right) = \frac{\sum_{j,c,t} \mathbf{1}\left(s_{j,t+1} = s'|s_{jt}, a_{jt}\right)}{\sum_{j,c,t} \mathbf{1}\left(s_{jt}, a_{jt}\right)}$$

- Estimate the conditional action probabilities with

$$\mathbb{P}\left(a_{Rt} = a|s_t\right) = \frac{\exp\left(q\left(s_t\right)' \omega_a\right)}{\sum_{a'} \exp\left(q\left(s_t\right)' \omega_a\right)}$$

- Now, the likelihood of observing a set of n chains of regulator actions is

$$\mathcal{L}\left(\theta_p\right) = \prod_n \prod_{t=1}^{T_{jn}} \mathbb{P}\left(a_{jnt}|s_{jnt}, \theta_P\right)$$

which can be maximized numerically

Duflo et al 2017: Estimating the Targeting Stage

- Parameters to estimate $\theta_T = \left\{ \underbrace{\phi}_{\text{pollution}}, \underbrace{\beta, \lambda_1, \lambda_2}_{\text{targeting rule}}, \underbrace{\mu_c}_{\text{mean abatement cost}}, \underbrace{\sigma_1, \sigma_2}_{\text{pollution shocks}} \right\}$
- Estimate by GMM using data on the number of inspections N_j, observables X_j, treatment status T_j, pollution P_j and abatement costs of abaters $c_j \times Run$ and the estimated $\hat{V}_0(p_j)$ from the penalty stage

Duflo et al 2017: Estimating the Targeting Stage

- Moments to use

1. Pollution equation: Let $Z_j = \begin{bmatrix} 1 & X_j & T_j \end{bmatrix}$

$$g_1(\phi) = Z_j' \left(\log P_j - \phi_0 - \phi_1 X_j - \phi_2 Run\right)$$

2. Expected inspections

$$g_2(\lambda, \beta) = \mathbf{1}' \left(\mathbb{E}\left[\mathcal{I}\left(u_{1j} | X_j, T_j, \lambda, \beta, \rho\right)\right] - I_j\right)$$
$$g_3(\lambda, \beta) = \mathbf{1}' \left(\mathbb{E}\left[\mathcal{I}^2\left(u_{1j} | X_j, T_j, \lambda, \beta, \rho\right)\right] - I_j^2\right)$$

3. Probability and cost of runnning abatement equipment

$$g_4(\phi, \mu, \sigma) = \mathbb{P}\left(Run = 1 | \phi, \mu, \sigma\right) - \frac{1}{N} \sum_j \mathbf{1}\left\{c_j > 0\right\}$$

$$g_5(\phi, \mu, \sigma) = \mathbb{E}\left[c_j | Run = 1, \phi, \mu, \sigma\right] - \frac{1}{\sum_j \mathbf{1}\left\{c_j > 0\right\}} \sum_j \mathbf{1}\left\{c_j > 0\right\} c_j$$

Duflo et al 2017: Estimating the Targeting Stage

4. Variance of pollution shocks and covariance with inspections

$$g_6(\beta,\phi,\sigma) = \mathbb{E}\left[\varepsilon_2^2|\beta,\phi,\sigma\right] - \sum_j \hat{\varepsilon}_2^2/N = \sigma_1^2 + \sigma_2^2 - \sum_j \hat{\varepsilon}_2^2/N$$

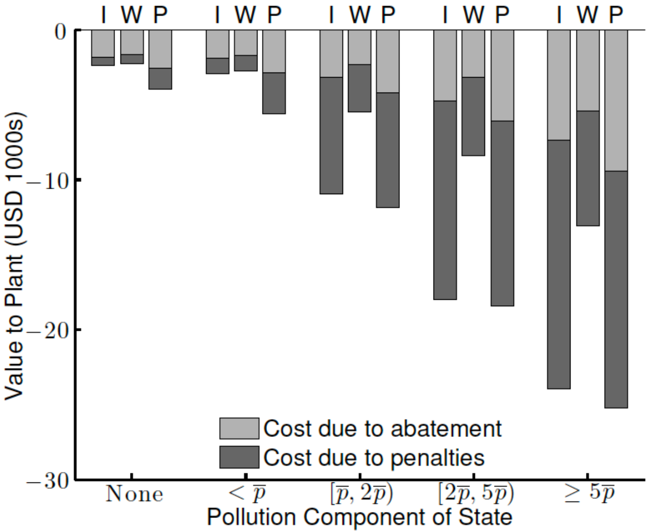
$$g_7(\beta,\phi,\sigma) = \mathbb{E}\left[\varepsilon_2 \cdot \mathcal{I}|\theta\right] - \sum_j \hat{\varepsilon}_{2j} \times I_j/N$$

- stack all these moments to form $g(\theta_T) = \begin{bmatrix} g_1' & g_2' & \ldots & g_7' \end{bmatrix}$
- Minimize gWg' to estimate $\hat{\theta}_T$.

Table 3: Structure of Penalty Stage Actions

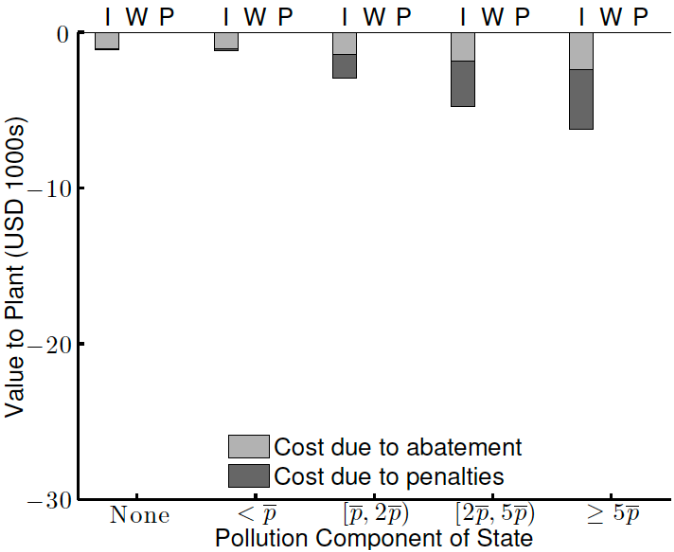
Round	Regulatory Action				Plant Action			
	Inspect	Warn	Punish	Accept	Ignore	Comply	N	% left
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1	100.0	0.0	0.0	0.0			7423	100.0
2					99.6	0.4	7423	
3	1.0	9.5	2.2	87.3			7423	100.0
4					92.8	7.2	941	
5	23.3	4.8	5.3	66.6			941	12.7
6					91.1	8.9	314	
7	18.8	11.8	9.9	59.6			314	4.2
8					83.5	16.5	127	
9	21.3	5.5	18.1	55.1			127	1.7
10					82.5	17.5	57	
11	26.3	3.5	10.5	59.6			57	0.8
12					87.0	13.0	23	
13	26.1	4.3	8.7	60.9			23	0.3
14					77.8	22.2	9	
15+	16.7	8.3	0.0	75.0	100.0	0.0	9	0.1
Total without inspections	0.0	4.6	1.6	42.7	50.2	0.9	7824	
Total	31.0	3.2	1.1	29.4	34.6	0.6	25217	

A. Value at $t = 6$



I W P
I W P
I W P
I W P
I W P
0
−10
−20
−30
Value to Plant (USD 1000s)
Cost due to abatement
Cost due to penalties
None
$< \overline{p}$
$[\overline{p}, 2\overline{p})$
$[2\overline{p}, 5\overline{p})$
$\geq 5\overline{p}$
Pollution Component of State

B. Value at $t = 2$



I W P
I W P
I W P
I W P
I W P
0
−10
−20
−30
Value to Plant (USD 1000s)
Cost due to abatement
Cost due to penalties
None
$< \overline{p}$
$[\overline{p}, 2\overline{p})$
$[2\overline{p}, 5\overline{p})$
$\geq 5\overline{p}$
Pollution Component of State

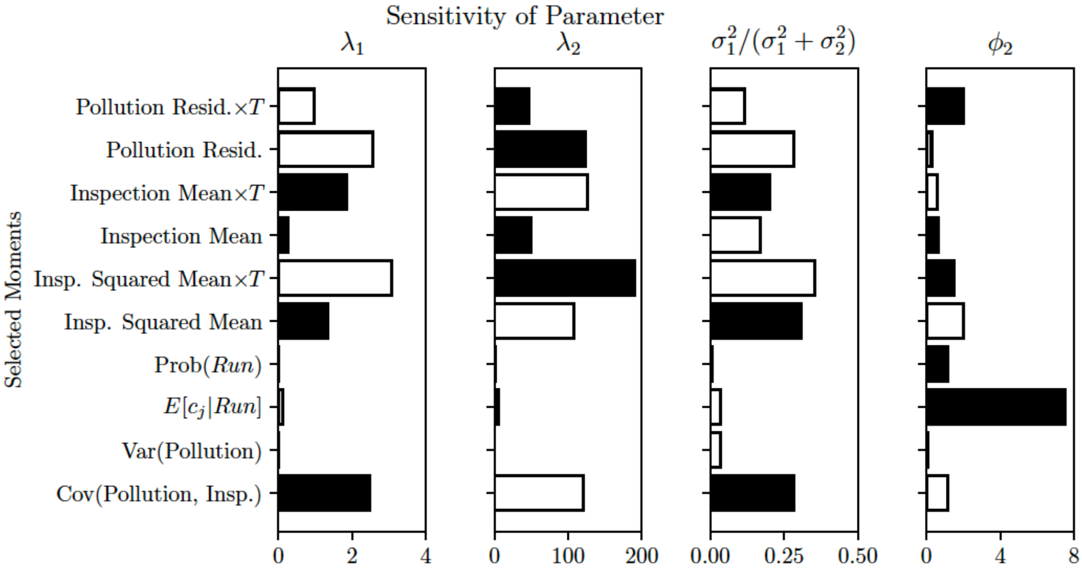
Table 4: Multinomial Logit Model of Action Choice Conditional on State

Party to move:	Regulatory Machine			Plant
	Inspect	Warn	Punish	Comply
	(1)	(2)	(3)	(4)
Lagged regulatory actions				
Warn, lag 1	0.33	-2.05***	-2.10***	-0.23
	(0.23)	(0.32)	(0.31)	(0.30)
Punish, lag 1	1.80***	-2.22***	-0.53*	1.29***
	(0.23)	(0.56)	(0.30)	(0.26)
Lagged plant actions				
Firm: Comply, lag 1	-1.80***	-1.03**	-0.82**	-0.53
	(0.32)	(0.47)	(0.37)	(0.66)
Last observed pollution reading				
0-1x	-0.38	-0.25	0.052	-0.18
	(0.23)	(0.16)	(0.24)	(0.38)
1-2x	-0.20	0.55***	0.37**	0.39*
	(0.16)	(0.098)	(0.18)	(0.23)
2-5x	-0.17	0.84***	0.70***	0.74***
	(0.17)	(0.10)	(0.17)	(0.22)
5x+	0.27	0.63***	1.15***	0.90***
	(0.21)	(0.16)	(0.21)	(0.26)

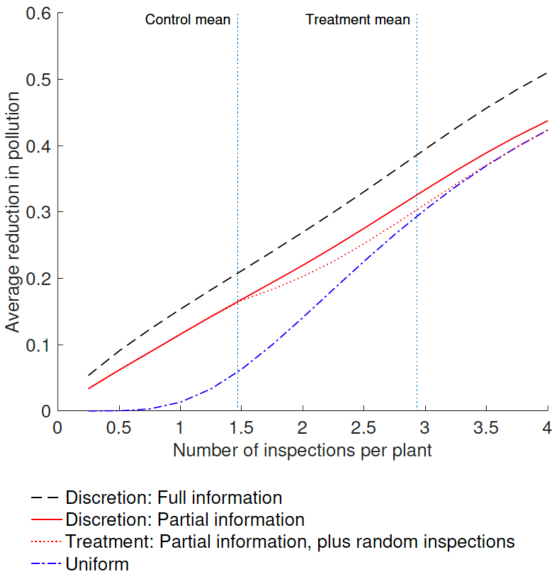
Table 6: Estimates of Targeting Stage Parameters

	Constrained		Unconstrained	
	Initial Inspections	Log Pollution	Initial Inspections	Log Pollution
	(1)	(2)	(3)	(4)
Panel A. Targeting and Pollution Equations				
Inspection treatment	0.095		0.162	
	(0.009)		(0.025)	
Run equipment (=1)		-1.902		-0.711
		(0.160)		(0.308)
Inspection targeting shift	-0.395		-0.220	
parameter (λ_1)	(0.003)		(0.066)	
Inspection targeting level	33.022		10.064	
parameter (λ_2)	(1.876)		(3.137)	
Constant		0.212		-0.009
		(0.109)		(0.102)
Panel B. Distributions of Pollution and Maintenance Cost Shocks				
Standard deviation of observed	0.069		0.111	
pollution shock (σ_1)	(0.003)		(0.022)	
Standard deviation of unobserved	1.033		0.864	
pollution shock (σ_2)	(0.047)		(0.042)	
Mean of log maintenance	2.388		1.833	
cost (μ_c)	(0.061)		(0.334)	
Panel C. Test of Targeting Optimality Constraints				
Distance metric test statistic χ^2_2	16.1039			
Test p-value	0.0003			

Figure 5: Sensitivity of Targeting Parameters to Moments



Value of Discretion: Abatement by Information Regime and Budget Constraint



Outline

Non-financial Incentives

- Bandiera, Best, Khan & Prat (WP 2020): *The Allocation of Authority in Organizations: A Field Experiment with Bureaucrats*
- Duflo, Greenstone, Pande & Ryan (WP 2017): *The Value of Regulatory Discretion: Estimates from Environmental Inspections in India*
- Khan, Khwaja & Olken (WP 2018) *Making Moves Matter: Experimental Evidence on Incentivizing Bureaucrats through Performance-Based Postings*

Khan et al 2018: Introduction

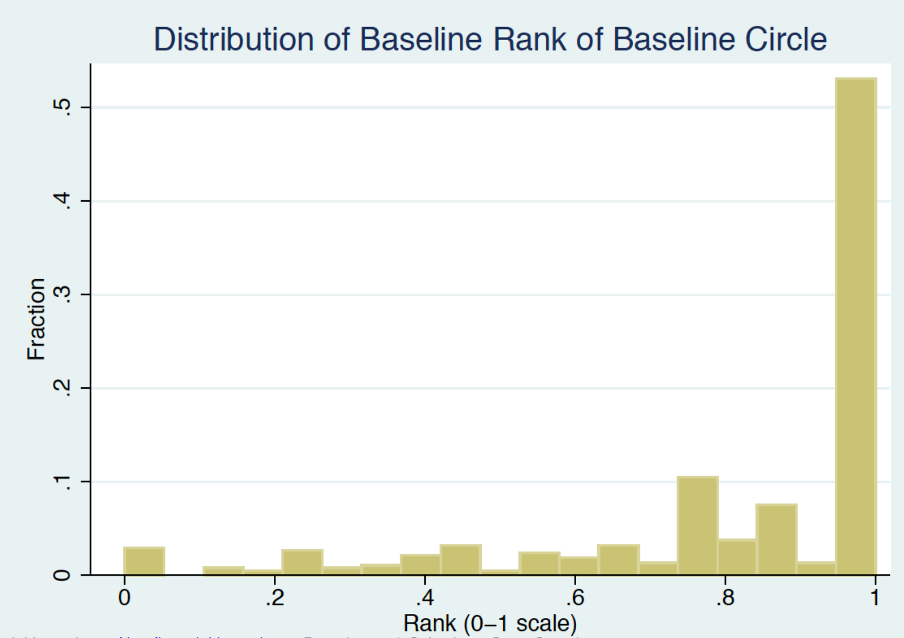
- Incentivizing workers in the government is doubly hard.
  - Scope for performance pay often limited
  - promotion often mechanical
  - punishment hard
- One thing we might try and leverage is transfers
  - bureaucrats are transfered all the time, but largely because of personal/political connections, idiosyncratic preferences, or arbitrariness
- Run an experiment in Punjab, Pakistan with property tax inspectors to test this.
- Design and implement a *performance-ranked serial dictatorship* (PRSD) mechanism

Khan et al 2018: Context

- Study urban property tax in Punjab, Pakistan (population 110 million)
- Same tax setting as in Khan et al 2016.
- About 1/3 inspectors is transferred each year
- Inspectors care about where they are posted, the circles are very heterogeneous.
  - Size: 90th %ile has 3x properties of 10th %ile
  - Ease of collecting taxes, opportunities for corruption
  - Amenities
- Current transfer process is opaque and subject to political influence, so limited use as incentive.

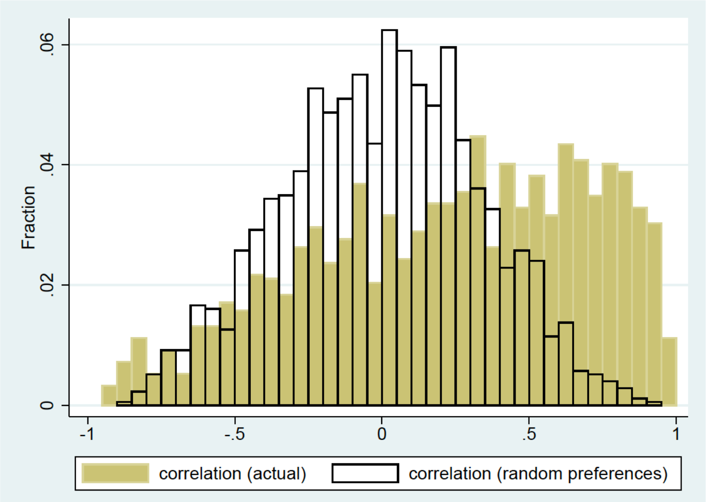
Khan et al 2018: Data

- Admin data on tax performance
  - quarterly inspector reports showing collections and tax base
- Preference data.
  - ask inspectors to rank all circles (average 10) in their district.
  - Told preferences would be used in treatments (incentive compatible to reveal true preferences) but before told treatment status
- Calculate the allocation that makes no inspector worse off, and no inspector or group of inspectors would want to deviate (Shapley & Scarf, 1974). 15% of inspectors would move, and these individuals would move 30%iles up their preference ordering.
- 15% relatively small, suggests that the incentive scheme will create both winners and losers.



Distribution of Baseline Rank of Baseline Circle
Fraction
0
.1
.2
.3
.4
.5
0
.2
.4
.6
.8
1
Rank (0–1 scale)

(b) Distribution of pairwise rank correlations



Khan et al 2018: Treatment

- Worked with Excise & Taxation department on the "Merit-Based Transfers and Postings" (MBTP) scheme
- Inspectors in MBTP scheme randomized into groups of 10 circles within districts.
- Inspectors told they would be ranked based on performance, and then based on ranking they would be given a choic of circles within their group.
- Performance based on 2 measures (randomized)
  1. Recovery: y-o-y percentage increase in tax collected
  2. Demand: y-o-y percentage increase in assessed tax base

Khan et al 2018: Modeling Incentives

- Inspector i gets utility u_{ij} from being assigned to circle j.
- Denote overall preference matrix by $\mathbf{P}$
- growth rate in revenue is $y_i = y_{i0} + e_i + \epsilon_i$, $\epsilon_i \sim iid$ with sd σ_ϵ
- For a vector of outcomes $\mathbf{y}$, the PRSD mechanism yields an allocation $r_i\left(\mathbf{y}, \mathbf{P}\right)$
- If effort has convex cost $c\left(e_i\right)$ then each inspector maximizes

$$\max_{e_i} \sum_{j=1}^{J} u_{ij} \mathbb{P}\left(j = r_i\left(\mathbf{y}, \mathbf{P}\right)\right) - c\left(e_i\right)$$

$\Rightarrow$ effort choice satisfies

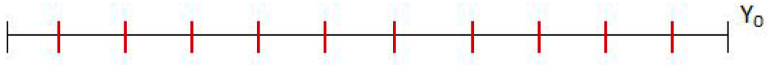
$$\frac{d\mathbb{E}\left[u\right]}{de_i} = \sum_{j=1}^{J} u_{ij} \frac{\partial \mathbb{P}\left(j = r_i\left(y_i, \mathbf{y}_{-i}, \mathbf{P}\right)\right)}{\partial y_i} = c'\left(e_i\right)$$

Khan et al 2018: Influences on Incentives

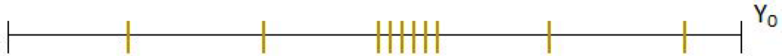
1. Preferences $\mathbf{P}$
   1.1 imagine everyone has the same preferences. Then all that matters is the inspector's rank in the performance distribution. $r_i = Rank\left(y_i, \mathbf{y}\right)$
   1.2 Imagine nobody shares a first-choice with anyone else. Then they all get their first choice, and scheme provides *no incentives*
2. Distribution of y_0.
   2.1 If they are all close together, then small effort changes can change ranks: strong incentives.
   2.2 If they are far apart, need lots of effort to change rank: weak incentives
3. Preferences u_{ij}
   3.1 With common preferences and common y_0, Lazear & Rosen (1981) show that you can replicate the efficient piece rate with a tournament
   3.2 With general preferences, it's not clear how close we are to a piece rate.

(a) When y_0 is concentrated, the marginal return to effort is high for all inspectors.

(b) When y_0 is spread out, marginal returns to effort are low.



(c) Within a group variation: inspectors with y_0 close together face strong incentives, whereas those with y_0 far apart face weaker.

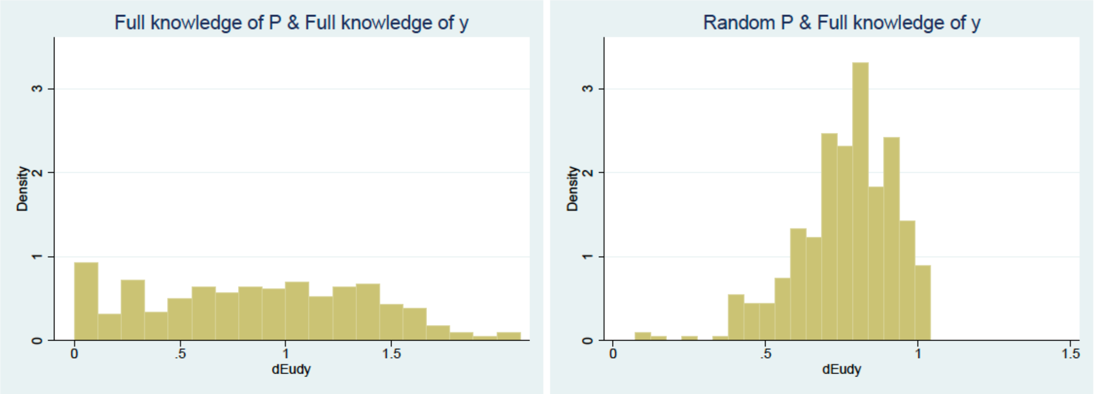


Khan et al 2018: Marginal return to effort

- Simulate the model under different assumptions about what the inspectors know about $\mathbf{P}$ and $\mathbf{y}_0$
- For a given effort vector $\mathbf{e}$, solve for the Nash equilibrium efforts.
- Need to parameterize
  - u_{ij}: take ordinal preferences and linearize $u_{ij} = 1$ for top circle and $u_{ij} = 0$ for bottom circle
  - $\mathbf{y}_0$: regress revenue changes on 2 lags of revenu and tax base in the control group. get $\mathbf{y}_0$ and σ_ϵ^2.
- Evaluate the lhs of the FOC $d\mathbb{E}\left[u\right]/de_i$ at $\mathbf{e} = 0$ under different assumptions:

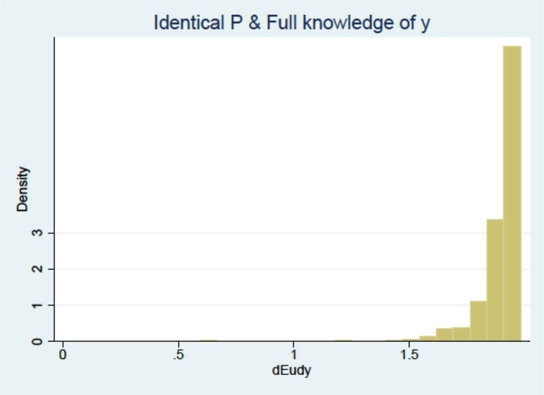
Khan et al 2018: Marginal return to effort

(a) Full Knowledge of P and y (b) Assuming random preferences P, full knowledge of y

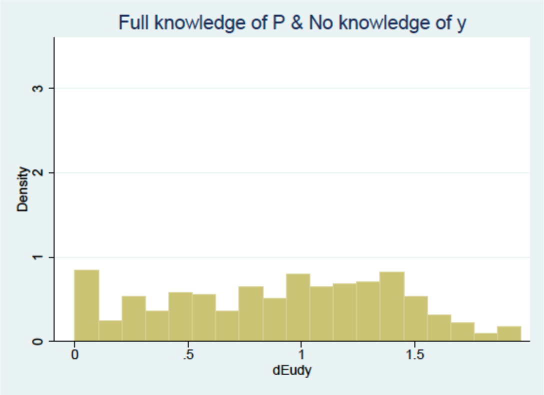


Khan et al 2018: Marginal return to effort

(c) Assuming identical preferences P, full knowledge of y

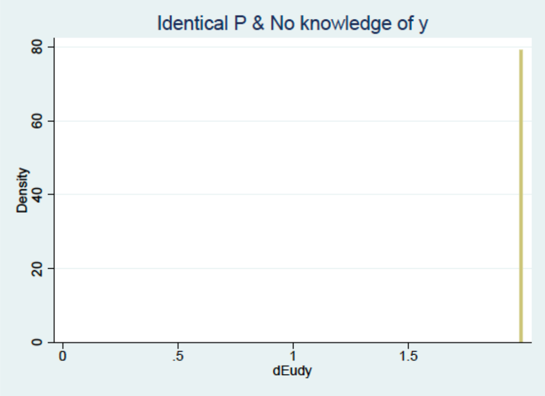


(d) Full knowledge of P, no knowledge of y

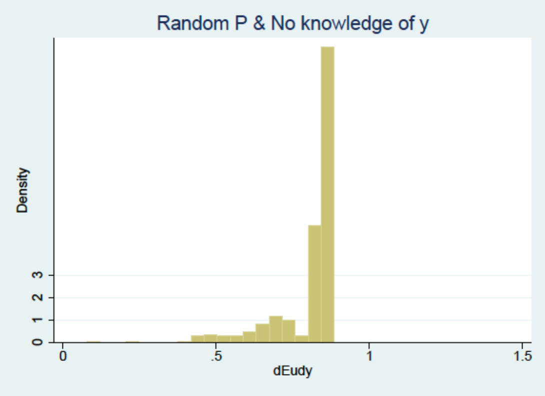


Khan et al 2018: Marginal return to effort

(e) Assuming identical preferences P, no knowledge of y



(f) Assuming random preferences P, no knowledge of y



Khan et al 2018: Experiment

- At start of year 1, circles randomly assigned into groups of 9-11 circles within metropolitan area. $\Rightarrow$ 41 groups.
- Randomized into treatment, control. Within treatment into recovery/demand measure of performance.
- In year 2, groups rerandomized into treatment/control stratifying by year-1 treatment status.

	Year 2 Control	Year 2 Treatment	Total
Year 1 Control	207	50	257
Year 1 Treatment	72	81	153
(Not included in Year 1 lottery)	96	19	115
Total	375	150	525

Khan et al 2018: Estimation

- Estimate treatment effects in circle c with

$$\log y_{ct} = \alpha_t + \gamma_t \log y_{c0} + \beta TREAT_c + \epsilon_{ct}$$

- Using model simulations, ge predicted equilibrium effort $\tilde{e}_i$ from each inspector and estimate

$$\log y_{ct} = \alpha_t + \alpha_g + \gamma_t \log y_{c0} + \beta_1 TREAT_c \times \tilde{e}_i + \beta_2 \tilde{e}_i + \beta_3 TREAT_c \times X_c + \beta_4 X_c + \epsilon_{ct}$$

- Estimate effect of re-randomizing with

$$\log y_{ct} = \alpha + \gamma \log y_{c0} + \beta_1 TREAT\_Y1_c + \beta_2 TREAT\_Y2_c + \beta_3 TREAT\_Y1_c \times TREAT\_Y2_c + \varepsilon_{ct}$$

Khan et al 2018: Revenue Results

Table 3: Treatment Effect on Log Tax Revenue

	Year 1 (Y1 Q4)			Year 2 (Y2 Q4)			Pooled		
	(1) Total	(2) Current	(3) Arrears	(4) Total	(5) Current	(6) Arrears	(7) Total	(8) Current	(9) Arrears
Treatment	0.049	0.048	0.065	0.092	0.069	-0.074	0.061	0.054	0.026
	(0.022)	(0.023)	(0.056)	(0.042)	(0.040)	(0.119)	(0.020)	(0.021)	(0.052)
	[0.009]	[0.023]	[0.259]	[0.036]	[0.142]	[0.594]	[0.002]	[0.004]	[0.653]
N	405	405	396	251	251	244	656	656	640
Mean growth in controls	0.117	0.154	-0.048	0.309	0.408	-0.337	0.203	0.268	-0.177

Notes: OLS regressions of log of tax revenue on treatment assignment. The unit of observation is a circle, as defined at the time of randomization. Specification controls for baseline values (FY 2013). Robust standard errors in parentheses. Standard errors are clustered by circle. Randomization inference based p-values in brackets.

	(1)	(2)	(3)	(4)
Panel A: Full knowledge of P, Y				
Treatment * Eq. effort	0.027	0.045	0.035	0.074*
	(0.023)	(0.039)	(0.023)	(0.040)
Treatment * Tax base at baseline		-0.046		-0.086
		(0.052)		(0.055)
Treatment * Recovery rate at baseline			-0.180	-0.246*
			(0.133)	(0.136)
Eq. effort	0.002	-0.001	-0.003	-0.014
	(0.015)	(0.017)	(0.016)	(0.016)
Panel B: Random P, full knowledge of Y				
Treatment * Eq. effort	0.020	0.037	0.051*	0.111**
	(0.025)	(0.032)	(0.028)	(0.045)
Treatment * Tax base at baseline		-0.036		-0.104**
		(0.044)		(0.053)
Treatment * Recovery rate at baseline			-0.217*	-0.332**
			(0.132)	(0.141)
Eq. effort	0.016	0.016	0.011	-0.010
	(0.024)	(0.028)	(0.026)	(0.027)

Panel C: Assume identical P, full knowledge of Y				
Treatment * Eq. effort	-0.007	-0.008	0.012	0.012
	(0.020)	(0.021)	(0.027)	(0.028)
Treatment * Tax base at baseline		-0.019		-0.038
		(0.036)		(0.037)
Treatment * Recovery rate at baseline			-0.185	-0.207
			(0.143)	(0.143)
Eq. effort	0.007	0.009	0.004	0.003
	(0.012)	(0.014)	(0.013)	(0.013)
Panel D: Full knowledge of P, no knowledge of Y				
Treatment * Eq. effort	0.024	0.037	0.023	0.047
	(0.024)	(0.040)	(0.026)	(0.038)
Treatment * Tax base at baseline		-0.037		-0.062
		(0.049)		(0.050)
Treatment * Recovery rate at baseline			-0.170	-0.207
			(0.137)	(0.139)
Eq. effort	-0.003	-0.005	-0.006	-0.012
	(0.013)	(0.014)	(0.013)	(0.014)
N	652	652	652	652
Mean of control group	16.078	16.078	16.078	16.078

	Y1 Preferences (Treatment)				Allocation		Difference in allocation			
	(1)		(2)		(3)		(4)		(5)	
	All circles		Top inspectors' circles		Treated inspectors		Treatment - Control (Revenue)		Treatment - Control (Tax base)	
	b / se	Mean	b / se	Mean	b / se	Mean	b / se	Mean	b / se	Mean
Log of tax base (Current)	0.167**	15.870	0.343*	15.873	0.312*	15.906	0.537*	16.055	0.173	16.050
	(0.070)		(0.177)		(0.179)		(0.304)		(0.278)	
Log of tax base (Arrears)	0.137	14.254	0.173	14.228	0.219	14.224	0.355	14.492	-0.092	14.552
	(0.128)		(0.413)		(0.407)		(0.677)		(0.667)	
Growth in tax base (Current)	0.001	0.101	0.004	0.099	0.011	0.094	0.006	0.113	-0.024	0.109
	(0.008)		(0.035)		(0.036)		(0.073)		(0.040)	
Growth in tax base (Arrears)	0.055	-0.321	-0.117	-0.335	-0.068	-0.361	-0.022	-0.362	-0.199	-0.317
	(0.086)		(0.199)		(0.220)		(0.414)		(0.216)	
Log of revenue (Current)	0.180**	15.565	0.376**	15.566	0.338*	15.605	0.635**	15.737	0.235	15.735
	(0.072)		(0.177)		(0.172)		(0.309)		(0.304)	
Log of revenue (Arrears)	0.151	13.848	0.113	13.814	0.152	13.821	0.669	14.023	-0.193	14.086
	(0.123)		(0.328)		(0.337)		(0.626)		(0.430)	
Growth in revenue (Current)	-0.003	0.142	0.024	0.140	0.029	0.138	0.057	0.172	0.040	0.158
	(0.011)		(0.036)		(0.037)		(0.067)		(0.060)	
Growth in revenue (Arrears)	0.068	-0.331	-0.192	-0.351	-0.164	-0.359	0.144	-0.353	-0.355	-0.312
	(0.093)		(0.220)		(0.242)		(0.435)		(0.231)	
Any unofficial payment	0.050*	0.395	0.040	0.395	0.034	0.404	-0.039	0.387	0.196	0.375
	(0.026)		(0.081)		(0.079)		(0.134)		(0.139)	
Log of unofficial payment rate	-0.043	0.704	-0.219*	0.728	-0.211*	0.705	-0.378	0.692	-0.402*	0.698
	(0.041)		(0.128)		(0.122)		(0.237)		(0.218)	
Log average p.c. expenditure	0.066	8.614	0.097	8.611	0.082	8.631	0.141	8.652	0.262	8.625
	(0.046)		(0.096)		(0.101)		(0.157)		(0.171)	
Properties for commercial use	-0.004	0.322	-0.072	0.325	-0.072	0.328	-0.016	0.367	-0.092	0.356
	(0.016)		(0.051)		(0.049)		(0.079)		(0.113)	
Properties for residential use	-0.006	0.424	0.114*	0.419	0.119*	0.413	0.056	0.377	0.150	0.381
	(0.015)		(0.068)		(0.068)		(0.102)		(0.153)	
Num of properties (in hundreds)	-5.497	65.585	-15.182**	68.221	-11.301*	63.547	-4.070	75.349	-30.497***	74.123
	(3.594)		(7.654)		(6.780)		(14.674)		(11.300)	
Log of average property value	0.204*	7.630	0.487	7.608	0.489	7.631	0.062	7.869	1.450**	7.809
	(0.114)		(0.347)		(0.341)		(1.103)		(0.573)	
N	1184		136		123		197		199	

	(1) Total	(2) Current	(3) Arrears
Y1 Treatment (β_1)	0.109	0.085	0.128
	(0.038)	(0.040)	(0.100)
	[0.003]	[0.020]	[0.200]
Y2 Treatment (β_2)	0.081	0.055	-0.074
	(0.043)	(0.041)	(0.119)
	[0.064]	[0.235]	[0.592]
Y1 AND Y2 Treatment (β_3)	-0.150	-0.085	-0.061
	(0.067)	(0.068)	(0.178)
	[0.014]	[0.203]	[0.733]
N	403	403	392
$\beta_1 = \beta_2$	0.564	0.560	0.167
$\beta_1 + \beta_3 = 0$	0.401	0.999	0.655
Mean growth in controls	0.309	0.408	-0.337

Outline

Theory

Financial Incentives

Non-financial Incentives

Recruitment & Selection

Open Questions

Outline

Recruitment & Selection

- Dal Bó, Finan & Rossi (QJE 2013) *Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service*
- Ashraf, Bandiera & Lee (2018) *Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services*
- Leaver, Ozier, Serneels & Zeitlin (AER 2021) *Recruitment, Effort and Retention Effects of Performance Contracts for Civil Servants: Experimental Evidence from Rwandan Primary Schools*
- Xu (AER 2018) *The Costs of Patronage: Evidence from the British Empire*

Dal Bó et al 2013: Introduction

- States need good people in order to function well.
- “Good” bureaucrats has many dimensions, how do pecuniary and non-pecuniary incentives of government jobs attract more or less of these dimensions?
- Run an experiment inside Mexico’s Regional Development Program to ask

1. Do higher wages attract higher quality applicants?
2. Do higher wages allow the government to hire more people?
3. How do job location disadvantages affect recruitment?

Dal Bó et al 2013: Context

- 2011: Regional Development Program (RDP) launched to increase state presence
- Create network of 350 community development agents & 50 coordinators to identify needs in the community and report to federal government to get resources to meet demands.
- Program implemented in 10 regions, 167 municipalities, 000s of localities.
- Each community development position is assigned to a particular municipality, selected on an index of socioeconomic characteristics.

Dal Bó et al 2013: Model

- Employer faces a residual labor supply of large number of workers.
- 2 periods, no discounting.

1. Worker decides whether to incur cost $c > 0$ of attending interview for job with posted wage w
2. If worker interviewed, receives offer w/pr $\rho \in (0, 1]$

- Individuals differ in 2 ways:

1. Market quality $v \in [0, \infty)$ with distribution $F(v)$
2. Public Sector Motivation (PSM) $\pi \in [0, \infty)$

- Utility from getting the job is $w + \pi$
- In period 2, accept if better than market opportunity with value $v + \varepsilon$.
  - Shock $\varepsilon \sim G(\varepsilon)$ realized before getting the offer

Dal Bó et al 2013: Model

- Solve the model backwards: Workers will accept if $v + \varepsilon \leq w + \pi$. This happens w/pr $G\left(w + \pi - v\right)$
- Action in the entry decision will be driven by relationship between π and v. 2 cases:
  1. π and v independent
  2. $v = m\left(\pi\right)$
- Assume ρ is fixed (at 1). In particular, doesn't depend on type of the worker.

Dal Bó et al 2013: Effects on Applicant Pool

PROPOSITION 1:

1. Given the assumptions of our model, an increase in wages increases the size and average quality of the applicant pool.
2. In the case where PSM and quality are independent in the population, an increase in wages decreases the average PSM of the applicant pool.
3. In the case where PSM and quality are positively correlated according to the function $m(\pi)$, an increase in wages increases the average PSM of the applicant pool.

Dal Bó et al 2013: Effects on Recruitment

- Simplify: $\pi = 0$. $\Rightarrow \exists \bar{v}$ type and all $v \leq \bar{v}$ attend the interview, $v > \bar{v}$ do not. $\partial \bar{v} / \partial w > 0$
- Overall labor supplys at wage w is

$$S = \int_0^{\bar{v}} G(w - v) f(v) \, dv = \underbrace{F(\bar{v})}_{\text{applicant pool}} \underbrace{\int_0^{\bar{v}} G(w - v) \frac{f(v)}{F(\bar{v})} dv}_{\gamma\text{: conversion rate}}$$

PROPOSITION 2:

1. The measure of recruited candidates in our model, and the labor supplied to the firm in equilibrium, can be written as the product $F(\bar{v}) \gamma(w, \bar{v})$ of the applicant pool size and the conversion rate. Thus, the elasticity of labor supply facing the employer is $\eta = \frac{dS}{sw} \frac{w}{S} = \frac{dF}{dw} \frac{w}{F} + \frac{d\gamma}{dw} \frac{w}{\gamma} \equiv \xi_{F(\bar{v})} + \xi_{\gamma(w, \bar{v})}$
2. Under the assumptions of the model, the elasticity of the applicant pool size is positive $\left(\xi_{F(\bar{v})} > 0\right)$ while the sign of the elasticity of the conversion rate $\xi_{\gamma(w, \bar{v})}$ is ambiguous

Dal Bó et al 2013: Experiment

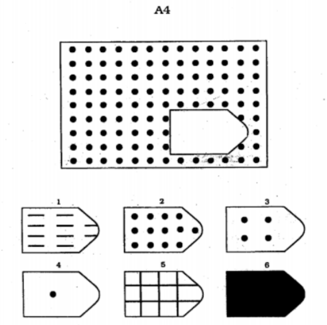
- Recruitment happened June - August 2011.
- recruitment sites: 113 schools in 106 localities in 10 target regions received job postings.
- Job ad provided general description of the job, and a toll-free number and email address.
- Respondents gave contact info and basic educ and employment history. Then, based on where they saw the ad, told the salary and a date/place for the interview.
- 1,920 individuals registered. 1,665 by phone, 208 by email, 47 did both.
- Experiment 1: Salaries randomly allocated across recruitment sites.
  - 65 localities posted 5,000 pesos/month (~US$500)
  - 41 localities posted 3,750 pesos
- Estimate effect on individual i applying in locality c in region r:

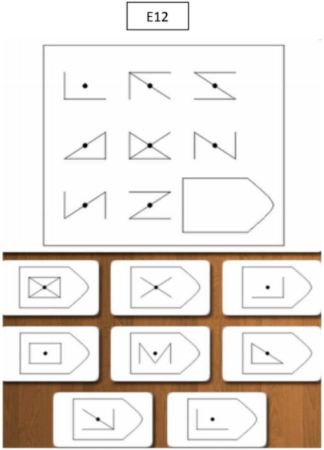
$$Y_{icr} = \beta_1 T_c + \zeta_r + \epsilon_{icr}$$

	Low-wage offer (1)	High-wage offer (2)	Difference (3)	Randomization inference p-value (4)
Latitude	19.359	19.754	0.396	.52
			[0.548]	
Longitude	99.088	100.136	1.048	.20
			[0.796]	
Altitude (m)	732.450	898.242	165.792	.30
			[158.536]	
Population (logs)	9.219	9.373	0.154	.67
			[0.331]	
Number of households (logs)	7.825	7.971	0.145	.71
			[0.335]	
Share of population between 15–65 years old	0.620	0.624	0.004	.56
			[0.007]	
Share of male population	0.480	0.482	0.002	.49
			[0.003]	
Share of indigenous population	0.275	0.160	−0.115	.05
			[0.061]*	
Illiteracy rate (% of illiterate among 15-year-olds and older)	0.104	0.096	−0.008	.45
			[0.013]	
Average years of schooling	8.335	8.251	−0.084	.75
			[0.249]	
Number of live births per woman	2.517	2.518	0.001	.99
			[0.081]	
Employment rate	0.965	0.960	−0.004	.88
			[0.008]	
Share of female-headed households	0.275	0.265	−0.009	.28
			[0.008]	
Share of households with access to electricity, water, and sanitation	0.715	0.756	0.040	.41
			[0.049]	
Share of households with a dirt floor	0.106	0.111	0.005	.80
			[0.018]	
Number of observations	41	65		

Dal Bó et al 2013: Experiment

- During the interview, applicants did a 3-hour exam measuring personal characteristics (more below).
- Raven matrices used to measure IQ





. . .

Dal Bó et al 2013: Experiment

- Scores below 7 deemed ineligible for job.
- 4 strata:
  1. high wage & high IQ (score above 9)
  2. high wage & normal IQ (score 7-9)
  3. low wage & high IQ
  4. low wage & normal IQ
- The 350 vacancies randomied to one of these strata and then offers made randomly to applicants in that stratum (sometimes conditioning on knowing the local indigenous language as well)
- Analyze acceptance A_{ics} by individual i in locality c and stratum s

$$A_{ics} = \gamma_1 T_c + X_i' \beta + \zeta_s + \epsilon_{ics}$$

- and effect of municipality characteristics (since individuals randomly assigned to municipalities)

$$A_{ics} = \gamma_1 T_c + \gamma_2 \left(T_c \times W_m \right) + \gamma_3 W_m + X_i' \beta + \zeta_s + \epsilon_{icms}$$

Dal Bó et al 2013: Personality data

1. Aptitude data
   - Outside wages: current/previous earnings.
   - Ravens matrices (above)
   - "Big 5" personality traits: openness to experience, conscientiousness, extraversion, agreeableness, neuroticism.
2. Public Service Motivation (PSM)
   2.1 Perry 1996 scale to elicit opinions of attractiveness of politics, public service, and prosocial activities.
   2.2 6 modules: Attraction to Policy Making, Commitment to Policy Making, Social Justice, Civic Duty, Compassion, Self-Sacrifice

Summary Statistics of Candidate Pool

	Observations (1)	Mean (2)	Standard deviation (3)	p10 (4)	p50 (5)	p90 (6)
Panel A: Sociodemographic characteristics						
Male	2,244	0.60	0.49	0.00	1.00	1.00
Age	2,231	27.34	6.89	20.00	26.00	37.00
Height	2,191	1.63	0.10	1.50	1.63	1.76
Indigenous	2,253	0.40	0.49	0.00	0.00	1.00
Wage in previous job	1,584	4,276.18	3,078.61	1,300.00	3,800.00	8,000.00
Previous job was white collar	1,784	0.30	0.46	0.00	0.00	1.00
Currently employed	2,250	0.14	0.34	0.00	0.00	1.00
Currently attending school	2,252	0.18	0.39	0.00	0.00	1.00
Years of experience in past 3 spells	2,237	1.36	2.45	0.00	0.25	4.00
Has work experience	2,237	0.56	0.50	0.00	1.00	1.00
Panel B: Aptitudes and skills						
Raven's score	2,254	8.77	2.69	5.00	9.00	12.00
Years of schooling	2,223	14.45	2.45	12.00	16.00	16.00
Chose dominated risk option	2,238	0.39	0.49	0.00	0.00	1.00

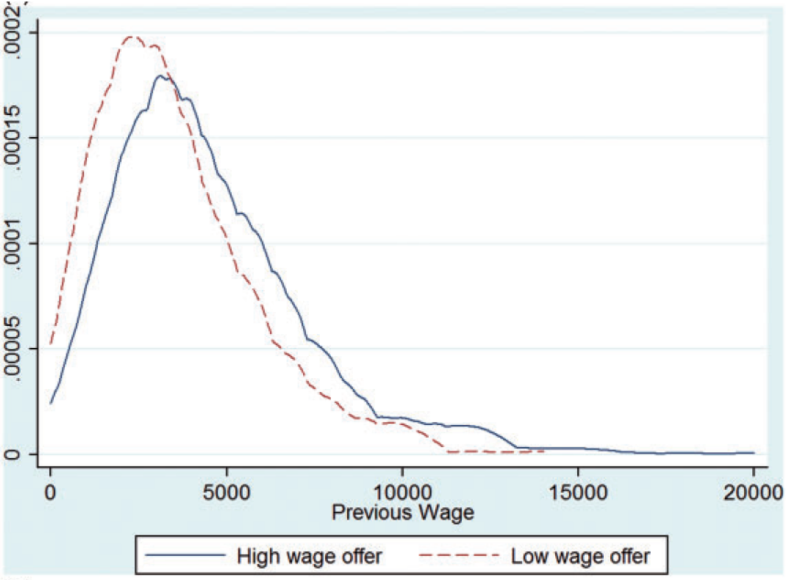
	Observations (1)	Mean (2)	Standard deviation (3)	p10 (4)	p50 (5)	p90 (6)
Panel C: Personality traits						
Conscientiousness	2,188	4.28	0.47	3.67	4.33	4.89
Extraversion	2,206	3.67	0.55	3.00	3.63	4.38
Openness	2,193	3.93	0.49	3.30	4.00	4.60
Agreeableness	2,214	4.11	0.43	3.56	4.11	4.67
Neuroticism	2,216	2.19	0.53	1.50	2.13	2.88
Big 5 Index	2,120	0.05	0.73	−0.87	0.09	0.96
Integrity: indirect measure	2,206	45.01	22.62	13.33	46.67	75.00
Integrity: direct measure	2,248	0.06	0.24	0.00	0.00	0.00
Panel D: Public service motivation						
Commitment	2,195	3.35	0.59	2.57	3.29	4.14
Social justice	2,204	3.70	0.57	3.00	3.80	4.40
Civic duty	2,183	3.94	0.63	3.14	4.00	4.71
Compassion	2,192	3.05	0.55	2.38	3.00	3.88
Self-sacrifice	2,192	3.72	0.61	3.00	3.75	4.50
Attraction	2,242	2.86	0.59	2.00	2.80	3.60
PSM index	2,096	0.07	0.72	−0.79	0.05	1.01

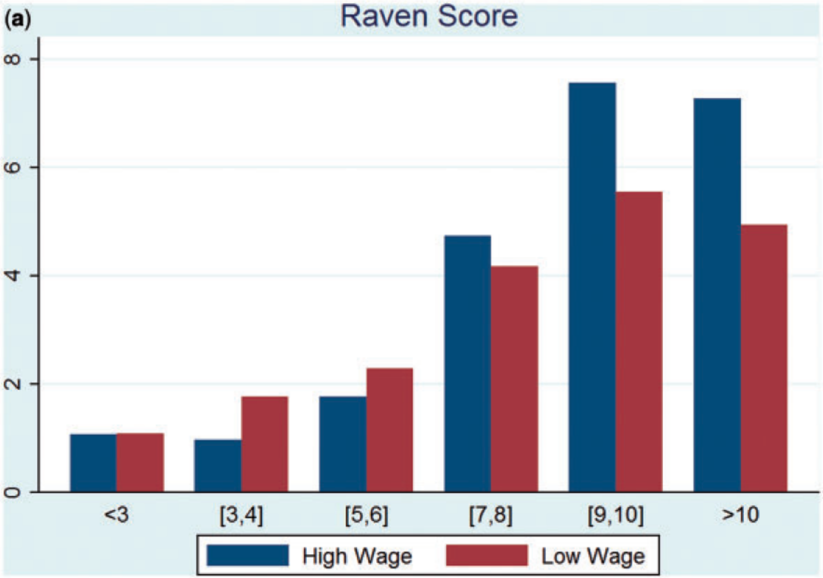
	Observations (1)	Mean (2)	Standard deviation (3)	p10 (4)	p50 (5)	p90 (6)
Panel E: Prosocial behavior						
Volunteered in the past year	2,249	0.71	0.46	0.00	1.00	1.00
Did charity work in the past year	2,248	0.54	0.50	0.00	1.00	1.00
Voted in last election	2,250	0.76	0.43	0.00	1.00	1.00
Belongs to a political party	2,250	0.10	0.30	0.00	0.00	0.00
Altruism	2,223	23.52	7.34	20.00	25.00	25.00
Negative reciprocity	2,231	0.55	0.50	0.00	1.00	1.00
Cooperation	2,182	26.40	10.71	10.00	25.00	40.00
Importance of wealth	2,048	3.22	1.38	1.40	3.20	5.20

Effects on Financial Incentives on Applicant Pool: Productive Attributes

	Observations (1)	Control (2)	Treatment effect (3)	Randomization inference p-value (4)	FDR q-value (5)
Number of applicants	106	18.093	4.714 [4.430]	.36	n/a
Panel A: Market skills					
Wage in previous job	1,572	3479.667	819.154 [174.703]***	.00	0.00
Previous job was white collar	1,170	0.243	0.069 [0.029]***	.01	0.02
Currently employed	2,225	0.104	0.053 [0.019]***	.01	0.02
Has work experience	2,212	0.459	0.167 [0.048]***	.00	0.00
Years of experience in past 3 spells	2,212	1.185	0.284 [0.171]	.08	0.06
IQ (Raven test)	2,229	8.488	0.506 [0.223]**	.01	0.02
Raven score ≥ 9	2,229	0.572	0.091 [0.039]**	.01	0.02
Chose dominated risk option	2,213	0.431	−0.064 [0.025]**	.01	0.02
Years of schooling	2,198	14.552	0.091 [0.308]	.40	0.14

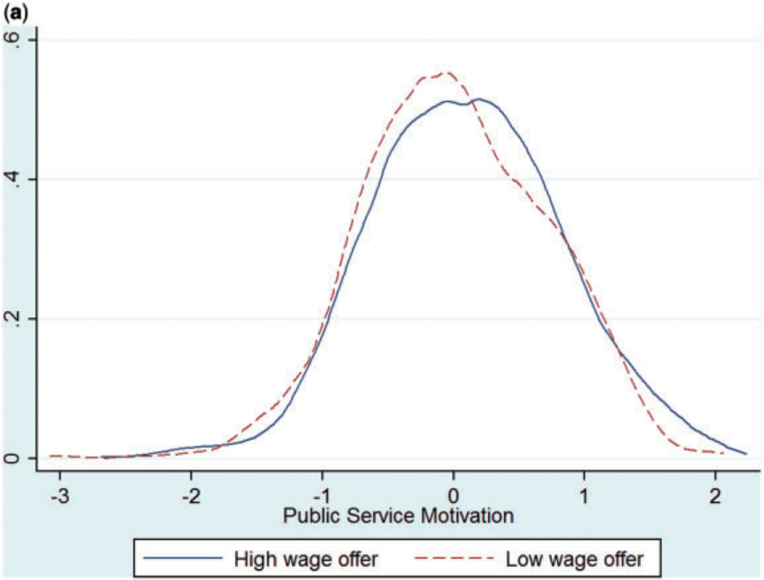
	Observations (1)	Control (2)	Treatment effect (3)	Randomization inference p-value (4)	FDR q-value (5)
Panel B: Personality traits					
Extraversion	2,189	3.674	0.013 [0.036]	.37	0.14
Agreeableness	2,167	4.107	0.004 [0.022]	.44	0.15
Conscientiousness	2,191	4.235	0.063 [0.030]**	.03	0.04
Neuroticism	2,168	2.254	−0.099 [0.033]***	.01	0.02
Openness	2,168	3.910	0.042 [0.028]	.08	0.06
Big 5 index	2,099	0.000	0.087 [0.049]*	.07	0.06
Integrity: direct	2,223	0.067	−0.009 [0.013]	.73	0.26
Integrity: indirect	2,099	44.424	0.602 [1.232]	.33	0.14


.0002
.00015
.0001
.00005
0
0
5000
10000
15000
20000
Previous Wage
High wage offer
Low wage offer


(a)
Raven Score
0
2
4
6
8
<3
[3,4]
[5,6]
[7,8]
[9,10]
>10
High Wage
Low Wage

	Observations (1)	Control (2)	Treatment effect (3)	Randomization inference p-value (4)	FDR q-value (5)
Panel A: PSM traits					
PSM index	2,074	0.000	0.092	.05	0.09
			[0.046]**		
Attractiveness	2,217	2.803	0.070	.05	0.14
			[0.041]*		
Commitment	2,170	3.316	0.045	.15	0.18
			[0.035]		
Social justice	2,180	3.646	0.075	.01	0.04
			[0.026]***		
Civic duty	2,158	3.924	0.027	.25	0.22
			[0.033]		
Compassion	2,168	3.001	0.066	.04	0.14
			[0.031]**		
Self-sacrifice	2,168	3.687	0.039	.15	0.18
			[0.034]		

Panel B: Prosocial behavior					
Altruism	2,199	23.491	0.039	.53	0.29
			[0.291]		
Negative reciprocity	2,206	0.508	0.075	.00	0.00
			[0.023]***		
Cooperation	2,157	26.174	0.675	.08	0.16
			[0.404]*		
Did charity work in the past year	2,223	0.605	−0.096	.01	0.05
			[0.041]**		
Volunteered in the past year	2,224	0.710	−0.006	.38	0.34
			[0.027]		
Importance of wealth	2,025	3.159	0.107	.14	0.18
			[0.087]		
Belongs to a political party	2,225	0.113	−0.026	.07	0.16
			[0.014]*		
Voted	2,225	0.758	0.019	.33	0.26
			[0.035]		


(a)
.6
.4
.2
0
-3
-2
-1
0
1
2
Public Service Motivation
High wage offer
Low wage offer

The Effects of Financial Incentives on Recruitment

	Accepted (1)	Accepted (2)	Rejected (3)	Not reachable (4)
High-wage offer	0.151	0.160	−0.017	−0.135
	[0.054]***	[0.054]**	[0.034]	[0.054]***
Characteristics				
Male		−0.080		
		[0.058]		
Years of schooling		−0.022		
		[0.008]**		
High IQ		−0.017		
		[0.053]		
Wage in previous job > 5,000 pesos		−0.007		
		[0.067]		
Big 5 index		0.042		
		[0.038]		
PSM index		−0.024		
		[0.044]		
Mean of dependent variable	0.55	0.55	0.13	0.32
Observations	350	343	350	350
R-squared	0.10	0.12	0.09	0.13

Effects of Municipal Characteristics on Acceptance Decisions

	Acceptance				
Dependent variable	(1)	(2)	(3)	(4)	(5)
High-wage offer	0.047	0.066	0.075	0.056	0.147
	[0.037]	[0.047]	[0.053]	[0.041]	[0.047]***
High wage offer × Distance	0.026			0.028	0.010
	[0.005]***			[0.007]***	[0.007]
Distance	−0.027			−0.028	−0.017
	[0.004]***			[0.007]***	[0.005]***
High wage offer × Drug-related		0.078		−0.033	−0.025
deaths per 1,000 inhabitants		[0.039]*		[0.037]	[0.040]
Drug-related deaths		−0.107		0.000	−0.022
per 1,000 inhabitants		[0.044]**		[0.045]	[0.040]
High wage offer × Human			−1.482	−0.913	−1.368
Development Index			[0.738]**	[0.645]	[0.589]**
Human Development Index			1.526	1.044	1.002
			[0.673]**	[0.598]*	[0.521]*
Observations	238	238	238	238	348
R-squared	0.25	0.21	0.2	0.26	0.18

Outline

Recruitment & Selection

- Dal Bó, Finan & Rossi (QJE 2013) *Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service*
- Ashraf, Bandiera & Lee (2018) *Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services*
- Leaver, Ozier, Serneels & Zeitlin (AER 2021) *Recruitment, Effort and Retention Effects of Performance Contracts for Civil Servants: Experimental Evidence from Rwandan Primary Schools*
- Xu (AER 2018) *The Costs of Patronage: Evidence from the British Empire*

Ashraf et al 2018: Introduction

- Over the course of the economic development of a nation, informal service providers are replaced with qualified, career professionals.
- Create and foster an identity based on a career as a civil servant.
- But does this displace providers whose identity is based on altruism towards beneficiaries?
- If so, how does this affect the quality of service delivery and the welfare of citizens?
- This paper: an experiment with the government of Zambia reruiting nurses to test these questions.

Ashraf et al 2018: Context

- Context: Delivery of health services in remote rural areas of Zambia
- Trained staff reluctant to be posted in remote areas so turnover is high
- Government created Community Health Assistant (CHA) post. Effectively formalizing informal community health workers.
- CHAs' main task is to visit households and refer them to health facilities
- CHAs join the civil service and can advance to higher-ranked and better paid (nurses paid about double for e.g.) cadres

Ashraf et al 2018: Experiment

- Hard to separate selection effect (who becomes a CHA) from the incentive effect (promotion incentives to perform well)
- Design an experiment to separate the two things

1. Selection: Use recruitment posters posted in public spaces. Randomly vary content of poster:
   1.1 Some posters emphasize career benefits of the job
   1.2 Othes emphasize community service
2. Shut down incentives: All successful applicants offered the same career opportunities on the job. All train together, all get the same information. $\Rightarrow$ any performance differences are only due to how they were selected.

REPUBLIC OF ZAMBIA
MINISTRY OF HEALTH

DESIGNATED HEALTH CENTRE:	FOR POSTING AT:

TRAINING OPPORTUNITY

ONE-YEAR COURSE IN COMMUNITY HEALTH

The Ministry of Health of the Republic of Zambia is launching a new national Community Health Worker (CHW) strategy and invites applicants to participate in the inaugural training of community health workers.

The training will begin on 30th **August 2010** and will be held at the Provincial level for selected applicants. All participation costs, including transportation, meals and accommodation will be covered by the Ministry of Health.

BENEFITS:

- Become a highly trained member of Zambia's health care system
- Interact with experts in medical fields
- **Access future career opportunities including:**
  - **Clinical Officer**
  - **Nurse**
  - **Environmental Health Technologist**

QUALIFICATIONS:

- Zambian National
- Grade 12 completed with two "O" levels
- Age 18-45 years
- Endorsed by Neighborhood Health Committee within place of residence
- Preference will be given to women and those with previous experience as a CHW

APPLICATION METHOD:

Submit to the **DESIGNATED HEALTH CENTRE** indicated above:

- Completed application form with necessary endorsements. If no blank forms are attached to this notice, kindly obtain a blank one at the nearest health centre.
- Photocopy of school certificate documenting completion of Grade 12 and two "O" levels.
- Photocopy of Zambian national registration card.

For more information: Contact the designated health centre indicated above.

CLOSING DATE: 30th JULY 2010.
Only shortlisted candidates will be contacted for interview.

REPUBLIC OF ZAMBIA
MINISTRY OF HEALTH

DESIGNATED HEALTH CENTRE:	FOR POSTING AT:

TRAINING OPPORTUNITY

ONE-YEAR COURSE IN COMMUNITY HEALTH

The Ministry of Health of the Republic of Zambia is launching a new national Community Health Worker (CHW) strategy and invites applicants to participate in the inaugural training of community health workers.

The training will begin on 30th **August 2010** and will be held at the Provincial level for selected applicants. All participation costs, including transportation, meals and accommodation will be covered by the Ministry of Health.

BENEFITS:

- Learn about the most important health issues in your community
- Gain the skills you need to prevent illness and promote health for your family and neighbors
- Work closely with your local health post and health centre
- Be a respected leader in your community

QUALIFICATIONS:

- Zambian National
- Grade 12 completed with two "O" levels
- Age 18-45 years
- Endorsed by Neighborhood Health Committee within place of residence
- Preference will be given to women and those with previous experience as a CHW

APPLICATION METHOD:

Submit to the **DESIGNATED HEALTH CENTRE** indicated above:

- Completed application form with necessary endorsements. If no blank forms are attached to this notice, kindly obtain a blank one at the nearest health centre.
- Photocopy of school certificate documenting completion of Grade 12 and two "O" levels.
- Photocopy of Zambian national registration card.

For more information: Contact the designated health centre indicated above.

CLOSING DATE: 30th JULY 2010.
Only shortlisted candidates will be contacted for interview.

Ashraf et al 2018: Model

- Applicants choose between CHA and their outside option.
- Material benefits of being a CHA are M
- Individuals differ in ability a and social preferences towards the community $s \in [0, 1]$
- If work as CHA, produce health $H(a)$ valued at $sH(a) \Rightarrow U(s_i H(a_i), M)$
- Applying requires paying cost c. Probability of getting job is $p(a)$ (concave).
- Outisde option is $V(a)$ where $V_a > U_a > 0 \; \forall a$
- Apply iff

$$E(a, s) = p(a) U(a, s) - c > V(a)$$

Ashraf et al 2018: Model

- Assume that $E(0) < V(0)$. Then there's a threshold $\underline{a}$ such that $a_i < \underline{a}$ don't apply. If curvatures of U and V suitable, then there's another threshold where $E(\bar{a}) = V(\bar{a})$ and only $\underline{a} < a_i < \bar{a}$ apply. Focus on this case.
- What happens when we increase M?
  1. Higher ability people apply now
  2. Low ability applicants might be discouraged since less likely to get the job: $\partial^2 p / \partial a \partial M > 0$

RESULT 1: *Increasing material benefits M will attract higher ability applicants who would not apply othersiwe $(\partial \bar{a} / \partial M > 0)$ and either (i) lower the ability of the lowest ranked applicant $(\partial \underline{a} / \partial M < 0)$ and increase the total number of applicants, or (ii) discourage low ability applicants $(\partial \underline{a} / \partial M > 0)$ and have an ambiguous impact on the total number of applicants.*

Ashraf et al 2018: Model

- Since $V_a > U_a$ the threshold $\bar{a}(s_i)$ will be increasing in s.
- Similarly $\underline{a}(s_i)$ is decreasing in s

$\Rightarrow$ Talent and pro-sociality are positively correlated for highest ability applicants, and negatively correlated for lowst ability applicants, even when they're not correlated in the population.

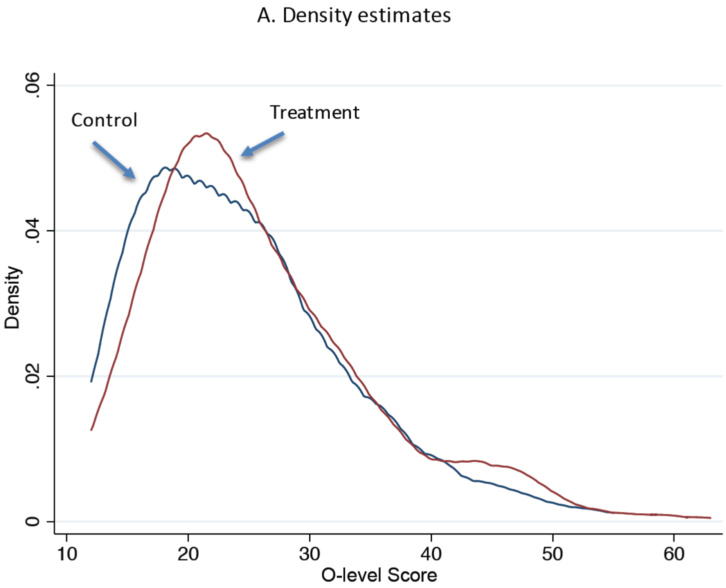
RESULT 2: *Under any M, the most able applicant is also the most prosocial. An increase in M leaves the prosociality of the marginal applicant unchanged and has an ambiguous effect on the prosociality of the average applicant*

Ashraf et al 2018: Applicants and Selection

- Recruitment drive $\Rightarrow$ 2,457 applications (average 7.3 for each position). 1,804 met eligbility requirements, 1,585 came for interview and answered questionnaire.
- Use grade 12 (o-level) exam scores as ability measure
- For non-cognitive skills, use career ambition (“what will you be doing in 5 years’ time”=”higher position in the Ministry”)
- For pro-sociality, use self-reported willingness to stay in the community and “Inclusion of Others in Self (IOS)” scale
- Selection panels chose 2 candidates to serve as CHA in the posts. Panels = district health official, health post’s health center representatice, 3 members of the neighborhood health committee.

$$s_{ih} = \sum_{j \in J} \alpha_j^c C_h X_i^j + \sum_{j \in J} \alpha_j^s \left(1 - C_h\right) X_i^j + \sum_{j \in J} \beta_j \bar{X}_h^j + \gamma N_h + \zeta_{ih}$$

where $s_{ih} = 1$ if selected, $C_h = 1$ for treatment posts, $X_i^j = 1$ if candidate i in top 3 of trait j, N_h is number of interviewees.


A. Density estimates
Control
Treatment
Density
.06
.04
.02
0
10
20
30
40
50
60
O-level Score

B. Quantile Treatment Effects on O-level Score

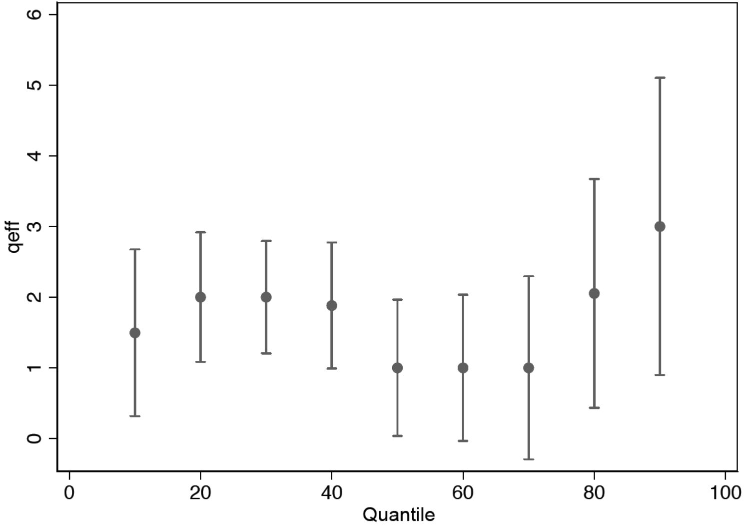


Table 1: The effect of career opportunities on the applicant pool

	treatment	control	p-value
Applicants per health post	**9**	**10**	**.228**
Cognitive skills (O-levels total exam score)	24.8	23.3	.019
Cognitive skills (number of science O-levels)	1.44	1.24	.006
Career motivation	0.25	0.19	.026
Pro-sociality	2.33	2.51	.236
farmer	.714	.683	.408
age	25.7	26.1	.446
female	.291	.304	.787

Table 2: The effect of career opportunities on candidates selection by panels

	=1 if selected	p-value	=1 if selected	p-value
=1 if top 3 in skills X treatment	0.121***		0.158***	
	(0.0287)	0.98	(0.0351)	0.54
=1 if top 3 in skills X control	0.122***		0.128***	
	(0.0374)		(0.0391)	
=1 if top 3 pro-sociality X treatment	0.0952**		0.0810*	
	(0.0386)	0.4	(0.0408)	0.63
=1 if top 3 pro-sociality X control	0.0576*		0.0560	
	(0.0291)		(0.0341)	
=1 if aims to higher rank X treatment	0.0973**		0.0933**	
	(0.0410)	0.57	(0.0378)	0.59
=1 if aims to higher rank X control	0.0698**		0.0654*	
	(0.0311)		(0.0335)	
=1 if connected to village leader X treatment			0.00983	
			(0.0389)	0.68
=1 if connected to village leader X control			0.0283	
			(0.0267)	
=1 if connected to health center staff X treatment			-0.0383	
			(0.0676)	0.62
=1 if connected to health center staff X control			-0.0009	
			(0.0395)	
Adjusted R-squared	0.106		0.110	
N	1519		1282	

Ashraf et al 2018: Service Delivery

- CHAs are required to keep records of all their visits, and issue receipts to households that they visit.
- CHAs send summary information by text message to Ministry of Health.
- At the end of each month each health post sends aggregated data to the ministry.
- Conduct validation visits to households.

$$v_{ihdp} = \alpha + \beta C_{id} + Z_h \gamma + \delta E_d + \rho_p + \epsilon_{ihdp}$$

where v_{ihds} is the number of visits CHA i in catchment area h in district d and province p does. C_{id} is treatment

Table 4: The effect of career opportunities on the number of visits

dependent variable	Household visits			
source	SMS receipts	SMS receipts	SMS receipts	SMS receipts
time horizon	months 1-18	months 1-6	months 7-12	months 13-18
unit of observation	CHA	CHA	CHA	CHA
	(1)	(2)	(3)	(4)
Treatment	93.95**	33.93**	29.56**	30.46**
	(37.19)	(15.97)	(13.49)	(12.92)
Area characteristics	Yes	Yes	Yes	Yes
Mean of dependent variable in control	318.6	167.1	92.1	59.8
Adjusted R-squared	0.112	0.115	0.064	0.105
N	307	307	307	307

Table 5: Compensation mechanisms

dependent variable	retention	visit duration	no of women and children visited per HH	no of unique HHs visited	no of visits per HH	community mobilization meetings	patients seen at health post	emergency calls
source	SMS receipts	SMS receipts	HMIS records	SMS receipts	SMS receipts	HMIS records	HMIS records	Time use survey
unit of observation	CHA	CHA	health post	CHA	CHA	health post	health post	CHA
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.0469	0.265	0.0437	36.35**	0.488*	17.06***	31.79	0.0469
	(0.0582)	(1.850)	(0.0947)	(15.49)	(0.246)	(5.220)	(260.4)	(0.0582)
Area characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dependent variable in control	0.796	33.9	2.06	179.4	1.817	20.32	1126.6	0.457
Adjusted R-squared	0.041	0.011	0.006	0.121	0.125	0.072	0.027	0.002
N	307	307	142	307	307	146	146	298

Notes: OLS Estimates, standard errors clustered at the district level. Treatment=1 if the health worker is recruited in a district where career opportunities were made salient. Retention=1 if the CHA still reports visits after 1 year. Visit duration is computed as end time minus start time in minutes. Emergency calls=1 if the CHA takes at least 1 out of hours call in a typical week. SMS receipts are sent by individual CHAs to MOH for each visit. The Health Management and Information System (HMIS) is the Zambian Ministry of Health's system for reporting health services data at government facilities. The two CHAs are required to submit monthly reports that summarize their activities at the health post/community level. The number of observations varies because some health posts do not submit the reports; these are equally distributed between treatments. The time use survey was administered in May 2013 during a refresher training program. All regressions include the stratification variables (province dummies and share of high school graduates in the district). Area characteristics include: number of staff in the health post, geographical distribution of households in the catchment area, and an indicator variable that equals 1 if the CHA reports to have good cell network coverage most of the time or all the time.

Ashraf et al 2018: Service Utilization

- The treatment increased health staff inputs. Does it also increase households' use of facilities? Health outcomes?
- Admin data on household visits to facilities

$$y_{hdpt} = \alpha + \beta C_{hd} + \gamma A_t + \delta C_{hd} \times A_t + Z_h \theta + E_d \phi + \rho_p + \xi_{hdpt}$$

where y_{hdpt} is the outcome at facility h in district d and province p in quarter t. A_t denotes quarters after the arrival of the CHAs (2012Q4).

Table 6: The effect of career opportunities on facility utilization

Dependent variable: total over each quarter 2011:1-2014:2	institutional deliveries	postnatal (0-6 weeks) visits	children under 5 visited	children under 5 weighed	children under 1 receiving BCG vaccinations	children under 1 receiving polio vaccinations	children under 1 receiving measles vaccinations	average standardized effect
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.134	-12.75	-65.96	-73.05	10.99	-0.374	1.707	-0.005
	(10.37)	(9.435)	(142.9)	(133.5)	(11.97)	(9.145)	(10.01)	(0.156)
After	4.408	15.47***	61.71	108.7*	-1.270	-1.177	-1.167	0.043
	(4.253)	(5.096)	(62.82)	(63.33)	(4.540)	(3.701)	(3.553)	(0.059)
Treatment*After	13.97**	7.919	312.0***	277.9**	7.147	14.65***	11.19	.277***
	(6.242)	(9.467)	(97.24)	(109.2)	(8.838)	(4.802)	(7.229)	(0.092)
Area characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dependent variable in control in year 1	46.7	49.9	1312.8	1261.5	89.8	73.9	73.6	na
Adjusted R-squared	0.353	0.213	0.253	0.253	0.151	0.151	0.118	na
Number of facilities	89	118	123	123	121	120	121	na
Number of observations	1268	1529	1618	1610	1518	1530	1535	1097

Notes: OLS Estimates, standard errors clustered at the district level. Treatment=1 if the health worker is recruited in a district where career opportunities were made salient. Data source is the Health Management and Information System (HMIS) available monthly from January 2011 until June 2014. Health center and health post staff are required to submit monthly reports that summarize their activities at the health post/community level. These are aggregated at the quarter level in the regressions. The variable in Column (1) is defined at the health center level because health centers are equipped for child births and health posts are not. The variables in Columns (2)-(7) are defined at the health post level if this reports data, at the health center otherwise. The average standardized treatment effect is computed using the methodology in Kling et al. (2001). After=1 after September 2012 (from 2012:4 onwards), when CHAs started working. All regressions include the stratification variables (province dummies and share of high school graduates in the district). Area characteristics include: number of staff in the health post, geographical distribution of households in the catchment area, and an indicator variable that equals 1 if the CHA reports to have good cell network coverage most of the time or all the time.

Table 7: The effect of career opportunities on health practices and outcomes

	Information	Health practices				Incidence of Illness			Anthropometrics				All
Dependent variable	% of correct answers in medical knowledge test	=1 if child under 2 yr old is breastfed	=1 if child's stool are safely disposed	number of deworming treatments	=1 if child exposed to CHA is on track with immunization schedule	=1 if child has experienced fever in the last two weeks	=1 if child has experienced diarrhea in the last two weeks	=1 if child has experienced cough in the last two weeks	=1 if weight for age z score <2 SD (moderately or severely undernourished)	=1 if weight for age z score <3 SD (severely undernourished)	=1 if MUAC<12.5 (moderately or severely wasted)	=1 if MUAC<11.5 (severely wasted)	average standardized effect
Treatment	0.002	0.051**	0.121***	0.225*	0.047**	-0.003	0.037	-0.070**	-0.053*	-0.028*	-0.023	-0.014	0.108***
	(0.010)	(0.023)	(0.039)	(0.129)	(0.020)	(0.037)	(0.027)	(0.033)	(0.030)	(0.015)	(0.015)	(0.014)	(0.036)
Household controls	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Child controls	no	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Mean of dep var in control	.740	.641	.595	1.44	.058	.469	.255	.450	.210	.051	.036	.014	
Adjusted R-squared	0.057	0.561	0.161	0.263	0.024	0.077	0.017	0.021	-0.006	0.003	0.018	0.017	
N	738	613	736	659	462	731	731	731	582	582	581	581	376

Notes: OLS estimates, standard errors clustered at the district level. Treatment=1 if the health worker is recruited in a district where career opportunities were made salient. The medical knowledge test contains 14 questions on topics that CHAs are supposed to cover; these questions were drafted by the researchers in consultation with CHA program officials and the CHA curriculum. Breastfeeding and stool disposal are self-reported. In line with UNICEF (2014), we define stools as safely disposed if flushed in toilet/latrine. Deworming, immunization data and schedule are as reported in the child health card. A child is defined as on track if they have completed all immunizations required for their age in months. The immunization sample is restricted to children who were 3 months or younger (including unborn) when the CHAs started working. Thresholds for weight-for-age and MUAC are taken from WHO guidelines; following these, data are restricted to children between 6-59 months. Household controls include size, education level of the respondent, and number of assets. Child controls include age and gender. All regressions include the stratification variables. The average standardized treatment effect is computed using the methodology in Kling et al. (2001) after recoding all variables so that higher values indicate better outcomes. For weight-for-age z score and MUAC we use the lowest thresholds.

Outline

Recruitment & Selection

- Dal Bó, Finan & Rossi (QJE 2013) *Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service*
- Ashraf, Bandiera & Lee (2018) *Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services*
- Leaver, Ozier, Serneels & Zeitlin (AER 2021) *Recruitment, Effort and Retention Effects of Performance Contracts for Civil Servants: Experimental Evidence from Rwandan Primary Schools*
- Xu (AER 2018) *The Costs of Patronage: Evidence from the British Empire*

Leaver et al (2021): Motivation

- Recruiting good people, motivating and retaining them are central challenges for any organization with significant labor inputs
- Might Pay for Performance (P4P) be a good way of doing this? We know it can motivate teachers (Muralidharan & Sundararaman 2011)
- Anti:
  - crowd out pro-social mottivation
  - screen in the wrong people (Benabou & Tirole, 2003, 2006)
- Pro:
  - classic contract theory (Lazear 2003)
  - evidence from jobs with readily measurable output (Lazear 2000)
- Work in Rwanda to conduct the first RCT to test for these effects together.

Leaver et al (2021): Setting

- Study looks at the cohort of teachers applying for state school positions in 6 districts in Rwanda in 2016
- Districts group vacancies into Math and Science (TMS); Modern Languages (TML); and Social Studies (TSS).
- Think of a district $\times$ subject as a labor market $\rightarrow$ 18 labor markets in the study.
- Small N for randomization, but spans 600+ hiring lines and >60% of hires that year.

Leaver et al (2021): Experimental Design

- Experiment in two stages

1. Randomize *advertizing* (with v without P4P) across labor markets
2. Once people assigned to schools, re-randomize *actual, experienced* contract.

		Advertised	
		FW	P4P
Experienced	FW	a	b
	P4P	c	d

FIGURE 1. TREATMENT GROUPS AMONG RECRUITS PLACED IN STUDY SCHOOLS

- a vs b and c vs d show compositional effects of P4P. a vs d shows *total* effect of P4P including both margins

Leaver et al (2021): First Tier Randomization

- Randomized the advertizing of the jobs across the 18 (district $\times$ subject family) labor markets
- Advertized by radio, poster, flyer and had a person at the district education office to explain the contracts
- Held three job fairs to promote the contracts dissemination through WhatsApp
- Applications submitted in December 2015 in response to these advertisements
- Districts held exams for the candidates
- Successful candidates placeed in jobs during Feb/March 2016

Leaver et al (2021): Second Tier Randomization

- After all the new teachers had been placed, all of the schools were randomized into either P4P (85 schools) or FW (79 schools)
- All teachers received the randomly assigned contract, both new and incumbents.
- P4P contract: top 20% of teachers receive FRw 100K (\$136 $\sim$ 15% of annual salary)
- Eligibility was equally weighted average of a) "pay for percentile": Students' measured performance gains controling for student composition; b) teacher presence; c) lesson preparation; d) observed pedagogy
- FW was bonus of FR2 20K for all teachers
- Potential worry: disappointment, so offered FRw 80K "retention bonus" to people who completed the year on top of P4P/FW

Leaver et al (2021): Hypotheses

- With this design, the paper seeks to test the following hypotheses

1. Advertising P4P induces different *application* qualities
2. Advertised P4P affects the observable skills of recruits *placed* in schools
3. Advertised P4P induces differentially intrinsically motivated recruits to be *placed* in schools
4. Advertised P4P selects in higher or lower performing teachers as measured by learning outcomes
5. Experienced P4P creates incentives that affect teacher performance as measured by their students' learning outcomes
6. These selection and effort effects are apparent in the composite performance metric

Leaver et al (2021): Analysis

- As we saw above, there are $N = 18$ labor markets, so power is a real concern
- Exactly which empirical specification has the highest power to test the hypotehses above depends on features of the data that are unknown **ex ante.**
- But we don't want to allow researchers to specification hunt (*p*-hack) **ex post.**
- Solution: Blinding
  - IPA blinded the data before giving it to the researchers.
  - Researchers explored pre-specified designs in the blinded data
  - After updating PAP, researchers received true data and implemented power-maximizing designs
- Solution: Randomization Inference. Central limit theorems won't work well with $N = 18$ clusters so use RI instead (Imbens & Rubin 2015) which works even in finite samples by providing exact inference.

Leaver et al (2021): Data

- Applications: 2,184 applications (1,962 have teacher training so valid)

TABLE 1—APPLICATION CHARACTERISTICS, BY DISTRICT

	Gatsibo	Kayonza	Kirehe	Ngoma	Nyagatare	Rwamagana	All
Applicants	390	310	462	380	327	315	2,184
Qualified	333	258	458	364	272	277	1,962
Has TTC score	317	233	405	337	260	163	1,715
Mean TTC score	0.53	0.54	0.50	0.53	0.54	0.55	0.53
SD TTC score	0.14	0.15	0.19	0.15	0.14	0.12	0.15
Qualified female	0.53	0.47	0.45	0.50	0.44	0.45	0.48
Qualified age	27.32	27.78	27.23	27.25	26.98	27.50	27.33

Leaver et al (2021): Data

- Teacher attributes collected in Feb/Mar 2016 during baseline
  - Measures of intrinsic motivation: modified dictator game (Ashraf, Bandiera & Jack 2014)
  - Grading task: Grade an exam script in 5 minutes. Used to estimate a two-parameter Item Response Theory (IRT) model
- Student learning at baseline, end of 2016 and 2017 school years.
  - developed comprehensive tests together with Ministry of Education
  - Students did oral exams for 50 mins
  - Estimate 2-parameter IRT model by Empirical Bayes for each student

Leaver et al (2021): Teacher Inputs Data

1. *Presence* measured during random spot checks
2. *Preparation* measured by reviewing teachers' weekly lesson plans. Teachers coached in filling in such plans beforehand. Often missing so presence is an indicator.
3. *Pedagogy* measured through classroom observation. Rubric based on Danielson Framework for Teaching widely used in US

- Teachers' scores for P4P based on average of these 3 plus student learning

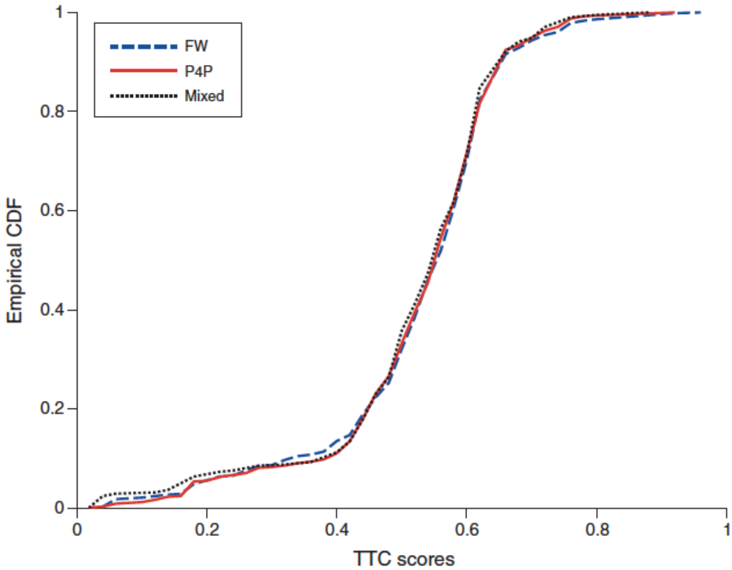
Leaver et al (2021): Composition Effects of P4P

- Use a Kolmogorov Smirnov test of the equality of distributions
- “discovered” during blinded phase

$$T^{KS} = sup_y \left| \hat{F}_{P4P}(y) - \hat{F}_{FW}(y) \right| = \max_{i=1,\dots,n} \left| \hat{F}_{P4P}(y_i) - \hat{F}_{FW}(y_i) \right|$$

- Do inference by randomization inference:
  - Calculate T^{KS} in the data.
  - Then, repeatedly sample from the set of potential treatment assignments.
  - For each assignment a calculate T_a^{KS}
  - $p-$value is the share of such tests with larger absolute value than the one calculated in the data
- Estimate $T^{KS} = 0.026$ with a $p-$value of 0.909

Leaver et al (2021): Composition Effects of P4P



Leaver et al (2021): Effects of P4P on skill and motivation

- Data on a range of attributes of teachers.
- Here focus on grading task (ability) and pro-sociality (dictator game)
- Estimate regressions of the form

$$x_{jqd} = \tau_A T^A_{qd} + \gamma_q + \delta_d + e_{jqd}$$

  where x_{jqd} is attribute of teacher j with qualification q in district d
- Test statistic: t-statistic on τ_A under the null hypothesis of no effects for any teacher. Estimated by RI.

Leaver et al (2021): Effects of P4P on skill and motivation

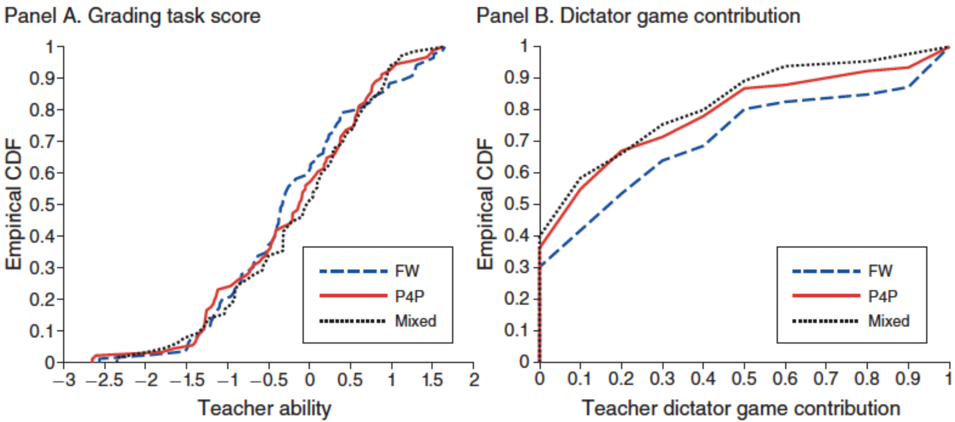


Figure 3. Distribution of Placed Recruit Attributes on Arrival, by Advertised Treatment Arm

Notes: In panel A, the t-statistic for a difference in mean grading task IRT score across the P4P and FW treatments is -0.184, with a p-value of 0.367. In panel B, the t-statistic for a difference in mean dictator game share sent across the P4P and FW treatments is -0.100, with a p-value of 0.029.

Leaver et al (2021): Student Learning

- Estimate dependence of students' results on teacher's advertised contract T_{qd}^{A} and experienced contract T_{s}^{E}

$$z_{ibksr} = \tau_A T_{qd}^{A} + \tau_E T_{s}^{E} + \lambda_I I_j + \lambda_E T_{s}^{E} I_j + \rho_{br} \bar{z}_{ksr-1} + \delta_d + \psi_r + e_{ibksr}$$

  where z_{ibksr} is performance of student i in subject b, stream k, school s and test round r; I_j is an incumbent dummy; and $\bar{z}_{ksr-1}$ is mean score in previous round for students in that subject.

- Also allow for interactions

$$z_{ibksr} = \tau_A T_{qd}^{A} + \tau_E T_{s}^{E} + \tau_{AE} T_{qd}^{A} T_{s}^{E} + \lambda_I I_j + \lambda_E T_{s}^{E} I_j + \rho_{br} \bar{z}_{ksr-1} + \delta_d + \psi_r + e_{ibksr}$$

- Estimate by mixed effects allowing for random effects at the round level.
- Calculate p-values for the τs by permuting the treatment assignments.

Leaver et al (2021): Student Learning

Table 3—Impacts on Student Learning, Linear Mixed Effects Model

	Pooled	Year 1	Year 2
Model A. Direct effects only			
Advertised P4P (τ_A)	0.01	−0.03	0.04
	[−0.04, 0.08]	[−0.06, 0.03]	[−0.05, 0.16]
	(0.75)	(0.20)	(0.31)
Experienced P4P (τ_E)	0.11	0.06	0.16
	[0.02, 0.21]	[−0.03, 0.15]	[0.04, 0.28]
	(0.02)	(0.17)	(0.00)
Experienced P4P × incumbent (λ_E)	−0.06	−0.05	−0.09
	[−0.20, 0.07]	[−0.19, 0.11]	[−0.24, 0.06]
	(0.36)	(0.54)	(0.27)
Model B. Interactions between advertised and experienced contracts			
Advertised P4P (τ_A)	0.01	−0.02	0.03
	[−0.05, 0.14]	[−0.06, 0.07]	[−0.05, 0.21]
	(0.46)	(0.62)	(0.22)
Experienced P4P (τ_E)	0.12	0.06	0.18
	[0.05, 0.25]	[−0.01, 0.19]	[0.08, 0.33]
	(0.01)	(0.10)	(0.00)
Advertised P4P × experienced P4P (τ_{AE})	−0.03	−0.01	−0.04
	[−0.17, 0.09]	[−0.15, 0.10]	[−0.22, 0.13]
	(0.51)	(0.65)	(0.58)
Experienced P4P × incumbent (λ_E)	−0.08	−0.05	−0.11
	[−0.31, 0.15]	[−0.30, 0.18]	[−0.36, 0.14]
	(0.43)	(0.56)	(0.38)
Observations	154,594	70,821	83,773

Leaver et al (2021): Effort effects

TABLE 4—ESTIMATED EFFECTS ON DIMENSIONS OF THE COMPOSITE "4P" PERFORMANCE METRIC

	Summary metric (1)	Preparation (2)	Presence (3)	Pedagogy (4)	Pupil learning (5)
Model A. Direct effects only					
Advertised P4P (τ_A)	−0.04	0.07	0.00	0.03	−0.02
	[−0.09, 0.01]	[−0.13, 0.32]	[−0.05, 0.07]	[−0.06, 0.10]	[−0.08, 0.02]
	(0.11)	(0.40)	(0.93)	(0.42)	(0.27)
Experienced P4P (τ_E)	0.23	0.02	0.08	0.10	0.09
	[0.19, 0.28]	[−0.13, 0.16]	[0.02, 0.14]	[−0.00, 0.21]	[0.03, 0.15]
	(0.00)	(0.84)	(0.01)	(0.05)	(0.00)
Experienced P4P × incumbent (λ_E)	0.03	0.07	−0.01	0.07	−0.00
	[−0.01, 0.07]	[−0.03, 0.18]	[−0.06, 0.05]	[−0.01, 0.16]	[−0.04, 0.03]
	(0.10)	(0.17)	(0.70)	(0.11)	(0.86)

Outline

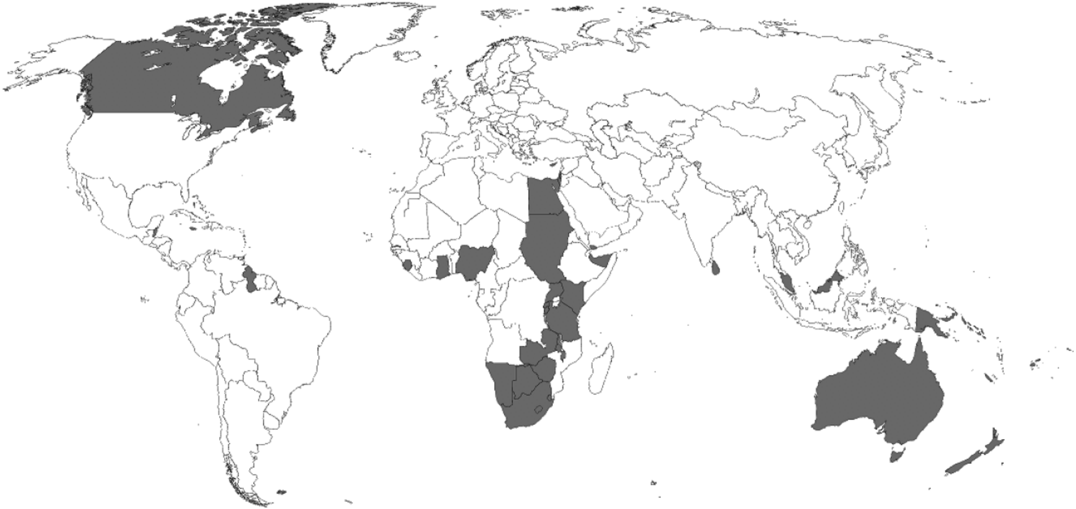
Recruitment & Selection

- Dal Bó, Finan & Rossi (QJE 2013) *Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service*
- Ashraf, Bandiera & Lee (2018) *Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services*
- Leaver, Ozier, Serneels & Zeitlin (AER 2021) *Recruitment, Effort and Retention Effects of Performance Contracts for Civil Servants: Experimental Evidence from Rwandan Primary Schools*
- Xu (AER 2018) *The Costs of Patronage: Evidence from the British Empire*

Xu 2017: Introduction

- Patronage has always been an important feature of the appointment of officials to public office.
- If principals have private information, patronage can be beneficial.
- But if it disincentivizes subordinates, it's detrimental.
- This project: Study impact of patronage on state performance in the Colonial Office of the British Empire.

Figure 1: Territories administered by the Colonial Office - 1905



Xu 2017: Context

- Study the Colonial Office. Founded in 1854 to administer "overseas possessions" At the peak of British colonialism, the bureaucracy spanned the globe, covering 1/5 of the world's land mass.
- Identifying variation comes from two sources

1. Ministerial turnover. The Colonial Office headed by the Secretary of State: a cabinet level political appointee. When the Prime Minister reshuffled the cabinet, or when there was an election, the connectedness of serving governors changed.
2. Changes in the appointment regime. The Secretary of State had discretion to appoint governors 1854-1930: the *patronage* period. In 1930 replaced by a formal system of open recruitment, the "Magna Carta of the Colonial Service".

Xu 2017: Data

- Digitized data to construct an individual-level personnel dataset on the Colonial Office

1. Colonial Lists. Data on postings, backgrounds, and salaries of governors 1860–1966
2. Blue Books 1821–1949. administrative statistics with detailed data on public finance (revenues, expenditures), demographics, trade and socio-economic statistic, and prices.
3. Genealogical data. Biographical information from UK Who-is-Who
4. Genealogical data and family tree data from The Peerage to crease a family network of the British elite. To measure connectedness.

Table 1: Descriptive characteristics of governors and British colonies

Panel A:	(1)	(2)	(3)	(4)	(5)	(6)
Governor characteristics	Pooled years		By year			
	Mean	SD	1860	1900	1930	1960
Peerage	0.085	0.280	0.047	0.157	0.027	0.000
Civil servant	0.843	0.363	0.809	0.921	0.810	1.000
Military	0.439	0.496	0.416	0.424	0.333	0.200
Politician	0.087	0.283	0.166	0.131	0.027	0.000
Eton	0.109	0.312	0.125	0.068	0.068	0.111
Oxford	0.178	0.383	0.136	0.151	0.303	0.100
Cambridge	0.150	0.358	0.103	0.171	0.242	0.600
Age at entry	48.652	8.990	41.600	46.078	50.800	48.900
Observations	456 (330)		42 (22)	38 (29)	37 (29)	10 (9)

Panel B:	(7)	(8)	(9)	(10)	(11)	(12)
Colony characteristics	Pooled years		By year			
	Mean	SD	1860	1900	1930	1960
(log) Total revenue	12.309	2.185	10.850	12.638	13.135	15.961
- Share customs revenue	0.470	0.206	0.550	0.467	0.431	0.575
(log) Total expenditure	12.333	2.166	10.879	12.551	13.236	15.964
(log) Population	11.689	1.995	10.823	12.037	12.071	13.052
(log) Governorship salary	7.928	0.795	7.739	7.961	8.078	8.877
Area tropics	0.652	0.423	0.564	0.591	0.720	0.742
(log) Distance from London	8.562	0.612	8.464	8.608	8.567	8.577
Observations	3,510 (2,595)		-	-	-	-
Number of colonies	70 (54)		42 (28)	39 (30)	37 (30)	10 (3)

Xu 2017: Measuring Connectedness

- A good measure has to be
  - objective. Hard because cannot directly observe ties
  - exogenous. If high ability people both get promoted and make ties, then ties measure ability.

1. Shared ancestry. Use family networks from family tree data as a proxy.
2. Membership of the aristocracy
3. Attendance of elite schools (Eton, Oxbridge)

Xu 2017: Effect on Salaries

- To estimate impact of social connections on governor i of colony s at time t's salary

$$\log w_{ist} = \beta c_{it} + \theta_i + x'_{it}\gamma + \tau_t + \varepsilon_{ist}$$

  where $c_{it} \in \{0, 1\}$ denotes connectedness to the Secretary of State
- Is the effect driven by increasing the salary of a given job, or by moving governors around to higher paid jobs?
  - estimate the same equation but with colony FEs
  - If it's moving, what types of moves drive salary increases?

Figure 2: Salary gap and the removal of patronage (Warren Fisher Reform 1930)

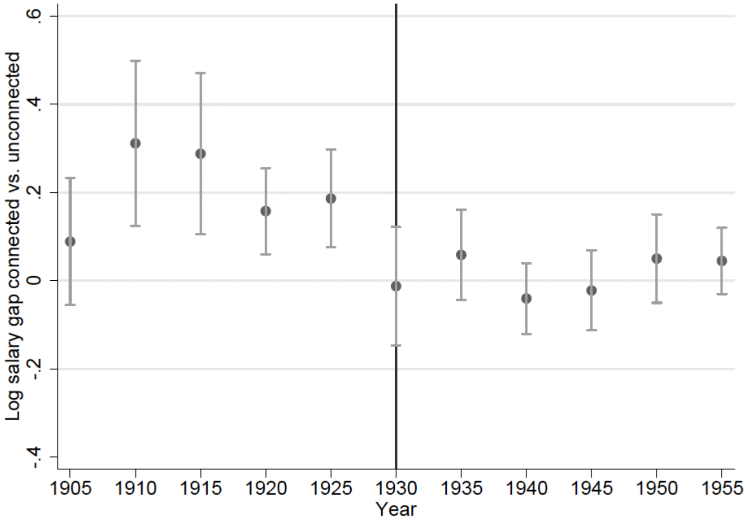


Table 3: Transfers and connectedness to Secretary of State

	(1)	(2)	(3)	(4)	(5)
			Fixed colony characteristics		
	log Governor salary (GBP)		log Initial revenue	Area in tropics	log Distance London
Mean of dep. var	7.929	7.929	10.74	0.653	8.563
No. colonies served	0.224***	0.034	0.737***	-0.017	0.063**
	(0.035)	(0.019)	(0.095)	(0.025)	(0.029)
Connected	0.098***	0.011	0.177*	0.014	-0.019
	(0.036)	(0.017)	(0.099)	(0.029)	(0.033)
Year FEs	Yes	Yes	Yes	Yes	Yes
Governor FEs	Yes	Yes	Yes	Yes	Yes
Colony FEs	-	Yes	-	-	-
Spell length FEs	Yes	Yes	Yes	Yes	Yes
Observations	3,510	3,510	3,510	3,510	3,510

Xu 2017: Effects of Abolition of Patronage

- Use the Warren Fisher reform of 1930 as a natural experiment

$$\log w_{ist} = \beta_0 c_{it} + \beta_1 c_{it} \times \mathbf{1}\{t \geq 1930\} + x'_{it}\gamma + \theta_i + \tau_t + \varepsilon_{ist}$$

Table 4: Warren Fisher 1930 - Removal of Patronage

	(1)	(2)	(3)	(4)
		Governor salary		
Mean of dep. var	7.929	7.929	7.929	7.929
Connected	0.097***	0.127***	0.205***	0.169***
	(0.036)	(0.043)	(0.059)	(0.060)
Reform dummy × Connected		-0.123**	-0.222***	-0.182**
		(0.056)	(0.079)	(0.084)
Connected + Reform dummy × Connected	-	0.004	-0.017	-0.013
		(0.040)	(0.043)	(0.048)
Year FEs	Yes	Yes	Yes	Yes
Governor FEs	Yes	Yes	Yes	Yes
Spell length FEs	Yes	Yes	Yes	Yes
Time-varying controls	Yes	Yes	Yes	Yes
Connected × Trend (centered 1930)	-	-	Yes	Yes
Reform dummy × Governor characteristics	-	-	-	Yes
Observations	3,510	3,510	3,510	3,027

Xu 2017: Performance

- If the performance of connected governors is also better, then the salary premium need not be problematic
- Estimate performance effects

$$y_{ist} = \beta c_{it} + \gamma' x_{it} + \nu_{is} + \tau_t + \varepsilon_{ist}$$

- note that the ν_{is} means use only within-appointment-spell shocks to connections (avoid problems with match-specific effects)

Table 5: Revenue performance and connectedness to Secretary of State

Panel A: Revenue	(1)	(2)	(3)	(4)
	Colony-level Public Finance			
	Public revenue			
	Overall		Trade	Internal
Mean of dep. var	12.31	12.31	11.47	11.58
Connected	-0.040**	-0.055***	-0.053**	-0.043
	(0.017)	(0.021)	(0.026)	(0.032)
Connected × Reform dummy		0.061*		
		(0.033)		
Connected + Connected × Reform dummy	-	0.005	-	-
		(0.026)		
Year FEs	Yes	Yes	Yes	Yes
Governor-Colony FEs	Yes	Yes	Yes	Yes
Spell length FEs	Yes	Yes	Yes	Yes
Time-varying controls	Yes	Yes	Yes	Yes
Observations	3,510	3,510	2,670	2,652

Panel B: Expenditure	(5)	(6)	(7)	(8)
	Public expenditure			
	Overall		Tax	Works
Mean of dep. var	12.33	12.37	9.015	10.32
Connected	-0.029	-0.042*	-0.089*	-0.107*
	(0.019)	(0.023)	(0.053)	(0.062)
Connected × Reform dummy		0.053		
		(0.034)		
Connected + Connected × Reform dummy	-	0.010	-	-
		(0.025)		
Year FEs	Yes	Yes	Yes	Yes
Governor-Colony FEs	Yes	Yes	Yes	Yes
Spell length FEs	Yes	Yes	Yes	Yes
Time-varying controls	Yes	Yes	Yes	Yes
Observations	3,510	3,510	1,742	2,588

Table 6: Tax ordinances, exemptions and connectedness to Secretary of State

	(1)	(2)	(3)	(4)	(5)
	Legislation	Broken down by ordinance type			
	ordinances	Direct tax	Customs	Exemptions	Works
Mean of dep. var	0.020	0.0105	0.0140	0.226	0.00698
Connected	0.085**	0.048	0.068**	0.202***	-0.011
	(0.037)	(0.031)	(0.031)	(0.063)	(0.019)
Connected	-0.083**	-0.051	-0.066**	-0.369***	0.013
× Reform dummy	(0.037)	(0.032)	(0.031)	(0.137)	(0.019)
Connected + Connected ×	0.001	-0.003	0.002	-0.167	0.002
Reform dummy	(0.005)	(0.004)	(0.004)	(0.125)	(0.003)
Year FE	Yes	Yes	Yes	Yes	Yes
Governor-Colony FEs	Yes	Yes	Yes	Yes	Yes
Spell length FEs	Yes	Yes	Yes	Yes	Yes
Time-varying controls	Yes	Yes	Yes	Yes	Yes
Data source	National Archives			Blue Book	N. Arch.
Observations	573	573	573	405	573

Xu 2017: Long-Run Effects

- Do connected governors cause long-term damage?
- Would like to estimate

$$y_s = \beta C_s + \gamma' x_s + \mu_{R(s)} + \varepsilon_s$$

  where $C_s = \sum_t c_{I(s,t),t}$ denotes the number of years under a connected governor, $i = I(s,t)$ returns the identity of the governor i serving in colony s at time t.
- But assignment of connected governors is endogenous.
- Construct an instrument based on
  1. Allocation rule: governorships limited to 6 years
  2. Turnover of the secretary of state
- Share of connected governors with at least 6 years of tenure when a colony becomes vacant is an instrument:

$$p_t = \frac{\sum_i \mathbf{1}\{T_{it} \geq 6\} \times c_{it}}{\sum_i \mathbf{1}\{T_{it} \geq 6\}}$$

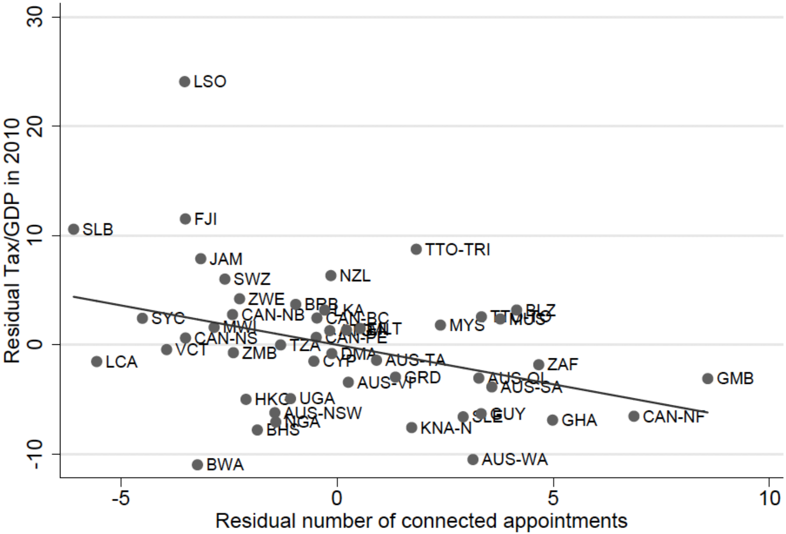
- Expected number of connected appointments is then $P_s = \sum_t p_{t-1} \times \mathbf{1}\left\{T_{I(s,t),t} = 1\right\}$

Table 8: Predicting connected appointments and years - First-stage

Panel A: Appointment level	(1)	(2)	(3)	(4)
	Connected	Connected years		
Mean of dep. var	0.306	1.442	1.436	1.423
Prob. connected appointment $t-2$				0.233
				(0.451)
Prob. connected appointment $t-1$	0.234***	0.892***	0.752**	0.808*
	(0.057)	(0.247)	(0.308)	(0.430)
Prob. connected appointment t				0.222
				(0.352)
Controls	Yes	Yes	Yes	Yes
Colony FEs	-	-	Yes	Yes
Observations	634	634	626	537

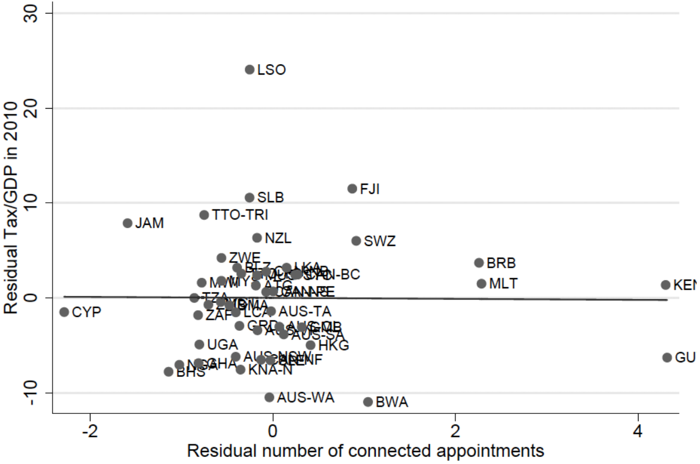
Panel B: Cross-colony level	(5)	(6)	(7)	(8)
	Total connected years			
	1854-1930		1931-1966	
Mean of dep. var	12.98	12.98	2.875	2.875
Expected # connected appointments 1854-1930	2.720***	2.739***		-0.031
	(0.539)	(0.534)		(0.272)
Expected # connected appointments 1931-1966		-0.857	3.734***	3.743***
		(3.977)	(1.340)	(1.342)
Controls	Yes	Yes	Yes	Yes
Continent FEs	Yes	Yes	Yes	Yes
Observations	48	48	48	48

Figure 4: Modern Tax/GDP and connected appointments under patronage



coef = -.72171264, (robust) se = .29592164, t = -2.44

Figure 5: Modern Tax/GDP and connected appointments in the post-patronage period



coef = -.05144496, (robust) se = .6579538, t = -.08

Table 10: Connected governors, revenue sources and the quality of tax systems in 2010

Panel A: Revenue sources	(1)	(2)	(3)	(4)	(5)
	Share of revenue (% of GDP) in 2010				
	Tax revenue	Direct tax	Indirect tax revenue		
			Total	GST	Trade
Mean of dep. var	20.62	9.897	10.64	7.473	3.258
Connected years 1854-1930	-0.427**	0.092	-0.523***	-0.117	-0.488***
	(0.187)	(0.097)	(0.153)	(0.096)	(0.135)
Connected years 1931-1966	0.426	0.220	0.010	0.164	-0.102
	(0.597)	(0.242)	(0.417)	(0.252)	(0.399)
Estimation	IV	IV	IV	IV	IV
Controls	Yes	Yes	Yes	Yes	Yes
Continent FEs	Yes	Yes	Yes	Yes	Yes
First-stage $F_{IV\,1854-30}$	26.17***	26.17***	26.17***	27.20***	27.20***
First-stage $F_{IV\,1931-66}$	8.03***	8.03***	8.03***	7.87***	7.87***
Data source	International Centre for Tax and Development (ICTD)				
Observations	48	48	48	47	47

Panel B: Quality of tax system	(6) Tariff rate	(7) # tariff lines	(8) Customs misreporting	(9) Customs hours	(10) Tax hours
Mean of dep. var	7.061	74.765	12.030	3.511	5.052
Connected years	-0.442**	4.234***	0.088***	0.060***	0.025*
1854-1930	(0.218)	(1.070)	(0.028)	(0.022)	(0.014)
Connected years	0.483	-4.730	0.005	-0.083	0.017
1931-1966	(0.299)	(3.552)	(0.061)	(0.059)	(0.053)
Estimation	IV	IV	IV	IV	IV
Controls	Yes	Yes	Yes	Yes	Yes
Continent FEs	Yes	Yes	Yes	Yes	Yes
First-stage $F_{IV\,1854-30}$	26.17***	21.59***	23.99***	25.31***	25.31***
First-stage $F_{IV\,1931-66}$	8.03***	3.79*	7.61***	7.42**	7.42**
Data source	World Integrated Trade Solution			Doing Business	
Observations	48	43	45	46	46

Outline

Theory

Financial Incentives

Non-financial Incentives

Recruitment & Selection

Open Questions

Open Questions

- How important is the policy that bureaucrats are implementing in determining their performance and who selectes into the bureaucracy?
- How should bureaucracies be organized? Hierarchically? How does corruption at one level of the bureaucracy interact with the other levels?
- What role should citizens play in monitoring the bureaucracy?
- How should bureaucracies innovate to deliver services more effectively?