Public Service Delivery, Exclusion and Externalities

Maitreesh Ghatak

(with Alex Armand, Britta Augsburg, and Antonella Bancalari)

London School of Economics

April 30, 2026

Motivation

- Improving quality of public services is a central policy objective for governments at all levels across the world
- Performance incentives are widely used to achieve this
- Because of the "public" aspect of public services, market provision is potentially:
  - Inefficient (due to externalities)
  - Inequitable (due to equity considerations relating to access)
- Yet user fees remain pervasive despite access concerns [Gertler et al., 1987; Dupas, 2014], and quality of provision is often poor

Motivation

- Public services can be provided in a host of ways, from direct government provision to provision by nonprofits/NGOs, contracting out to private providers, public-private partnerships etc
- Throws up a bunch of interesting questions relating to incentives and organizational aspects of public service delivery (see Besley and Ghatak, 2001, 2017)
- There has also been a fair bit of empirical work on various aspects of improving the quality of provision of public services (Finan, Olken, Pande, 2017)
- Today I want to talk about a RCT that was carried out in the Indian state of Uttar Pradesh, one of the poorest states, but with a population of 240 million, which would make it the fifth largest if it were a country
- The aim of the study was to see how to improve the quality and usage of community toilets using, among other things, incentives for caretakers

Motivation: This Paper

- We study what happens when direct quality incentives are provided to service providers whose fees are set exogenously
- Quality bonuses seem an obvious win when fees are held constant — but:
  - They raise demand, which requires resources
  - This creates incentives for **stricter fee enforcement**, risking user exclusion
  - Excluded users resort to inferior outside options that generate **negative externalities** (e.g. open defecation), offsetting social gains
- Setting: community toilets in slums of Uttar Pradesh, India — among the poorest states in India with a population of 240m (would be the world's fifth largest country)

Central Question

How do quality incentives interact with user fees, exclusion, and externalities?

What This Paper Does

1. **Theory:** a model of fee-funded public service delivery from the perspective of a policymaker whose objective accounts for both the social benefits of increased usage and the negative externalities of the outside option
2. **Evidence:** a large-scale field experiment in slums of two cities in UP, boosting hygiene and sanitation quality at community toilets with **no change in user fees**, tracking provider and user behaviour with rich data

Main finding: quality incentives improve service quality but induce stricter fee collection, reducing access and generating welfare-reducing externalities.

Contribution

- **State capacity and frontline incentives:** financial rewards for frontline workers [Duflo, 2012; Ashraf et al., 2014; Behrman et al., 2015; Besley and Ghatak, 2018; Caria et al., 2025], bureaucrat performance [Bandiera et al., 2021; Akhtari et al., 2022; Besley et al., 2022; Fenizia, 2022; Best et al., 2023], and state capacity to internalise externalities [Besley and Persson, 2009]
- **User fees and exclusion:** theoretical and empirical evidence that fees are regressive and exclusionary [Gertler et al., 1987; Norman, 2004; Andrabi et al., 2020; Romero et al., 2020; Dupas and Jain, 2024; Bronsoler et al., 2025] — we show similar exclusion can arise from *fee enforcement incentives* even without fee changes
- **Basic service underprovision in LMICs:** rationing due to non-payment and remedies [Guiteras et al., 2015; Jack and Smith, 2015, 2020; Burgess et al., 2020; Coville et al., 2020; Cameron et al., 2022]

Outline

1. Theoretical framework
2. Field experiment
3. Data
4. Specification
5. Results
6. Conclusions

Setup: Users and Demand

- Discrete public service; users decide whether to use it
- Fee p is fixed but *imperfectly enforced*: enforcement effort e_1 yields expected fee $\tilde{p} = pe_1$
- Quality equals maintenance effort: $q = e_2$
- Linear demand (micro-founded; intensive and extensive margins available):

$$D(e_1, e_2) = \alpha e_2 - \frac{1}{2}\beta p e_1 + \varphi$$

- α: quality sensitivity; β: price sensitivity; φ: baseline demand
- Non-users resort to an outside option generating negative externalities (open defecation, unsafe water, polluting fuel use etc.)
- $s \geq 0$ denotes the net social gain per use and total social gain is $s\,D(e_1, e_2)$

Provider Effort and Costs

- For this part, we take a variant of the multitasking model
- Provider undertakes two efforts:
  - $e_1 \in [0,1]$: **fee enforcement** (monitoring payment)
  - $e_2 \in [0,1]$: **quality maintenance**
- Cost function:

$$c(e_1, e_2) = \frac{1}{2}e_1^2 + \frac{1}{2}e_2^2 - \eta\, e_1 e_2$$

- $\eta > 0$: complements — tasks share inputs or are performed simultaneously
  $\eta < 0$: substitutes — tasks compete for time and attention
  $\eta = 0$: independent
- The sign of η is not a modelling convenience — it is the key empirical and economic question: does maintaining quality make enforcement easier or harder?
- Convexity of cost function requires $|\eta| < 1$

First-Best: Social Planner

- Planner chooses e_1, e_2 directly; $\omega \equiv 1 - \tau \geq 0$ where τ is the welfare weight on user payments:

$$\max_{e_1, e_2} W = \left[s + \omega\, p e_1\right] D(e_1, e_2) - c(e_1, e_2)$$

First-order conditions:

$$(1 + \omega\beta p^2)\, e_1 - (\eta + \omega\alpha p)\, e_2 = \omega p\varphi - \tfrac{1}{2}\beta p s \tag{1a}$$

$$-(\eta + \omega\alpha p)\, e_1 + e_2 = \alpha s \tag{1b}$$

Three insights:

1. **Public-good motive:** quality condition — e_2 proportional to αs even without enforcement
2. **Two complementarity channels:** efforts interact via $(\eta + \omega\alpha p)$ — cost channel η and revenue channel $\omega\alpha p$
3. **Enforcement trade-off:** enforcement condition — revenue motive $\omega p\varphi$ vs. social exclusion cost $-\frac{1}{2}\beta p s$; when $\omega = 0$, free access is optimal

First-Best Insights

- In **high-resource** settings ($\tau = 1$, $\omega = 0$): optimal policy fully subsidizes access and sets quality proportional to s
- In **low-resource** settings ($\tau < 1$, $\omega > 0$): user fees are part of an optimal financing strategy
- **Complementarity between price and quality**: higher quality raises demand $\Rightarrow$ higher marginal return to charging a fee; higher fee $\Rightarrow$ higher marginal return to quality

Interpretation

User fees can be socially optimal under fiscal constraints even when the service is socially valuable. The revenue and cost channels jointly govern the relationship between enforcement and quality effort.

Second-Best: Agency Problem and Provider Payoff

- e_1, e_2 **not contractible**; provider does not internalise s; retains $\lambda > 0$ of fee revenue
  ▸ Details
- Policymaker offers bonus b on observable quality $q = e_2$; provider maximises $\pi = \lambda p e_1 D + b e_2 - c(e_1, e_2)$

First-order conditions:

$$(1 + \lambda \beta p^2)\, e_1 - (\eta + \lambda \alpha p)\, e_2 = \lambda p \varphi \quad (2a)$$

$$-(\eta + \lambda \alpha p)\, e_1 + e_2 = b \quad (2b)$$

Three distortions vs. first-best:

1. **Quality motive privatised:** $\alpha s \to b$ — provider ignores social externality
2. **Revenue motive privatised:** $\omega \to \lambda$ — fee revenue treated as private income
3. **Enforcement cost disappears:** $-\frac{1}{2}\beta p s$ vanishes — no private counterpart

Implication: even at $b = \alpha s$, $\lambda = \omega$, first-best fails — distortion 3 is uncorrectable by the bonus alone.

Quality Incentives: The Demand Decomposition

Raising the bonus b has two opposing effects on demand:

$$\frac{dD}{db} = \underbrace{\alpha \frac{de_2^*}{db}}_{\substack{\text{quality effect} \\ \text{(always positive)}}} - \underbrace{\tfrac{1}{2}\beta p \frac{de_1^*}{db}}_{\substack{\text{enforcement effect} \\ \text{(positive when } \eta>0\text{)}}}$$

Substituting closed-form solutions:

$$\frac{dD}{db} = \frac{\alpha(1+\beta\lambda p^2) - \frac{1}{2}\beta p(\alpha\lambda p + \eta)}{\Delta}, \qquad \Delta > 0$$

- When $\eta \leq 0$: enforcement effort does not rise with the bonus $\Rightarrow$ demand unambiguously increases
- When $\eta > 0$: bonus raises e_2, which via cost complementarity raises e_1, which raises $\tilde{p} = pe_1$ and suppresses demand — the **perverse enforcement effect**
- The net effect is **ambiguous** and depends on deep parameters

Proposition

Proposition

When efforts are complements ($\eta > 0$), the quality bonus ***reduces demand*** *if and only if:*

$$\frac{2}{\beta p} + \lambda p \; < \; \frac{\eta}{\alpha}$$

When this holds, the enforcement effect dominates and welfare falls despite improved quality.

What makes this interesting?

- A quality bonus with unchanged nominal fees can *reduce* usage and welfare — by strengthening the incentive to exclude
- Non-monotone in p: the risk is **greatest at intermediate fee levels**, not at high fees where exclusion seems most obvious
- Smaller λ does *not* protect: cost complementarity transmits the bonus into enforcement even when the revenue stake is low
- Mechanism is absent when $\eta \leq 0$ — a clean testable prediction

When Does the Mechanism Apply?

- The perverse effect requires cost complementarity ($\eta > 0$) and revenue complementarity condition — when is this plausible?

More likely:

- Same agent delivers quality *and* collects payment — tasks are physically integrated
- Community toilets, water stations, rural health posts: caretaker cleans, maintains, and monitors entry
- Electricity distribution: linemen maintain supply and enforce payment

Less likely:

- Tasks separated across roles: teachers do not collect fees; clinicians do not process billing
- $\eta \leq 0$ when quality effort competes directly with enforcement time

The model yields a testable prediction

The perverse demand effect should be concentrated in settings with task integration and fee dependence. This is precisely what the experimental design can test — and what we take to the data.

Other Contracting Solutions?

Why not make the provider the residual claimant?

- Would work without social returns, but here $s > 0$: provider would over-enforce and under-provide quality relative to social optimum
- Also: risk aversion and liquidity constraints limit lump-sum contracting in low-resource settings

Why not reward usage directly?

- Creates incentives to *waive* fee collection to attract users — undermining financial sustainability

What would get you the first-best?

- Set $p = 0$ and $b = \alpha s$
- Will not work in low-resource environments

Theory: Summary

	$\eta \leq 0$	$\eta > 0$ (condition holds)
Quality (e_2)	$\uparrow$	$\uparrow$
Enforcement (e_1)	unchanged / $\downarrow$	$\uparrow$
Demand (D)	$\uparrow$	$\downarrow$
Externalities	$\downarrow$	$\uparrow$
Welfare	$\uparrow$	ambiguous / $\downarrow$
Joint optimisation: shade b downward; lower p to reduce enforcement stakes		
Limited liability: amplifies distortion; feasible bonus set shrinks		

- A single, spare model — two efforts, one bonus, one fee — generates a rich set of predictions
- The empirical design is built directly around the model's key conditions and comparative statics

Outline

1 Theoretical framework

2 Field experiment

3 Data

4 Specification

5 Results

6 Conclusions

Field experiment

- Field experiment in the two largest cities of **Uttar Pradesh, India**
- **Pay-to-use community toilets (CTs)**
  - Public services offer essential access to hygiene and sanitation through communal facilities targeting specific group of residents.
  - Informal settlements (or *slums*) ⇒ overcrowding, limited space and inadequate housing constrain access to safely managed private toilets.
- **Outside option**: unimproved facilities or resorting to open defecation (OD) ⇒ significant negative externalities.
  - Infectious diseases and mortality [Geruso & Spears 2018, Coffey et al. 2018, Pickering et al. 2018].
  - Stunted human capital [Miguel and Kremer, 2004; Bleakley, 2007; Adukia, 2017; Augsburg and Rodriguez-Lesmes, 2018; Orgill-Meyer and Pattanayak, 2020; Spears, 2020].
  - Damaged environmental quality [Greenstone and Jack, 2015].

Field experiment

- Field experiment in the two largest cities of **Uttar Pradesh, India**
- **Pay-to-use community toilets (CTs)**
  - Public services offer essential access to hygiene and sanitation through communal facilities targeting specific group of residents.
  - Informal settlements (or *slums*) ⇒ overcrowding, limited space and inadequate housing constrain access to safely managed private toilets.
- **Outside option**: unimproved facilities or resorting to open defecation (OD) ⇒ significant negative externalities.
  - Infectious diseases and mortality [Geruso & Spears 2018, Coffey et al. 2018, Pickering et al. 2018].
  - Stunted human capital [Miguel and Kremer, 2004; Bleakley, 2007; Adukia, 2017; Augsburg and Rodriguez-Lesmes, 2018; Orgill-Meyer and Pattanayak, 2020; Spears, 2020].
  - Damaged environmental quality [Greenstone and Jack, 2015].

Field experiment

- Field experiment in the two largest cities of **Uttar Pradesh, India**
- **Pay-to-use community toilets (CTs)**
  - Public services offer essential access to hygiene and sanitation through communal facilities targeting specific group of residents.
  - Informal settlements (or *slums*) $\Rightarrow$ overcrowding, limited space and inadequate housing constrain access to safely managed private toilets.
- **Outside option**: unimproved facilities or resorting to open defecation (OD) $\Rightarrow$ significant negative externalities.
  - Infectious diseases and mortality [Geruso & Spears 2018, Coffey et al. 2018, Pickering et al. 2018].
  - Stunted human capital [Miguel and Kremer, 2004; Bleakley, 2007; Adukia, 2017; Augsburg and Rodriguez-Lesmes, 2018; Orgill-Meyer and Pattanayak, 2020; Spears, 2020].
  - Damaged environmental quality [Greenstone and Jack, 2015].

Data

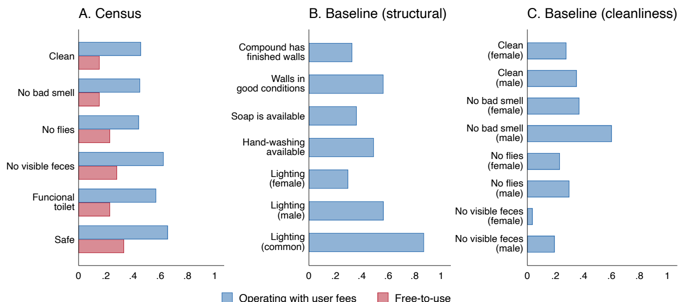
- In 2017 a census was conducted of all the facilities in the study area
- These data formed the basis for selecting the facilities that operate with user fees and are used primarily by residents, excluding those that are permanently closed or abandoned.
- More details to come

Field experiment

- Prevalent nationwide and operate with user fees (public or PPPs)
- Managed by a **caretaker** $\Rightarrow$ collect fees + maintenance
- Conditions of service delivery:
    - Poorly maintained and dirty.
    - Non-payment among users is common and WTP is low.

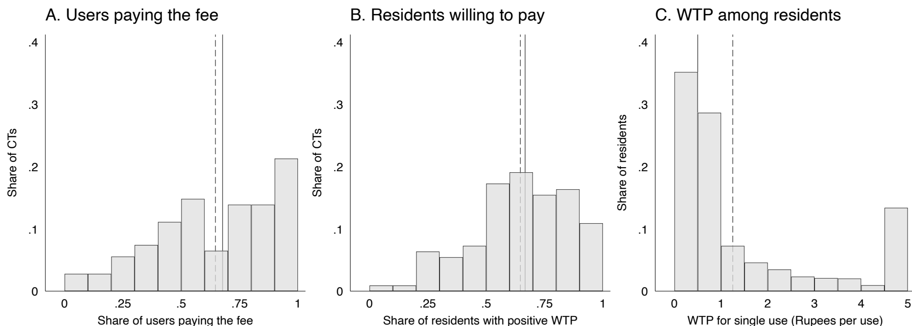
Field experiment

- Prevalent nationwide and operate with user fees (public or PPPs)
- Managed by a **caretaker** $\Rightarrow$ collect fees + maintenance
- Conditions of service delivery:
  - Poorly maintained and dirty.
  - Non-payment among users is common and WTP is low.



Field experiment

- Prevalent nationwide and operate with user fees (public or PPPs)
- Managed by a **caretaker** $\Rightarrow$ collect fees + maintenance
- Conditions of service delivery:
  - Poorly maintained and dirty.
  - Non-payment among users is common and WTP is low.



Applying the model to CTs in slums

Goal: reduce OD and encourage the use of CTs.

- CT use depends on:
  - β price sensitivity, which depends on fee-collection efforts (e_1)
  - α quality sensitivity, which depends on efforts to improve quality (e_2)
- High negative externalities associated with OD (low s)
- Instruments to improve use:
  1. Improving quality: either directly shifting an exogenous component of quality ($q = e_2 + a$) or indirectly incentivising the caretaker's effort to improve quality ($\uparrow e_2$ through b)
  2. Information campaign to increase awareness and that may boost demand: $\uparrow \varphi$

Applying the model to CTs in slums

- The **net welfare effect is potentially ambiguous** in these interventions given the core tension between exclusion due to increased monitoring and quality improvement
- In our context, the contract of caretakers does not include any performance-based financial rewards for quality improvement.
- However, to retain their job, they are aware that fee collection is essential:
  - As their salary is funded through fee revenues, caretakers need to monitor fee payments to avoid being relocated or fired.
  - The resources available to the caretakers for operating the facility are directly tied to fee revenues, which are used to fund cleaning agents, tools, and cleaners based on the amount collected.

Field experiment: experimental design

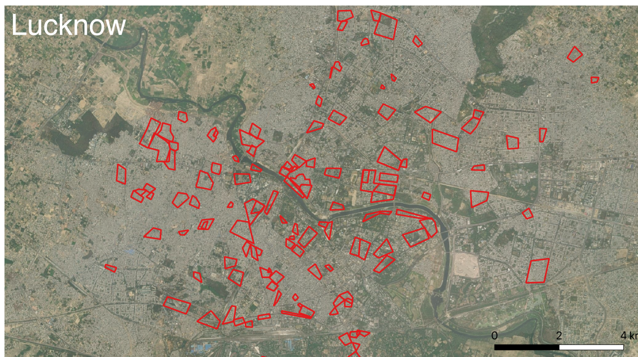
- Identified all CTs serving slums ⇒ census of slums and CTs
- 110 CTs randomly allocated to:
  1. **Control** group (40)
  2. **Maintenance** group (70) ⇒ boost quality in public service delivery

Field experiment: experimental design

- Identified all CTs serving slums ⇒ census of slums and CTs
- 110 CTs randomly allocated to:
  1. **Control** group (40)
  2. **Maintenance** group (70) ⇒ boost quality in public service delivery

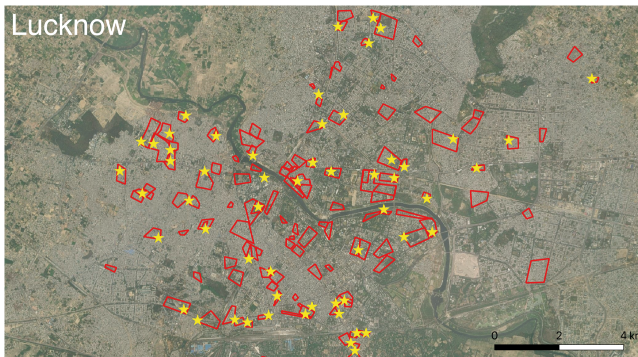
Field experiment: experimental design

- Identified all CTs serving slums ⇒ census of slums and CTs
- 110 CTs randomly allocated to:
  1. **Control** group (40)
  2. **Maintenance** group (70) ⇒ boost quality in public service delivery



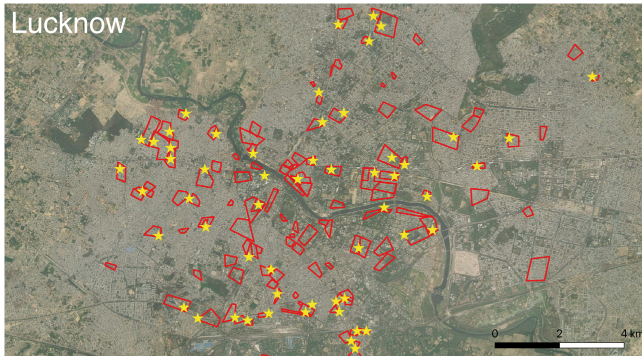
Field experiment: experimental design

- Identified all CTs serving slums ⇒ census of slums and CTs
- 110 CTs randomly allocated to:
    1. **Control** group (40)
    2. **Maintenance** group (70) ⇒ boost quality in public service delivery



Field experiment: experimental design

- Identified all CTs serving slums ⇒ census of slums and CTs
- 110 CTs randomly allocated to:
  1. **Control** group (40)
  2. **Maintenance** group (70) ⇒ boost quality in public service delivery

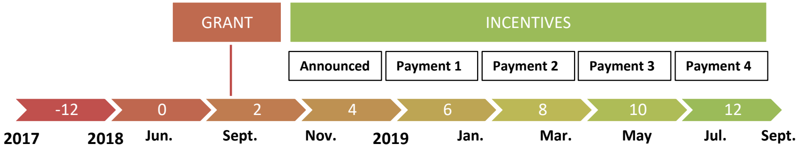


Maintenance intervention

Two components

1. **Grant** (*a*): one-off grant to rehabilitate the facility.
   - Caretaker(s) chose between different packages. ▸ examples
   - ≈ 90% of monthly O&M cost of adequate-quality CT.

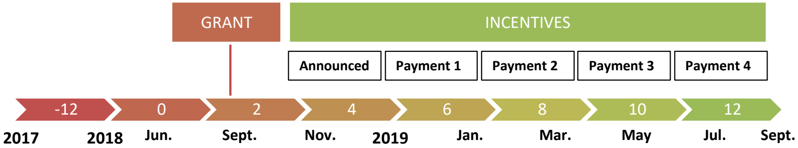
2. **Incentive** (*b*): financial rewards for routine maintenance
   - Paid to caretakers conditional on objective cleanliness
   - 40% of average monthly × 4 payments (≈ 13% annual salary)



Maintenance intervention

Two components

1. **Grant (*a*)**: one-off grant to rehabilitate the facility.
   - Caretaker(s) chose between different packages. ▸ examples
   - ≈ 90% of monthly O&M cost of adequate-quality CT.
2. **Incentive (*b*)**: financial rewards for routine maintenance
   - Paid to caretakers conditional on objective cleanliness
   - 40% of average monthly x 4 payments (≈ 13% annual salary)



Outline

1 Theoretical framework

2 Field experiment

3 Data

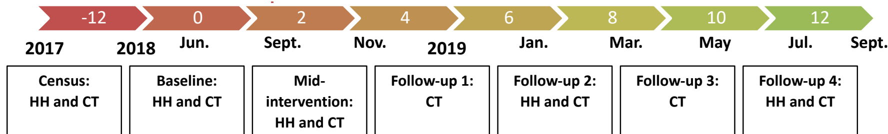
4 Specification

5 Results

6 Conclusions

Data I: public service delivery

1. **Surveys** (BL + 4 follow-ups): 110 CTs
2. **Objective measurements**:
   1. CT survey: administered to caretakers.
   2. Observations: number of users and payment, structural quality and cleaning status.
   3. Laboratory tests for bacteria presence.



Data II: residents

Creating a sampling frame around all pay-to-use CTs in Lucknow and Kanpur

- Slum borders in each CT catchment area + census of residents within
- *Residents*:
  - Using the CT or practising OD
  - Sample restricted by proximity to facility ▸ Distance

Data II: residents

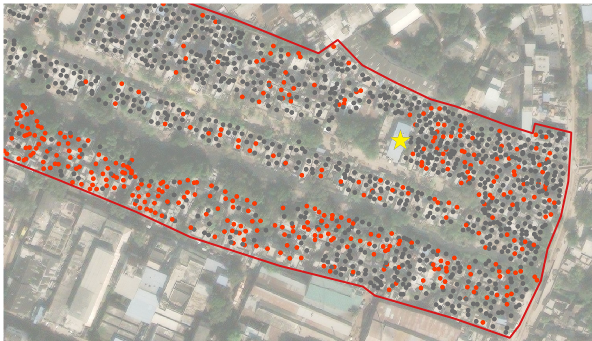
Creating a sampling frame around all pay-to-use CTs in Lucknow and Kanpur

- Slum borders in each CT catchment area + census of residents within
- *Residents*:
    - Using the CT or practising OD
    - Sample restricted by proximity to facility

Data II: residents

Creating a sampling frame around all pay-to-use CTs in Lucknow and Kanpur

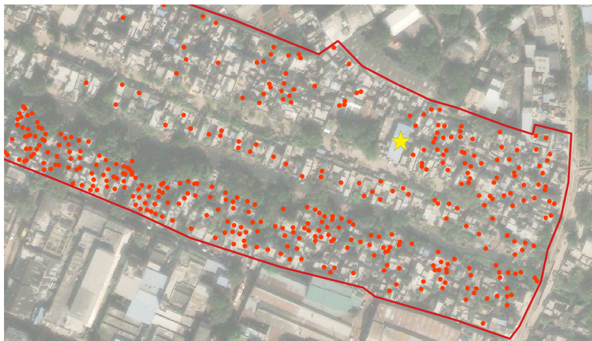
- Slum borders in each CT catchment area + census of residents within
- *Residents*:
  - Using the CT or practising OD
  - Sample restricted by proximity to facility ▸ Distance



Data II: residents

Creating a sampling frame around all pay-to-use CTs in Lucknow and Kanpur

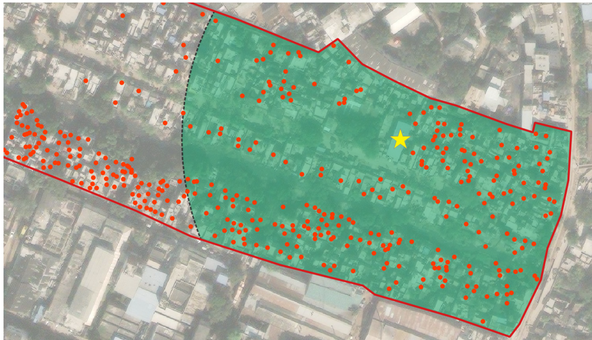
- Slum borders in each CT catchment area + census of residents within
- *Residents*:
  - Using the CT or practising OD
  - Sample restricted by proximity to facility Distance



Data II: residents

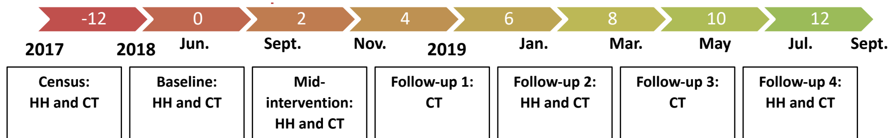
Creating a sampling frame around all pay-to-use CTs in Lucknow and Kanpur

- Slum borders in each CT catchment area + census of residents within
- *Residents*:
  - Using the CT or practising OD
  - Sample restricted by proximity to facility Distance



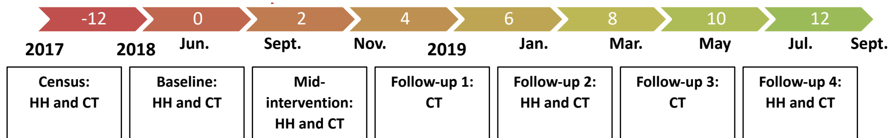
Data II: users and potential users

1. **Surveys** (BL + 3 follow-ups): 1500 residents
2. **Lab-in-the-field experiments**:
   - List randomization to measure outside option ⇒ OD is a sensitive behaviour



Data II: users and potential users

1. **Surveys** (BL + 3 follow-ups): 1500 residents
2. **Lab-in-the-field experiments**:
    - List randomization to measure outside option ⇒ OD is a sensitive behaviour



Outline

1. Theoretical framework
2. Field experiment
3. Data
4. Specification
5. Results
6. Conclusions

Specification

Maintenance (T_j) vs control comparison

$$Y_{ij,t} = \beta_0 + \beta_1 Tj + \beta_x X_{ij} + \delta_t + \epsilon_{ij,t} \tag{3}$$

- Separate estimates for **2 periods**: BL and FUs $\Rightarrow$ assumes β_1 is constant within these periods.
- Pool follow-up measurements to reduce noise [McKenzie, 2012].

Robustness

- Separate estimates for maintenance only vs maintenance plus sensitization
- Baseline balance in all observables CT/caretaker HH
- Attrition orthogonal to treatment allocation Attrition

Outline

1 Theoretical framework

2 Field experiment

3 Data

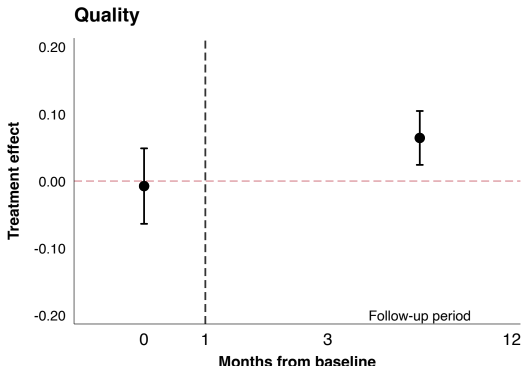
4 Specification

5 Results

6 Conclusions

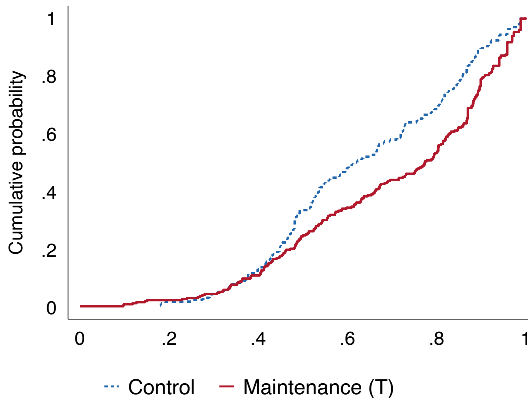
Quality of service delivery

- Index using **objective measurements of service delivery**:
  - Structural quality of the facility + cleanliness + presence of bacteria.
- Higher-quality provision ↑ 10.4% over control mean ▸ Factors ▸ Table
- Shift towards the top of the distribution



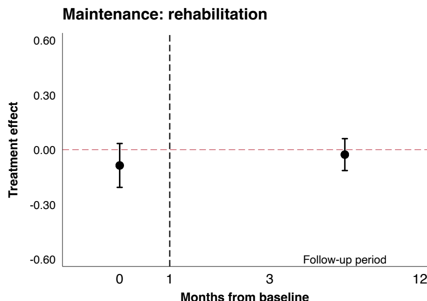
Quality of service delivery

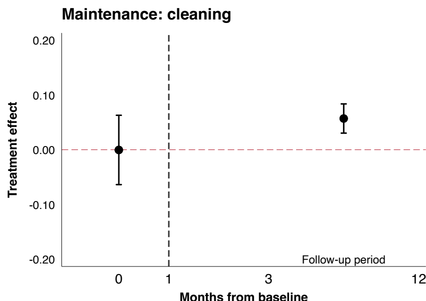
- Index using **objective measurements of service delivery**:
    - Structural quality of the facility + cleanliness + presence of bacteria.
- Higher-quality provision ↑ 10.4% over control mean Factors Table
- Shift towards the top of the distribution



Inputs to service delivery

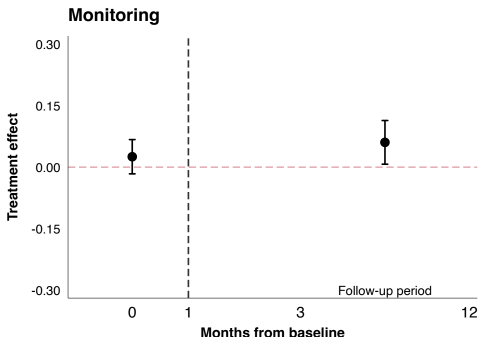
- **Maintenance** ↑
  - Rehabilitation is unaffected, but ↑ inputs used in cleaning and the correct implementation.
- Monitoring ↑
  - Caretakers respond to incentives by increasing share of time spent on monitoring (9.5%).





Inputs to service delivery

- **Maintenance** $\uparrow$
  - Rehabilitation is unaffected, but $\uparrow$ inputs used in cleaning and the correct implementation.
- **Monitoring** $\uparrow$
  - Caretakers respond to incentives by increasing share of time spent on monitoring (9.5%).



Users and payment

1. **Users**
   1. **Aggregate**: observed users during rush hour
   2. **Residents**: self-reported number of uses
2. **Payment**
   1. **Aggregate**: observed payment during rush hours.
   2. **Residents**: incentive-compatible WTP $\Rightarrow$ multiple price list
      - Random draw from 13 questions and payment based on corresponding choice.
      - Market price = Rs. 5 $\Rightarrow$ CT expenditure $\approx$ to 8% of HH income.

Option A	Option B
10 tickets	0 Rs
10 tickets	5 Rs (giving up 0.5 Rs/ticket)
10 tickets	10 Rs (giving up 1Rs/ticket)
10 tickets	15 Rs (giving up 1.5Rs/ticket)
10 tickets	20 Rs (giving up 2Rs/ticket)
10 tickets	25 Rs (giving up 2.5Rs/ticket)
.	.
.	.
10 tickets	60 Rs (giving up 6Rs/ticket)

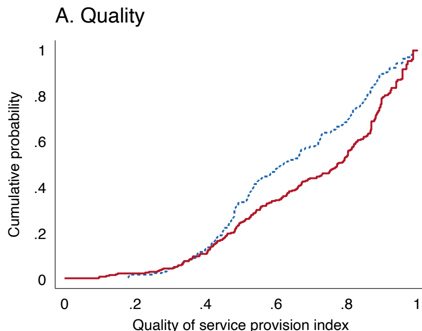
Use and payment

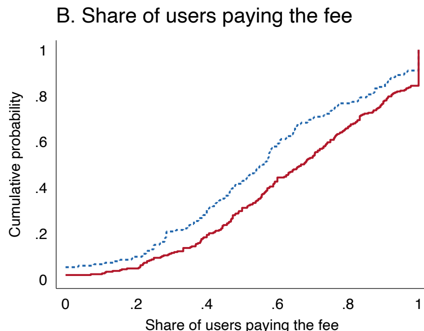
Dep. variable:	All users (residents and passersby)		Residents			
	Users during the peak hours	Share paying the fee	Users during the peak hours	Number of daily uses	Share paying the fee	WTP (amount)
	(1)	(2)	(3)	(4)	(5)	(6)
Maintenance (T)	−1.781	0.100	−2.258	-0.117	0.111	0.008
	(1.518)	(0.032)	(1.250)	(0.059)	(0.034)	(0.080)
	[0.24, 0.23]	[0.00, 0.01]	[0.07, 0.10]	[0.05, 0.08]	[0.00, 0.09]	[0.92, 0.90]
Mean (control group)	33.903	0.556	27.519	1.198	0.489	1.205
Observations	434	434	434	3303	434	6001
Catchment areas	110	110	110	109	110	109
Observation rounds	4	4	4	2	4	2
Level of analysis	CT	CT	CT	Household	CT	Respondent
Measurement	Observed	Observed	Observed	Self-reported	Observed	Incentivized

Note. 'All users' refers to anyone using the service, while 'residents' focuses only on people who live in the slum and do not have access to private sanitation (see Section 4.2). Estimates based on CT-level OLS regressions using equation (9). Standard errors clustered by catchment area are reported in parentheses. The p-values are presented in brackets, the first from individual testing, the second adjusting for testing that each treatment is jointly different from zero for all outcomes presented in the table (see Section 5 for details). Dependent variables by column: (1) total number of users observed during the peak hours; (2) observed share of users who pay the entry fee during the peak hours; (3) total number of users living in the slum observed during the peak hours; (4) number of times the respondent used the CT for defecation in the day previous to the interview; (5) observed share of users living in the slum who pay the entry fee during the peak hours; (6) incentivized willingness to pay for a single CT use (in rupees), elicited for a bundle of 10 tickets and divided by 10 to get a single-use WTP. All specifications include indicator variables for data collection rounds and randomization strata. Additional details about the variables are presented in Online Appendix B.

Quality and payment

- Results consistent with price-elasticity effect > quality effect
- Increase in quality and payment comes at the cost of user selection
- No effect on revenues revenues





Control   Maintenance (T)

Outside option and externalities

1. **Practiced OD:** list randomization on behavior in the previous day
   - Randomly allocated to short or long list of statements.
   - Difference in average number of items between B and A gives prevalence.

Short (A)	Long (B)
- I cooked yesterday	- I cooked yesterday
- I bought milk yesterday	- I bought milk yesterday
- I watched TV yesterday	- I watched TV yesterday
	- I defecated in the open yesterday

Outside option and externalities

Dep. variable:	Outside option	Health consequences			
	Practiced OD	Morbidity	Health expenditure	Preventive healthcare	Curative healthcare
	(1)	(2)	(3)	(4)	(5)
Maintenance (T)	0.228	0.035	0.040	−0.003	0.047
	(0.071)	(0.023)	(0.070)	(0.003)	(0.021)
	[0.00, 0.01]	[0.14, 0.35]	[0.57, 0.48]	[0.31, 0.48]	[0.03, 0.12]
Mean (control group)	0.210	0.451	6.937	0.992	0.636
Observations	817	3323	3306	3322	3298
Catchment areas	107	109	109	109	109
Observation rounds	1	2	2	2	2
Level of analysis	Respondent	Household	Household	Household	Household
Measurement	List randomization	Self-reported	Self-reported	Self-reported	Self-reported

Note. Estimates refer to residents, that is, people living in the slum and not having access to private sanitation (see Section 4.2). Estimates based on household-level OLS regressions using equation (9). Standard errors clustered by catchment area are reported in parentheses. The p-values are presented in brackets, the first from individual testing, the second adjusting for testing that each treatment is jointly different from zero for all outcomes presented in the table (see Section 5 for details). Dependent variables by column: (1) share of study participants who practiced OD the day before the interview, obtained using the list randomization technique applied to the most senior male and female household member in follow-up 4; (2) an indicator variable equal to 1 if any household member had a fever, diarrhea, or a cough during the two weeks previous to the interview, and 0 otherwise; (3) total health expenditure during the month previous to the interview, reported in logarithms; (4) an indicator variable equal to 1 if the respondent spent on preventive healthcare, and 0 otherwise; and, (5) indicator variable equal to 1 if the respondent spent on curative healthcare, and 0 otherwise. Column 1 includes only 107 catchment areas because, due to the randomization of lists to respondents, a number of areas do not have respondents with the list of items including OD. Columns 2–5 include only 109 catchment areas in the sample because the dependent variables were measured only in rounds 3 and 5, after one slum was displaced. All specifications include indicator variables for data collection rounds and randomization strata. Additional details about the variables are presented in Online Appendix B.

Alternative mechanisms

We address and rule out several alternative explanations linking $\downarrow$ CT use and $\uparrow$ OD.

1. User exclusion is not driven by overcrowding caused by a surge in demand, as we don't find $\uparrow$ traffic, not even after the grant scheme ▸ CT use
2. No evidence that the users adjusted the timing of their visits to the CT ▸ Timing
3. Not driven by caretaker absenteeism, changes in opening hours or closure of facilities, as suggested by our unannounced visits ▸ Attrition

Outline

1 Theoretical framework

2 Field experiment

3 Data

4 Specification

5 Results

6 Conclusions

Conclusions

- **Theoretical framework** highlights how, for the service provider, fee payment monitoring and maintenace efforts are complementary
- **Field experiment** shows mechanisms behind this effect:
  - Exogenous boost to maintenance increases **quality**
  - $\uparrow$ both efforts: **maintenance** and **monitoring** of fee-payment
  - Increase in **payment** is accompanied by **user exclusion**
  - Reliance on **outside option** increases, which entails negative health **externalities**.
- Boosting quality of public services can exclude users

Thank you!

Policy discussion

- When the policymaker can contract the effort of the service provider, the FCOs show that in the absence of net social gains from the service ($s = 0$), the optimal fee collection effort is strictly positive ($e_1 > 0$).
- Large net social gains cause this effort to be equal to zero ($e_1 = 0$), and so access to the service should be free whenever the net social gain is sufficiently high.
- In the presence of large social gains from the service, the use of a bonus to improve the quality of the service would improve social welfare only in the free-to-use scenario; otherwise, the welfare effects would be ambiguous.
- Funding free-to-use, and securing quality, requires fully subsidizing operations and maintenance through alternative means, such as taxation –a dilemma in LMICs due to challenges in expanding tax bases, particularly for property taxes [Weigel, 2020; Balan et al., 2022; Dzansi et al., 2022].

Why Does the Provider Care About Fee Revenue?

Common arrangements in fee-funded public services:

- **Salary funded from fees:** provider must collect to get paid
- **Inputs funded from fees:** cleaning agents, tools, operating costs
- **Profit/revenue sharing:** contracted providers retain a share as compensation
- **Job security:** poor collection makes the service financially unviable

In our setting

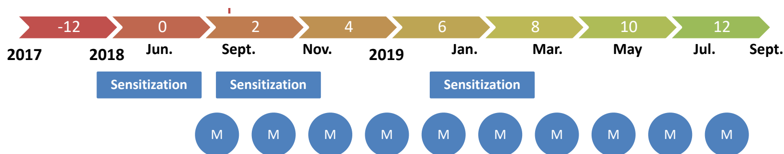
Caretaker salaries (INR 5,000/month) and all operating inputs are funded from CT fee revenues. Poor collection risks dismissal. $\lambda > 0$ is institutional reality, not an assumption.

Back

Sensitization campaign among potential users

Raise awareness about the **returns of a well-maintained facility**

1. Door-to-door campaign
2. Distribution of leaflets
3. Posters placed in CTs
4. Monthly reminder voice messages (M)



Sensitization campaign

- **Effective at reaching individuals**, but **no effect on behavior**

Table D9: Exposure to the interventions, by component

	Maintenance		**Sensitization campaign**		
	Transfer to the ...		Recall of WASH campaign		Awareness
	CT	Caretaker	Interactive activities	Posters at CT	
	(1)	(2)	(3)	(4)	(5)
Panel A					
Maintenance (T)	4.739	0.761	0.053	0.090	0.031
	(0.060)	(0.034)	(0.020)	(0.028)	(0.018)
	[0.00]	[0.00]	[0.01]	[0.00]	[0.10]
Panel B					
Maintenance only (T1)	4.645	0.746	0.023	0.019	0.008
	(0.081)	(0.045)	(0.025)	(0.031)	(0.022)
	[0.00]	[0.00]	[0.35]	[0.54]	[0.71]
Maintenance + sensitization (T2)	4.839	0.776	0.083	0.160	0.053
	(0.074)	(0.047)	(0.021)	(0.029)	(0.020)
	[0.00]	[0.00]	[0.00]	[0.00]	[0.01]
T1 = T2 (p-value)	0.063	0.636	0.009	0.000	0.042
Mean (control group)	0.315	0.063	0.646	0.327	0.660
Std. dev. (control group)	0.358	0.025	0.478	0.469	0.474
Observations	560	560	4844	3323	4793
Catchment areas	110	110	110	109	110
Observation rounds	5	5	3	2	3

Sensitization campaign

- **Effective at reaching individuals**, but **no effect on behavior**

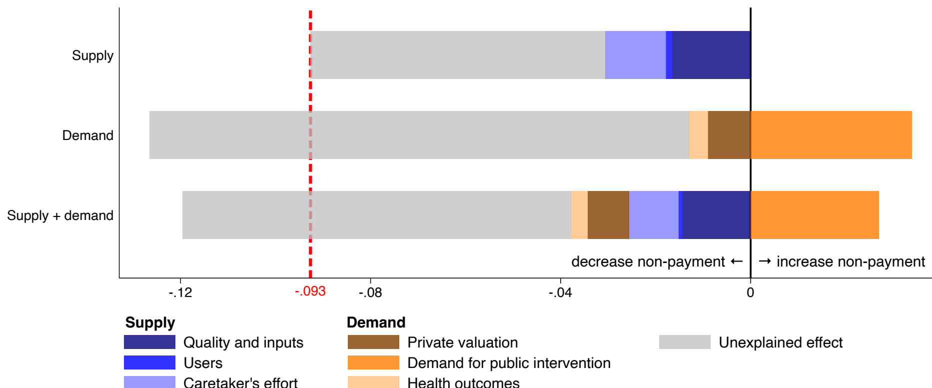
Table 4: The effect of sensitization

	Maintenance only			Maintenance + sensitization			$T1 = T2$
	β	se	p-value	β	se	p-value	p-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Quality	0.07	0.03	0.01	0.05	0.03	0.09	0.58
Maintenance: cleaning	0.06	0.02	0.00	0.06	0.02	0.00	0.85
Maintenance: rehabilitation	-0.04	0.06	0.47	-0.01	0.06	0.85	0.60
Monitoring	0.05	0.04	0.22	0.07	0.04	0.04	0.35
Share of users paying	0.08	0.05	0.09	0.11	0.05	0.03	0.54
Share of residents with positive WTP	0.01	0.03	0.74	-0.03	0.02	0.28	0.22
WTP among residents	0.09	0.11	0.41	-0.07	0.10	0.49	0.16
Users	-2.61	1.85	0.16	-1.25	1.81	0.49	0.42
Number of uses among residents:							
Regular users	-0.06	0.05	0.28	-0.16	0.06	0.01	0.13
Other	-0.23	0.11	0.04	-0.16	0.11	0.16	0.58
Morbidity	0.03	0.03	0.36	0.03	0.03	0.34	1.00
Health expenditure:							
Curative (extensive)	0.04	0.03	0.17	0.06	0.03	0.04	0.52
Curative (intensive)	31.25	227.29	0.89	-98.73	226.54	0.66	0.58
Preventive (extensive)	-0.00	0.00	0.41	-0.00	0.00	0.60	0.73
Preventive (intensive)	20.09	64.90	0.76	-10.44	63.43	0.87	0.61
Practiced OD	0.19	0.10	0.05	0.16	0.09	0.08	0.71

Note. In columns (1)–(6), estimates are based on CT-, respondent- or household-level OLS regressions using equation (6) for each outcome. $p-$values are presented in columns (3) and (6), the first from individual testing, the second adjusting for jointly testing that each treatment is different from zero for all outcomes presented in the table. Column (7) presents a test based on equality of coefficients of the effects of T1 and T2. Standard errors are clustered by catchment area for CT-level outcomes and by catchment-area–round for respondent- and household-level outcomes. The dependent variables are indicated in the rows and are defined in Appendix A. All specifications include indicator variables for data collection rounds, and strata indicators for the city and the provider of the CT.

Full payment

- Cost of improved services are 1.3–2.8x current cost ⇒ fully covered by eradicating non-payment at the market fee.
- Mediation analysis on the effect of the interventions on non-payment.
- While supply-side mediators mainly ↓ non-payment, demand-side factors ↑.

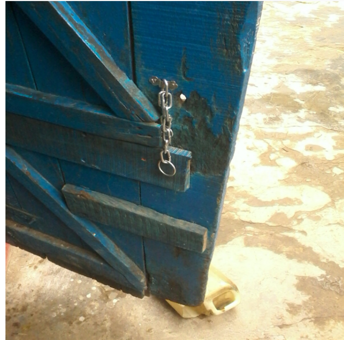


Intervention - CT

- **One-off CT grant scheme**
- Example of deep cleaning:

Intervention - CT ▸ Back

- **One-off CT grant scheme**
- Example of repair:



Intervention - sensitization

Back

- **Door-to-door information campaign**

Intervention - sensitization

Back

- Distribution of leaflets

सामुदायिक शौचालयों का उपयोग करने के लाभ

1. सामुदायिक शौचालय सीवरेज या सेप्टिक टैंक से जुड़ा होता है, जिससे खुले में शौच से होने वाले संक्रमण नहीं होते है.
2. उचित शुल्क, दैनिक साफ़ सफाई, रखरखाव, मरम्मत, बिजली और संचालन आदी सेवाएं प्रदान करता है.
3. महिलाओं के लिए सुरक्षित, दरवाजे और ताले के साथ रोशनी प्रदान करते हैं.
4. अतिरिक्त सेवाएं: हाथ धोने की सुविधा, साबुन, स्नानघर, कपड़े धोना और नियमित पानी की आपूर्ति.
5. सामुदायिक शौचालय आप, आपके परिवार और आपके समुदाय के लिए बेहतर स्वास्थ्य, सुरक्षा और गरिमा का प्रतीक हैं.

Institute for Fiscal Studies

सामुदायिक शौचालय बेहतर बनाने के लिए आप और आपका परिवार क्या कर सकता हैं?

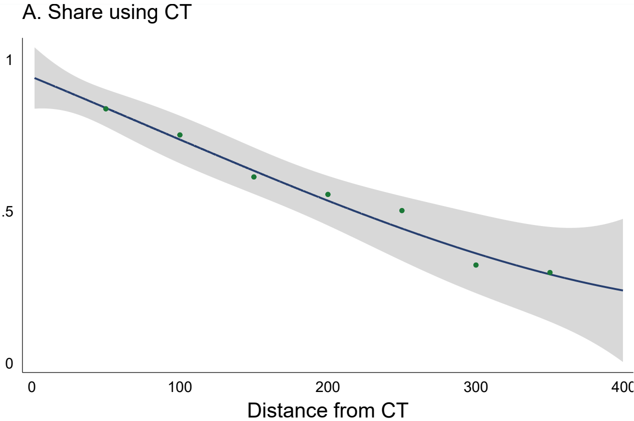
1. नियमित रूप से भुगतान
2. सामुदायिक शौचालयों के कूड़ेदानों में ही कचरा फेके।
3. उपयोग के बाद फ्लश करने के लिए पानी का उपयोग करें.
4. शौचालय परिसर को दूषित न करें.
5. शौचालयों का उपयोग करते समय वृद्ध लोगों और बच्चों की सहायता करें।
6. अँधेरे में महिलाओं और लड़कियों को अकेले न जाने दें.
7. शौचालय उपयोग करने के बाद स्वच्छ रखें जैसा की आप अपने उपयोग के लीये चाहते है.

Institute for Fiscal Studies

- Posters placed in CTs

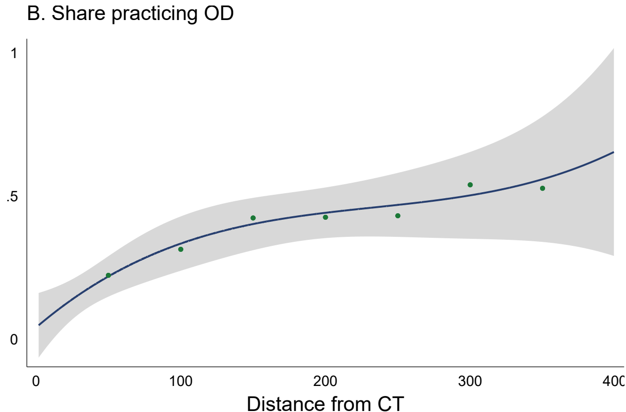
Distance and use

Back



Distance and use

Back



Baseline balance: CT characteristics ▸ Back

Table D1: CT characteristics at baseline, by treatment group

	Descriptive statistics		Differences from control group, by treatment group			
	All	Control	Any treatment	Improvement	Improvement + sensitization	P-value joint test (4)-(5)
	(1)	(2)	(3)	(4)	(5)	(6)
Year of construction	1997.11	1995.26	2.87	2.91	2.83	0.32
	[8.81]	[9.29]	(1.87)	(2.22)	(2.17)	
Distance to closest CT	0.54	0.58	-0.06	-0.04	-0.07	0.77
	[0.45]	[0.67]	(0.09)	(0.11)	(0.10)	
Surrounding: Market	0.33	0.36	-0.04	0.00	-0.08	0.69
	[0.47]	[0.49]	(0.10)	(0.11)	(0.11)	
Surrounding: Road	0.83	0.87	-0.06	-0.02	-0.09	0.54
	[0.37]	[0.34]	(0.08)	(0.09)	(0.09)	
Surrounding: Government office	0.25	0.21	0.07	0.10	0.04	0.64
	[0.44]	[0.41]	(0.09)	(0.10)	(0.10)	
Single caretaker	0.80	0.82	-0.04	0.03	-0.10	0.39
	[0.40]	[0.39]	(0.08)	(0.10)	(0.09)	
% Woman caretaker	0.18	0.22	-0.06	-0.05	-0.07	0.70
	[0.37]	[0.39]	(0.07)	(0.09)	(0.08)	
Caretaker is cleaner	0.27	0.28	-0.02	-0.04	-0.00	0.92
	[0.45]	[0.46]	(0.09)	(0.11)	(0.10)	
Caretaker from community	0.44	0.49	-0.07	-0.12	-0.01	0.54
	[0.50]	[0.51]	(0.10)	(0.12)	(0.12)	
Months caretaker in CT	125.28	129.91	-6.94	-0.61	-12.52	0.86
	[103.45]	[109.34]	(22.51)	(26.53)	(25.71)	
% Time collecting fees	0.35	0.36	-0.01	-0.02	-0.00	0.74
	[0.11]	[0.11]	(0.02)	(0.03)	(0.03)	
% Time cleaning	0.20	0.21	-0.01	-0.01	-0.01	0.64
	[0.06]	[0.06]	(0.01)	(0.01)	(0.01)	
Clean frequently	0.86	0.87	-0.02	-0.02	-0.01	0.96
	[0.35]	[0.34]	(0.07)	(0.08)	(0.08)	

Note. Columns (1)–(2) report sample mean and standard deviation in brackets for the whole sample and control group, respectively. Column (3) reports the difference with the control group with all treatment groups pooled together using an OLS regression of the correspondent outcome on the treatment indicator. Columns (4)-(5) report the difference with the control group for each treatment group. Standard errors clustered at slum level are reported in parentheses. Column (6) present a joint test of significance of the coefficients for each treatment dummy.

Baseline balance: HH characteristics

Back

Table D2: Household characteristics at baseline, by treatment group

	Descriptive statistics		**Differences from control group, by treatment group**			
	All	Control	Any treatment	Improvement	Improvement + sensitization	P-value joint test (4)-(5)
	(1)	(2)	(3)	(4)	(5)	(6)
Head, age	45.43	46.02	-0.93	-0.96	-0.90	0.55
	[12.82]	[13.42]	(0.84)	(1.00)	(0.94)	
Head, male	0.75	0.73	0.03	0.05	0.01	0.29
	[0.43]	[0.44]	(0.02)	(0.03)	(0.03)	
Head, educ < primary	0.54	0.56	-0.03**	-0.09**	0.03	0.03
	[0.50]	[0.50]	(0.04)	(0.05)	(0.04)	
Head, married	0.77	0.76	0.01	0.01	0.01	0.88
	[0.42]	[0.43]	(0.03)	(0.03)	(0.03)	
Number of children below 6 years old	0.47	0.50	-0.05	-0.04	-0.06	0.69
	[0.77]	[0.82]	(0.06)	(0.07)	(0.07)	
Number of adult members	4.47	4.44	0.05	0.04	0.06	0.87
	[1.83]	[1.92]	(0.11)	(0.14)	(0.12)	
Muslim	0.17	0.12	0.08*	0.11*	0.06	0.13
	[0.37]	[0.32]	(0.04)	(0.06)	(0.05)	
General caste	0.07	0.05	0.03	0.03	0.02	0.29
	[0.26]	[0.23]	(0.02)	(0.02)	(0.02)	
Asset index	0.53	0.53	0.00	0.01	-0.00	0.79
	[0.15]	[0.16]	(0.02)	(0.02)	(0.02)	
Piped water	0.71	0.70	0.01	-0.01	0.04	0.72
	[0.45]	[0.46]	(0.06)	(0.07)	(0.07)	
Private toilet	0.08	0.07	0.01	0.01	0.02	0.71
	[0.27]	[0.26]	(0.02)	(0.02)	(0.02)	
CT expense	180.53	173.42	11.20	-2.50	24.23	0.65
	[244.52]	[221.41]	(22.92)	(22.57)	(31.01)	

Note. Columns (1)–(2) report sample mean and standard deviation in brackets for the whole sample and control group, respectively. Column (3) reports the difference with the control group with all treatment groups pooled together using an OLS regression of the correspondent outcome on the treatment indicator. Columns (4)-(5) report the difference with the control group for each treatment group. Standard errors clustered at slum level are reported in parentheses. Column (6) present a joint test of significance of the coefficients for each treatment dummy.

Random attrition and replacements

Back to attrition

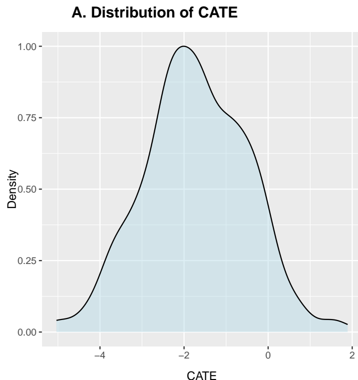
	Follow-up interviews per baseline household	**Interviewed at baseline and not re-interviewed in ...**				**Replacements**
		Any follow-up	Follow-up 1	Follow-up 3	Follow-up 5	Household is replacement
	(1)	(2)	(3)	(4)	(5)	(6)
Maintenance (T1)	0.029	0.004	0.013	-0.026	-0.016	0.008
	(0.072)	(0.011)	(0.022)	(0.037)	(0.035)	(0.015)
	[0.69]	[0.73]	[0.57]	[0.48]	[0.65]	[0.60]
Maintenance + sensitization (T2)	0.013	0.008	0.003	-0.014	-0.002	-0.000
	(0.078)	(0.014)	(0.021)	(0.041)	(0.034)	(0.014)
	[0.87]	[0.54]	[0.87]	[0.73]	[0.96]	[0.99]
T1 = T2 (p-value)	0.807	0.754	0.678	0.706	0.656	0.594
Attrition rate	2.575	0.025	0.090	0.194	0.142	0.161
Observations	1575	1575	1575	1575	1575	6711

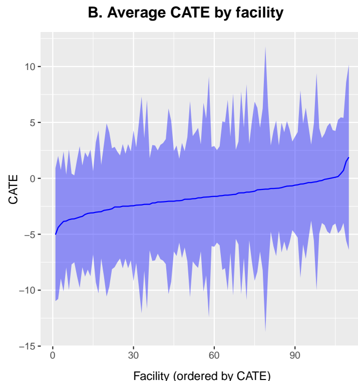
Note. Figure B2 provides the timing of each follow-up survey. Dependent variables by column: (1) indicator variable equal to 1 if the household was interviewed at baseline and was not re-interviewed in any of the follow-ups, and zero otherwise; (2) indicator variable equal to 1 if the household was interviewed at baseline and was not re-interviewed in two out of three follow-ups, and 0 otherwise; (3)–(5) indicator variable equal to 1 if the household was interviewed at baseline and was not re-interviewed at follow-up 1 or follow-up 2 or follow-up 3, and 0 otherwise; (6) indicator variable equal to 1 if the household is part of the replacement sample (it was interviewed in any of the follow-ups, but it was not interviewed at baseline), and 0 otherwise. In columns (1)–(5), the sample is restricted to baseline observations, while in column (6) the sample is restricted to follow-up observations. All specifications include strata indicators for city and the provider of the CT. Standard errors clustered by catchment area are presented in parenthesis in columns (1)–(5). Standard errors clustered by catchment area and follow-up round are presented in parenthesis in column (6).

Back to alternative mechanisms

Heterogeneity of effect on non-payment

Conditional ATE of the maintenance treatment on non-payment computed using the causal forest procedure of Basu et al. (2018) and Athey and Wager (2019)



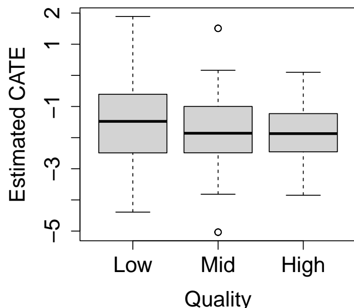


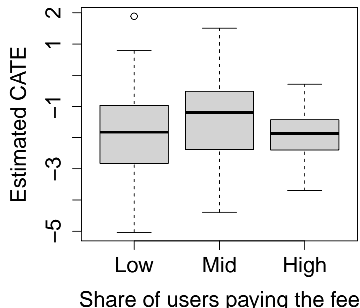
▸ Back

Heterogeneity of effect on non-payment

Conditional ATE of the maintenance treatment on non-payment computed using the causal forest procedure of Basu et al. (2018) and Athey and Wager (2019)

C. CATE by main baseline characteristics

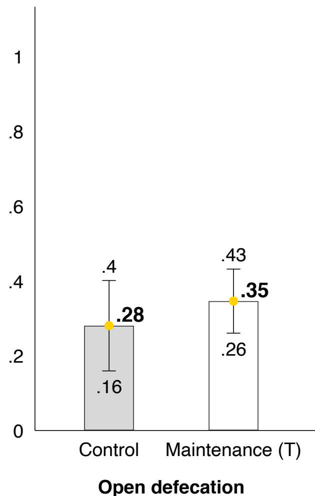


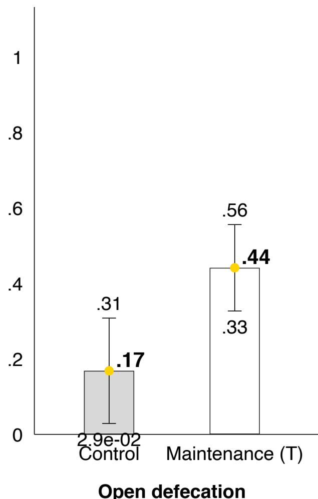


Back

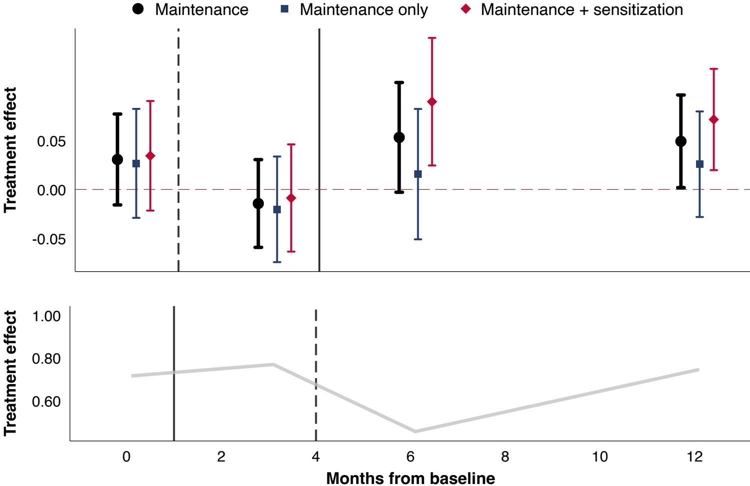
OD prevalence, by gender

Back





Increased awareness of externalities from OD



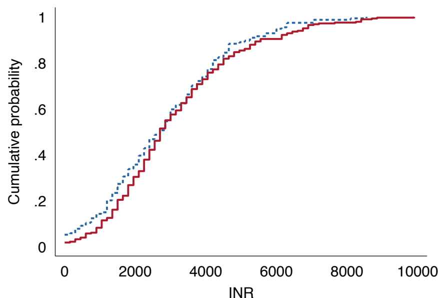
Sensitization campaign

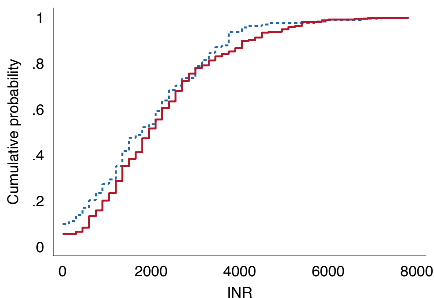
- The campaign was **effective at reaching individuals**

	Recall of WASH campaign		Voice messages
	Interactive activities	Posters at CT	Exposure
	(1)	(2)	(3)
Maintenance only (T1)	0.023	0.017	-0.038
	(0.024)	(0.030)	(0.047)
	[0.33]	[0.58]	[0.42]
Maintenance + sensitization (T2)	0.083	0.158	0.827
	(0.023)	(0.029)	(0.086)
	[0.00]	[0.00]	[0.00]
T1 = T2 (p-value)	0.014	0.000	0.000
Mean (control group)	0.645	0.327	0.188
Std. dev. (control group)	0.479	0.469	0.347
Observations	4793	3301	4793
Catchment areas	328	218	328
Observation rounds	3	2	3

Revenues

Back

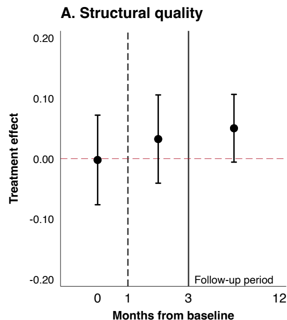


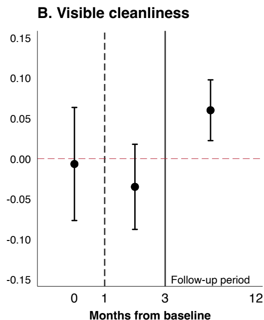


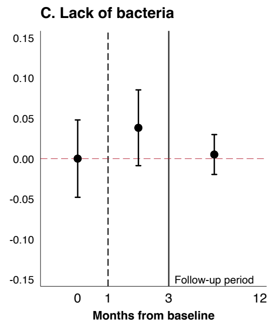
Control Maintenance (T)

Quality of service delivery: factors

Back







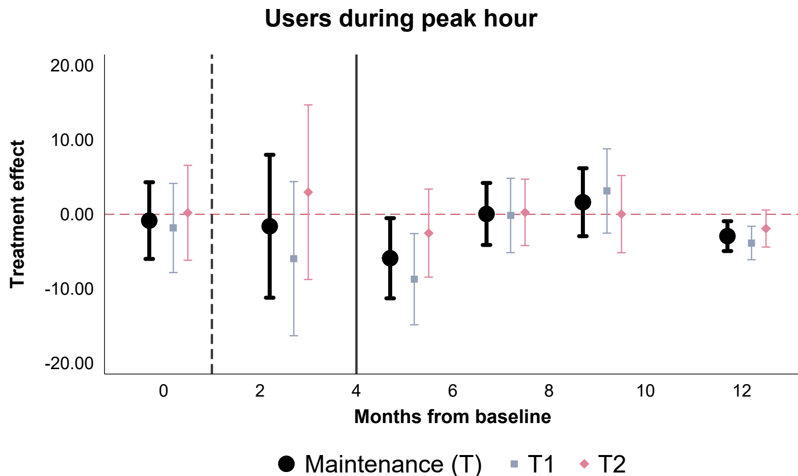
Service delivery

▸ Back

Dep. variable:	Quality	Efforts in service delivery		
		Maintenance		Monitoring
		Cleaning	Renovation	
	(1)	(2)	(3)	(4)
Maintenance (T)	0.066	0.061	-0.012	0.067
	(0.022)	(0.014)	(0.048)	(0.028)
	[0.00, 0.01]	[0.00, 0.00]	[0.80, 0.80]	[0.02, 0.05]
Mean (control group)	0.636	0.513	0.625	0.707
Observations	434	434	434	434
Catchment areas	110	110	110	110
Observation rounds	4	4	4	4
Level of analysis	CT	CT	CT	CT
Measurement	Observed	Self-reported	Self-reported	Self-reported

Note. Estimates based on CT-level OLS regressions using equation (9). Standard errors clustered by catchment area are reported in parentheses. The p-values are presented in brackets, the first from individual testing, the second adjusting for testing that each treatment is jointly different from zero for all outcomes presented in the table (see Section 5 for details). Dependent variables by column: (1) quality, an index computed by aggregating indicator variables about the structural quality of the facility, its cleanliness, and the lack of bacteria, and re-scaled to be between 0 (lowest in-sample quality) and 1 (highest in-sample quality); (2) cleaning, an index including the number of tools, materials, and cleaners used during the last cleaning of the facility and the caretaker's knowledge about this process, normalized to be between 0 and 1 (see Table D10 for individual components); (3) renovation, an indicator variable equal to 1 if the CT received repairs and/or deep cleaning of the infrastructure in the month previous to the visit, and 0 otherwise; (4) monitoring, the share of worked hours allocated by the caretaker to collect fees and supervise cleaners, rather than conducting activities away from the entrance or off-site. All specifications include indicator variables for data collection rounds and randomization strata. Additional details about the variables are presented in Online Appendix B.

Timing of effects on users during peak hours



Back

Services revenues, alternative timing use, and payment enforcement

Back

Dep. variable:	Monthly revenues during rush hour		Alternative timing	Caretaker ever refused entry		Refused entry for not paying	
	Total	From residents	Users during afternoon	All CTs	Low payment	All CTs	Low payment
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Maintenance (T)	335.722	268.163	-0.383	0.019	0.069	0.009	0.051
	(195.127)	(143.555)	(0.470)	(0.020)	(0.028)	(0.020)	(0.022)
	[0.09]	[0.06]	[0.42]	[0.34]	[0.02]	[0.64]	[0.03]
Mean (control group)	2840.260	1954.870	12.955	0.076	0.044	0.074	0.041
Observations	434	434	434	1641	812	1641	812
Catchment areas	110	110	110	109	53	109	53
Observation rounds	4	4	4	1	1	1	1

Note. Columns (1)–(3) show estimates based on CT-level OLS regressions using equation (9), while columns (4)–(7) show estimates based on household-level OLS regressions using equation (9). Standard errors clustered by catchment area are reported in parentheses and p-values in brackets. Dependent variables are defined in Appendix B. *Low payment* restricts the sample to CTs that at baseline presented a share of users paying the fee below the sample median. All specifications include indicator variables for data collection rounds and randomization strata. Details about the variables are presented in Appendix B.

Consequences for public health

- $\uparrow$ 7% in positive curative expenditures over control.