SUPPLEMENTARY MATERIAL – NOT FOR ONLINE PUBLICATION

Public Service Delivery and Free Riding: Experimental Evidence from India

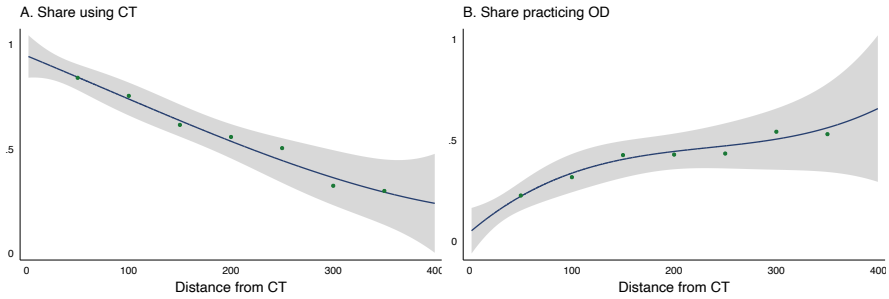
Alex Armand, Britta Augsburg, Antonella Bancalari

This supplementary material provides detailed information about the context in which the study is implemented (S.1), a detailed description of interventions (S.2), their cost (S.3), sampling, data collection and measurement (S.4).

S.1 An analysis of CTs in the study area

The location of a CT in the slum is a strong predictor of sanitation-related behavior among eligible households. Using self-reported data from the census of slum residents (see Supplementary Material S.4.1), we study how distance from a facility affects the use of the service versus free riding among study households. Figure S.1.1 presents cubic fits for the relationships between the distance from the facility and self-reported use of the service (panel A), or OD (panel B). The use of a CT reduces rapidly with distance from the facility: at 200 meters, only 50 percent of eligible households use the facility, and more than 40 percent practice OD. At 400 meters from a CT, very few households use the service, while over 50 percent of respondents practice OD.

Figure S.1.1: Sanitation behavior in the slum, by distance from a facility

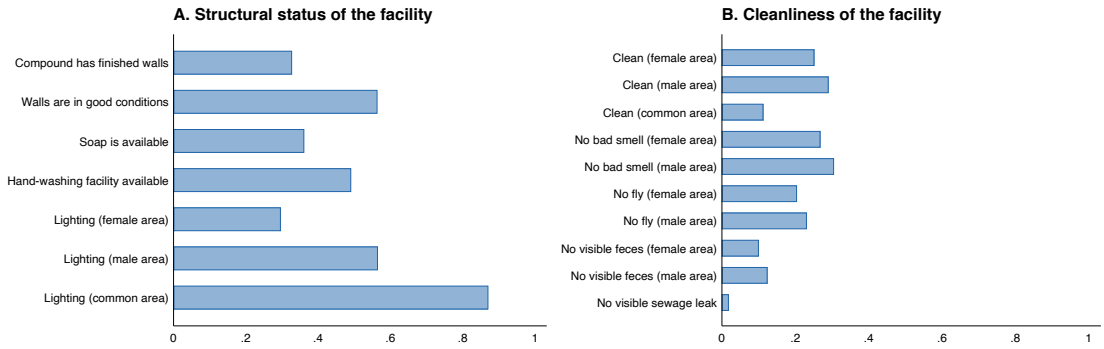


Note. Data source is the slum resident census (see Supplementary Material S.4). The figures present cubic fits of the share using CT (panel A) and of the share practicing OD (panel B) on distance from the closest CT. Dots show the average of the variable indicated on the vertical axis for equally-spaced intervals on the variable indicated on the horizontal axis. The shaded area presents the 90 percent confidence intervals, assuming standard errors are clustered at the slum level. The sample includes all households in the census that are considered eligible for the study (see Supplementary Material S.4.1 for details).

In the urban slums of Uttar Pradesh, CTs are often found in poor condition. Figure S.1.2 summarizes the average status of the facilities as collected by observers at baseline (Supplementary Material S.4.2). Panel A refers to the structural status of the facility, while panel B refers to the cleanliness of the facility. CTs are characterized by the poor

quality of the construction, by the lack of hand-washing facilities, and by a general lack of cleanliness.

Figure S.1.2: Status of facilities at baseline

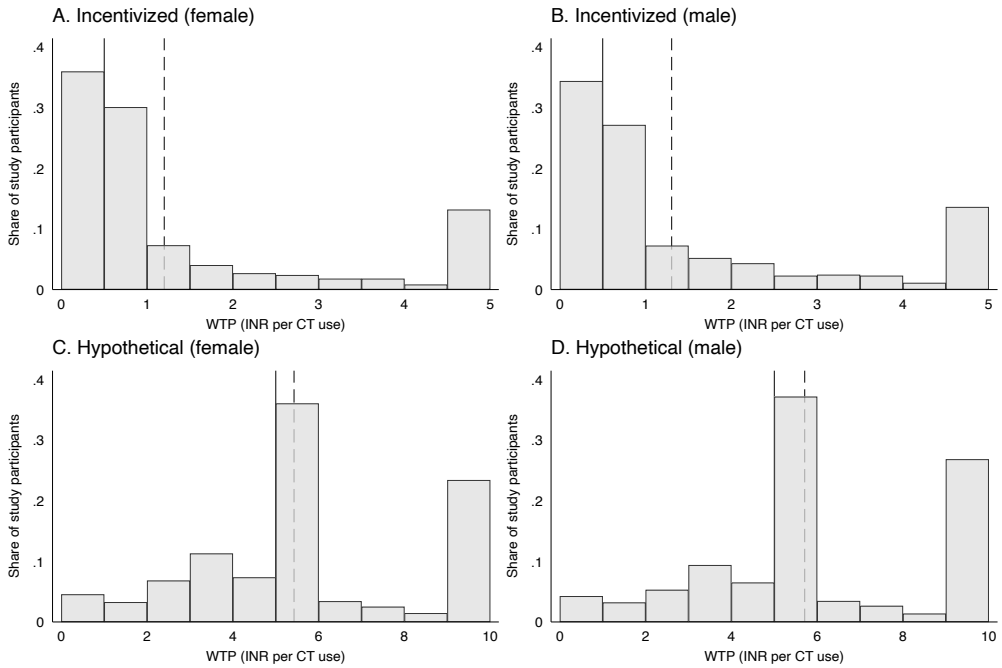


Note. Share of CTs that have or have access to the corresponding characteristic. Information is measured at baseline by observers. Supplementary Material S.4 provides details about measurement.

The role of slum residents. Less than half of households reporting that they like the services offered in the local CT, 36 percent reporting the CT is clean, 15 percent reporting that they like the facility, and 28 percent reporting they considered it safe. Figure S.1.3 shows the distribution of WTP among male and female respondents, measured using the incentivized elicitation of WTP (panels A and B) and a non-incentivized elicitation for a hypothetical higher-quality CT (see Supplementary Material S.4.5). On average, slum residents are willing to pay INR 1.40 to use the CT. WTP is slightly higher for male respondents (INR 1.46 versus 1.36 for female respondents). For a hypothetical CT with a high level of cleanliness, good hand-washing facilities, and well-lit and locked cubicles, the average WTP is INR 5.57, slightly higher for men (INR 5.71 versus INR 5.44 for women). Non-payment is often found to be rampant. Figure S.1.4 documents the distribution of the share of women and men who use the CT without paying during the rush hour. On average, only 66 percent of users pay the CT fee. This is mainly driven by women, among whom 50 percent do not pay the fee, as compared with 24 percent among men.

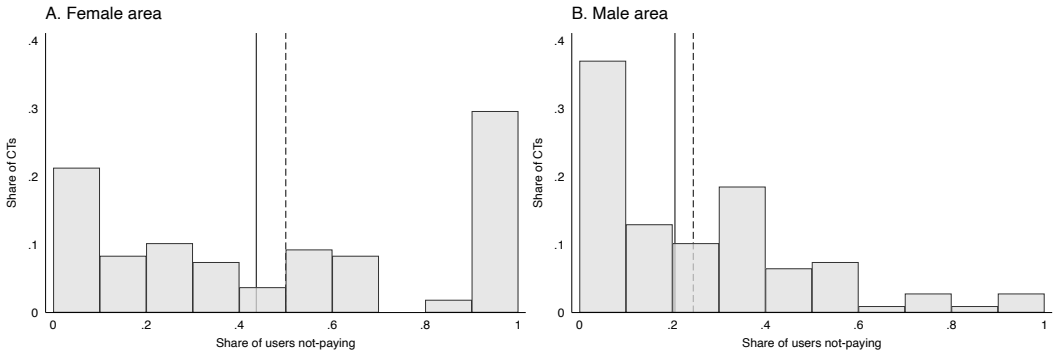
The role of caretakers. Figure S.1.5 presents an unconditional correlation matrix measured at baseline for a variety of indicators associated with the CT, including caretakers' characteristics and average characteristics of slum residents in the catchment area. Caretakers' labor supply is positively related to the opening hours of the CT and the share of time allocated to managing the CT. Caretakers who are also working as cleaners tend to be female, have lower pro-social motivation and work in worse CTs (as measured by the share of functioning toilets). Recent improvements positively correlate with the

Figure S.1.3: WTP for service use at baseline, by gender



Note. Data collected at baseline. Panel A (panel B) shows the distribution of the WTP for a single CT use among female (male) slum residents, measured using the incentivized elicitation of WTP. Panel C (panel D) shows the distribution of the WTP for a single CT use among female (male) slum residents, measured in a non-incentivized setting for a hypothetical higher-quality CT. The solid vertical lines represent the sample median, and the dashed vertical lines represent the sample mean. Additional details about the variable are presented in Appendix A. Supplementary Material S.4 provides details about measurement.

Figure S.1.4: Non-payment at baseline, by gender-specific area

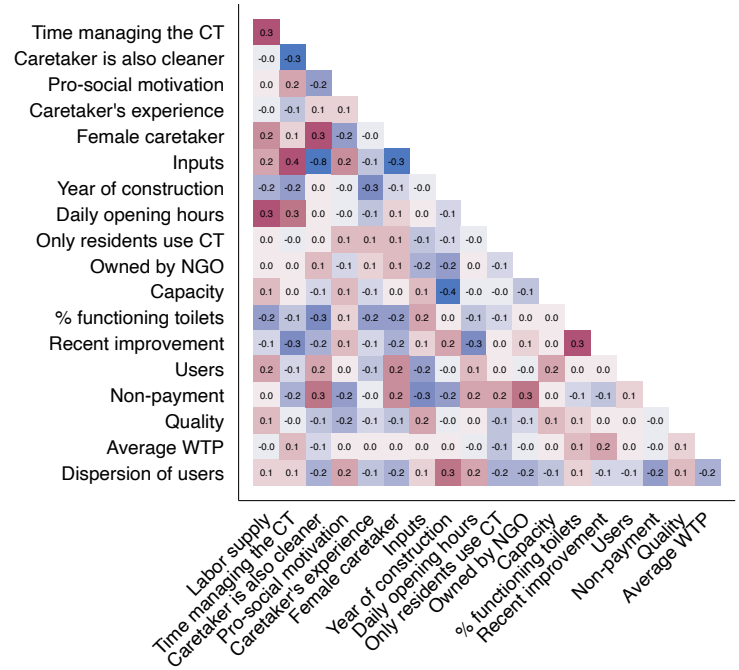


Note. Data collected at baseline. Panel A (panel B) reports the share of female (male) users who do not pay the fee for the use of the female (male) area of the CT during 1 hour at dawn (rush hour), measured by observers. The solid vertical lines represent the sample median, and the dashed vertical lines represent the sample mean. Additional details about the variable are presented in Appendix A. Supplementary Material S.4 provides details about measurement.

share of functioning toilets, which is higher where caretakers have higher experience. Non-payment is especially concentrated in CTs owned by an NGO, and in CTs used only by residents, but is lower where caretakers are male, they are more pro-socially motivated, they spend a higher share of time managing the CT, they do not work as cleaners, and they use better inputs. Overall, these statistics indicate the importance of

the caretaker for the quality of service delivery.

Figure S.1.5: Correlation matrix for caretaker and CT characteristics



Note. The figure presents unconditional correlations between CT indicators. Variables are measured at baseline. Refer to Appendix A for variable definitions. *Caretaker's experience* is the number of months the caretaker has worked in the CT. *Dispersion of users* is the square of the average distance between a potential free rider and the CT in the catchment area.

S.2 Details about the interventions

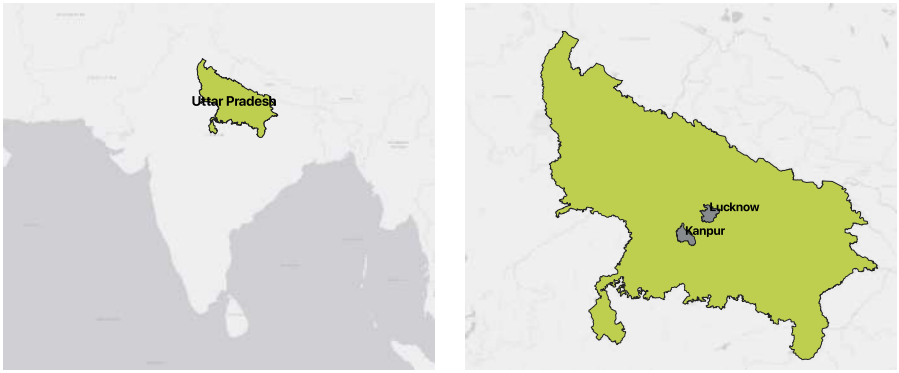
Figure S.2.1 shows the study location and the timeline of activities.

S.2.1 The maintenance intervention

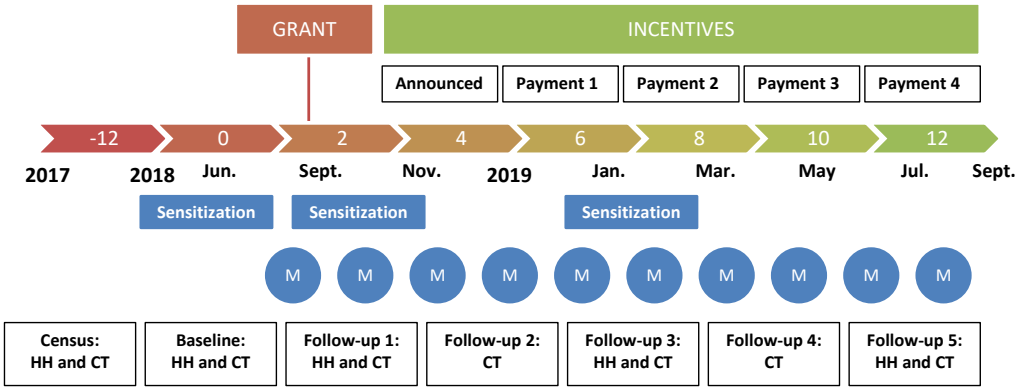
The maintenance intervention consisted of two subsequent components: a grant and a financial reward. The **grant** offered three packages of similar monetary value from which the caretaker(s) could select one. *Deep cleaning* includes septic tank sewage removal, unclogging latrines and sewerage pipes, and cleaning walls, floors and inside toilets. *Repairs* includes sanitation/water connection repairs and/or infrastructure refurbishment. *Cleaning tools and agents* included four pairs of gloves, five floor cleaners, four toilet disinfectants, five liquid soaps, four toilet-cleaning brushes, two wipes, four nose masks, two brooms, two bucket and mop sets, three detergents, two hand-washing dispensers, two dustpans and two dustbins. The training was provided as a theoretical session followed by a practical session about cleaning practices. For CTs that selected

Figure S.2.1: Location and timeline of the study

A. Location



B. Timeline



Note. Panel A shows the location of the state of Uttar Pradesh and of Lucknow and Kanpur within the state. Panel B shows the timeline of activities. M indicates the delivery of voice messages. HH and CT indicates the collection of the household and CT surveys, respectively. Basemap source: Esri (see Supplementary Material S.4.9 for details and attributions). Details about data collection activities are reported in Supplementary Material S.4.

repairs or deep cleaning, pictures of the CT area to be improved were taken before the work was done. Also, in this visit, a date was set for the works to be conducted. Based on this information, our partner FINISH arranged and supervised the work with an external contractor, which was used in all CTs.

The **financial reward** was introduced in order to improve the quality of the service rendered by the caretaker. Two months after completion of the grant scheme works, we announced the financial reward scheme to caretakers in order to incentivize them to keep the CT clean. Caretakers could receive the following rewards: INR 500 conditional on soap availability in hand-washing facilities for both genders; INR 500 conditional on visible cleanliness of latrines, defined by whether cubicles were free from visible feces (both inside and outside the latrines); INR 1,000 conditional on bacteria counts being kept to a minimum standard (i.e., being below the median of the demeaned baseline distribution by city). Caretakers were informed that an external agent would return to

Figure S.2.2: Examples of grant use

A. Pre-grant B. Post-grant

Note. Example of deep cleaning of walls and repair of locks in a CT in Lucknow. Panel A shows the status before the intervention, while panel B shows the status after the deep cleaning. Source: Antonella Bancalari.

measure each condition on a random day and time within the following two months, and that we would pay the financial reward depending on what the external agent measured. In CTs with more than one caretaker (20 percent of the sample), the financial reward was split among them. After two months and with a bi-monthly frequency, the conditions were verified by observers following a manual, and the incentives paid accordingly. In each round, we reminded the caretaker(s) of the conditions to be awarded the financial reward. In each payment round, we informed caretakers of their past cleanliness performance.

S.2.2 The sensitization campaign

The sensitization campaign targeted all members of study households, in particular the heads of participant households and their spouses. The campaign was designed in conjunction with our NGO partner FINISH to provide information that was accessible to participants with low literacy levels. We provided key messages regarding the risks of unsafe sanitation behavior and the importance of paying the fee to fund operation and maintenance of the CTs through four different means. First, **door-to-door visits** used a flip chart with cartoons and messages targeted at all household members, especially

household heads and spouses. This session covered the following sections: how open defecation affects your community; how open defecation affects your family; benefits of using CTs; what you and your family can do to make the CT better; your rights when you pay the fee for using the CT. The cartoons were made by a local graphic designer considering the context of urban slums in India. Figure S.2.3 shows the flip chart cover used for the campaign, and an example of delivery.

Figure S.2.3: Door-to-door campaign

A. flip chart cover

B. Delivery of the campaign

Note. Panel A shows the cover of the flip chart used to communicate key messages to residents in slums. It translates from Hindi as 'Awareness campaign to encourage CT use and maintenance in India". Panel B shows a moment of the sensitization campaign, in which a household head and spouse pay attention to the flip chart during a household visit in Lucknow. Source: Morsel.

Second, the main messages of the door-to-door campaign were summarized into a four-page **leaflet** (Figure S.2.4) distributed among study households. The key messages provided during the door-to-door visits were also summarized in a series of **posters** (Figure S.2.5). We placed three medium-sized and two large posters in the entrance to CTs, in the area close to the hand-washing facilities and in each gender-specific area.

Third, reminders in the form of **voice messages** were sent to participants' mobile phones. We sent a total of 10 rounds of voice message between month 1 and 11 of the study. The monthly frequency was chosen because it has been shown to be adequate to induce behavioral change (Cortes et al., 2021). This component was implemented using an purposely designed tracking app pre-populated with all mobile phone numbers. Households listened on average to 7 of the 10 monthly rounds of messages (panel A in Figure S.2.6), and listened to a good proportion of the message (panel B). More than 20 percent of the information messages highlighting public and private health risks of unsafe sanitation, as well as supply-side messages, were heard. Concerning voice message reminders, to disentangle the effects of receiving voice messages and the effect of receiving messages about the sensitization campaign, all study participants received the

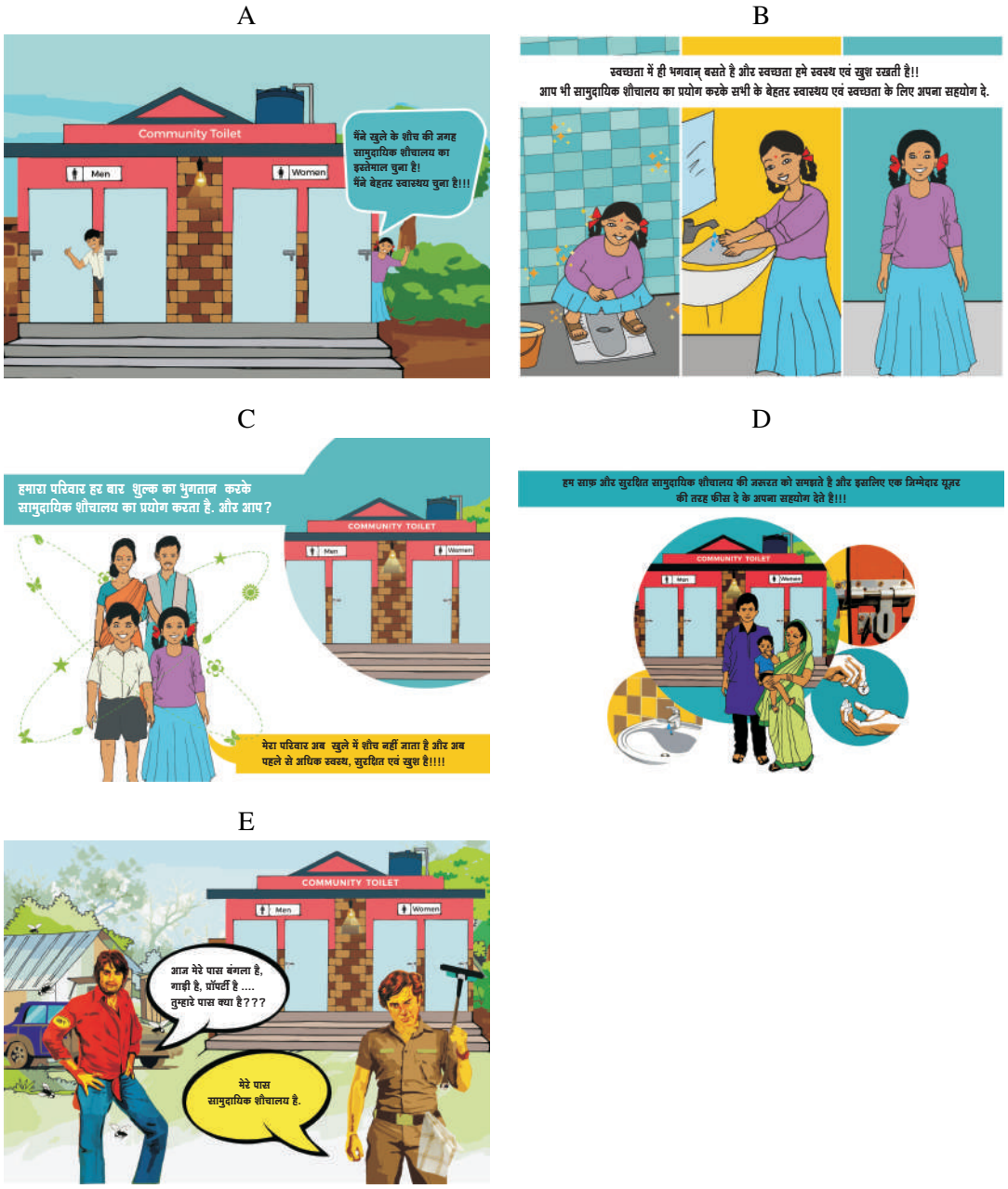
Figure S.2.4: Leaflet

Note. The figure presents the leaflet circulated during the sensitization campaign. The first page from the left presents the 'benefits of CTs' and includes: (1) improved sanitation facilities; (2) operation and maintenance of infrastructure; (3) safety with doors, locks and lights; (4) hand-washing facilities; and (5) gender-specific areas. The second page presents 'duties of users' and includes: (1) paying the fee to use the CT; (2) not throwing trash into the latrines; (3) flushing after using; (4) not spitting; (5) helping the elderly in the family; (6) accompanying females in the family during darkness; and (7) keeping the facility clean. The third page presents the 'rights of users' and includes: (1) caretakers not allowing free riders; (2) regular cleaning; (3) repairs; (4) respecting opening hours; (5) functional doors, locks and lights; (6) keeping men out of female areas; and (7) respecting and giving priority to females with children and the elderly. The final page, the cover, is the same as the one provided in the flip chart, shown in Figure S.2.3, and provides the title of the campaign.

following voice message with no content related to the sensitization campaign: *the community toilet is open from early morning until late evening*. In addition, study participants in the maintenance treatment group received the message: *your community toilet has been granted aid to improve its quality. We hope you get to enjoy this better service.* Study participants in the maintenance plus sensitization group received the following messages:

> *Do you know open defecation is one of the biggest causes of diarrhea which can even kill your children? Adopting good sanitation behavior will ensure a healthier future for*

Figure S.2.5: Posters placed on CT walls

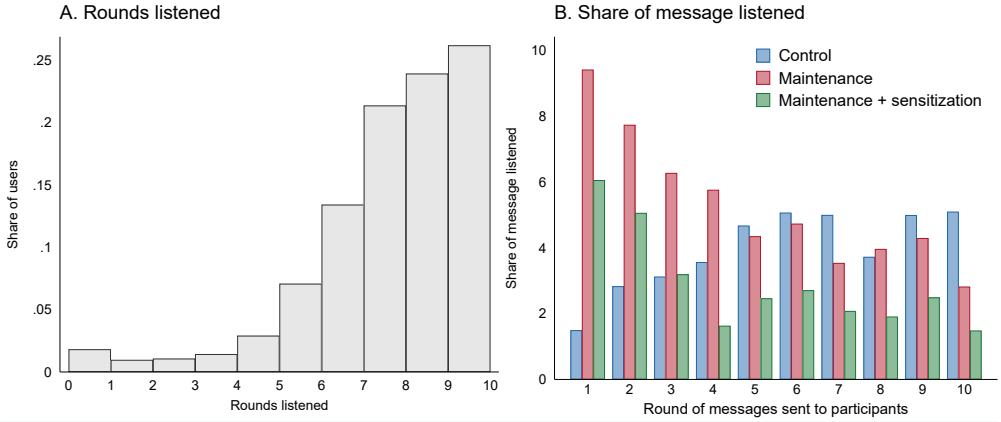


Note. The five posters placed on the walls of CTs read in Hindi: A, 'I choose to always defecate in CTs, I choose better health'; B, 'Health is happiness and cleanliness is godliness. Do your bit by using CTs'; C, 'We always pay and use CTs, do you? My family moved away from open defecation and now is healthier, safer and happier'; D, 'I value a clean and safe CT, that's why I pay the fee'; and E replicates a Bollywood scene but replacing the words to make it relevant to CTs. The villain, depicted as a dirty man says 'I have buildings, properties, vehicles, what do you have?' and the hero replies 'I have my CT'.

your family. / Open defecation is a big risk for your family's as well as your neighbors' health. Use community toilets to defecate instead of polluting and contaminating your community with open defecation. / Health is wealth! By not defecating in the open you are keeping your health safe and reducing expenses on medicines and treatment. / Cleanliness is godliness! By using community toilets, you are contributing towards the cleanliness and

health of your community. / Do you know how unsafe it is for women and girls in your family to go for open defecation? Be the change and adopt the use of community toilets. / Using community toilets ensures dignity of women in your community. Women should not feel ashamed of going to community toilets... It is way better than open defecation. / Using community toilets improves the health of your children and keeps medicines and doctors away.

Figure S.2.6: Voice messages



Note. Panel A indicates the number of rounds in which at least one household member answered the phone. Panel B shows the average share of the message that is listened to by participants, distinguishing across treatment arms and rounds. We correct for the differential duration of messages per treatment arm by multiplying the mean share of the message listened to by the maximum length per treatment arm.

S.3 Cost of intervention and quality scenarios

Table S.3.1 presents a summary of the cost associated with each activity falling under the maintenance (panel A) and sensitization interventions (panel B). Note that these are total costs throughout the project, while individual components have different timelines for implementation. Based on input from our implementing partner FINISH Society, as well as Lucknow Municipal Corporation, Table S.3.2 provides information on O&M costs for the median CT in our study sample, defined by its age (20 years), size (four female WCs, six male WCs and two urinals), and number of daily users (average of 150). Cost items include salaries for a caretaker and cleaner(s), cleaning supplies, as well as electricity and costs for minor repairs. Notice that an eligible household has the potential to provide monthly revenues of at least INR 600 if all members over 5 years of age use the CT once per day and pay the market fee of INR 5.

The monthly maintenance cost for the current scenario (which we term as 'status quo') is INR 10,200 (US$ 144.85). Under the current scenario, salaries represent 78 percent of the total budget, and cover the costs for a full-time caretaker and for one cleaner

Table S.3.1: Cost of interventions

	Total expenditure		Cost per facility	
	INR	**US$**	**INR**	**US$**
A. Maintenance intervention				
Management	324,000	4,601	4,629	66
Implementation of grant scheme	1,688,500	23,678	24,121	343
Incentives for caretakers	267,000	3,792	3,814	54
Laboratory tests	210,000	2,982	3,000	42.60
Total	**2,489,500**	**35,352**	**35,564**	**505**
B. Sensitization intervention				
Management	81,000	1,150	2,314	32.86
Design and printing of material	50,000	710	1,429	20
Door-to-door campaign	440,770	6,259	12,593	179
Voice messages	21,662	308	619	8.79
Total	**593,432**	**8,427**	**16,955**	**241**

Note. For conversion of Indian rupees into US$, we assume an exchange rate of 70.42 INR/US$. The implementation of the grant component includes subcontracting, material for repairs, human resources, transportation and the overall management of the intervention. Door-to-door campaign includes transportation costs. Cost per facility is computed assuming 70 CTs in the maintenance intervention, and 35 in the sensitization intervention. Details about the interventions are provided in Supplementary Material S.2.1 and S.2.2 .

Table S.3.2: Monthly O&M costs and grant and incentive costs per CT

	Maintenance level			
	Poor (status quo)		**Improved**	
	INR	US$	INR	US$
Panel A. O&M COSTS				
Salaries				
Caretaker (full-time)	5,000	71.00	12,000	170.41
Cleaner(s)	3,000	42.6	6,000	85.2
Supplies				
Cleaning agents	500	7.10	4,000	56.80
Cleaning equipment	200	2.84	2,200	31.24
Other				
Electricity	500	7.10	2,600	36.92
Minor repairs	1,000	14.20	2,000	28.40
Total	**10,200**	**144.85**	**28,800**	**408.97**
Total per eligible household	**300**	**4.26**	**847**	**12.03**
Panel B. INTERVENTION				
Maintenance grant				
Implementation	2,010	28.54		
Management	193	2.74		
Incentive scheme				
Amount paid to caretaker	477	6.77		
Management	289	4.11		
Laboratory tests	375	5.33		
Total	**3,344**	**47.49**		
Total per eligible household	**98**	**1.40**		
TOTAL (A + B)	**13,544**	**192.33**	**28,800**	**408.97**
TOTAL (A + B) per eligible household	**398**	**5.66**	**847**	**12.03**

Note. For conversion of INR into US$, we assume an exchange rate of 70.42 INR/US$. We assume that the grant is provided once a year and that incentives are provided on an ongoing basis every two months. We allocate 50 percent of total management cost to the maintenance grant implementation and 50 percent to the incentive scheme. To compute the total per eligible household, we consider the median number of households in the catchment area (34), and we assume no other household is using the CT.

performing a daily routine clean. We consider one alternative cost scenario that was deemed to support an 'improved' maintenance level. Under this scenario, we assume that the number of users remains constant. The scenario introduces a higher salary for

the caretaker (which allows hiring a more experienced caretaker), higher input costs, and a yearly investment into cleaning machinery, such as a pressurized water cleaner, which costs about INR 20,000 (US$ 284.01). This scenario leads to a total of INR 28,800 (US$ 408.97) per month, with salaries representing 63 percent of the total. It is important to note that we do not claim that this scenario is optimal, and it can be improved further. The table also shows cost per eligible household (see Supplementary Material S.4.1 for eligibility and proximity criteria), of which there are 34 in the median CT. In panel B of Table S.3.2 we convert the total intervention expenditures of the maintenance intervention (Table S.3.1) into monthly expenditures. Adding these costs, the total monthly costs become INR 13,544 (US$ 192.33) per CT.

S.4 Sampling, data collection and measurement

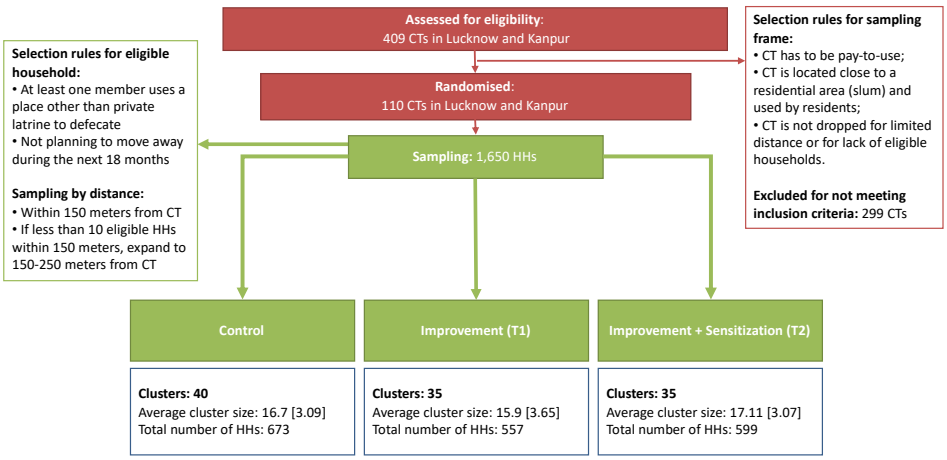
S.4.1 CT census, slum resident census, and sampling strategy

Figure S.4.1 summarizes the sampling procedure for CTs and households. In order to obtain the sampling frame, during the first half of 2017, we performed mapping of slums and a census of all CTs in both study cities. The definition of *slums* follows official criteria used in the Indian Census. According to Government of India (2011), an *identified* slum is 'a compact area of at least 300 people or about 60-70 households of poorly built congested tenements, in unhygienic environment usually with inadequate infrastructure and lacking in proper sanitary and drinking water facilities.'

The census questionnaire was administered to caretakers and/or supervisors, and collected information on the geolocation and the main characteristics of CTs, such as main users, building characteristics, ownership, management structure, and payment system. We identified a total of 201 CTs in Lucknow and 208 CTs Kanpur. Out of these, we dropped CTs with the following characteristics: free to use, located outside slum areas, and used primarily by non-residents (generally located near market areas). To avoid cases in which residents can choose between different CTs, we drop CTs that are closer than 300 meters to each other, and CTs that have two other CTs closer than 350 meters. In addition, we also dropped very small complexes, i.e. CTs in whose catchment areas are living fewer than eight eligible households. This resulted in a total of 110 CTs, which were all selected for the study.

The second level of the sampling frame is characterized by all households living in proximity to the CT without access to a private toilet. To identify these households, during the second half of 2017, we performed a census of all households living within slum bor-

Figure S.4.1: Sampling frame definition and sampling procedure

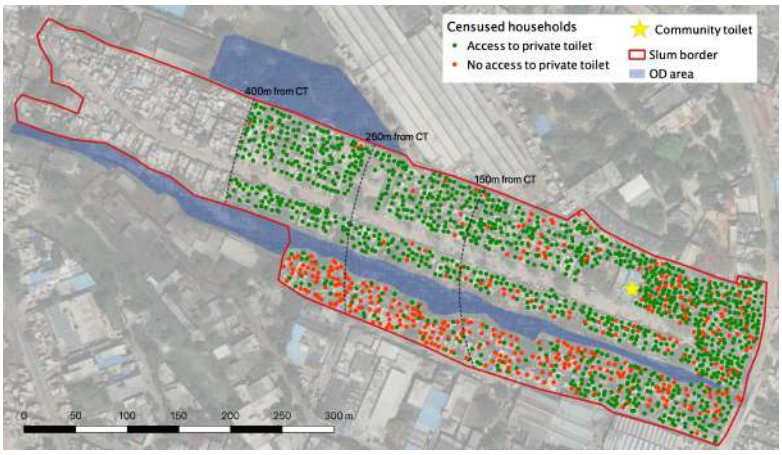


Note. The flowchart summarizes the procedure followed for the selection of CTs and the sampling of households within their catchment areas. Details of the procedure are discussed in Appendix S.4.1.

ders and within 400 meters of the selected CTs. The distance bound was selected based on qualitative evidence about the maximum distance one person would walk to opt for CT use versus open defecation (Armand et al., 2020a). The questionnaire was administered to household heads and gathered information about demographic and dwelling characteristics (including geolocation), and sanitation-related behavior for more than 30,000 households. To identify potential free riders, we defined a household to be *eligible* for the study if all these conditions are met: the household lives in the catchment area of a selected CT, defined by the area within the slum and within 150 or 250 meters in straight distance from the CT building; at least one household member reports using a CT or a shared toilet, or practicing open defecation; the household does not intend to migrate during the 18 months following the census interview. Figure S.4.2 shows an example of this selection process. Figure S.4.3 provides the spatial distribution of CTs in the study.

Within each of the 110 catchment areas, we sampled up to 17 eligible households. For catchment areas with fewer than 10 eligible households available within 150 meters, we selected all households within this bound, and randomly selected the remaining ones (up to 17 households) from the households living between 150 and 250 meters from the CT. In total, we obtained a sample of 1,650 households living in 110 catchment areas. Table S.4.3 provides a comparison between sampled households and the average characteristics of slum residents across all states of India and in Uttar Pradesh.

Figure S.4.2: Definition of sampling frame: an example



Note. The figure shows an example of the selection process for constructing the sampling frame using a hypothetical slum. Each dot represents a censused household. The area within the slum border but more than 400 meters from the CT was not covered by the census. Distance bounds are computed as straight distance from the CT. Basemap source: Esri (see Appendix S.4 for details and attributions).

Table S.4.3: Descriptive statistics of slum populations

	2011 Census of India		**Study sample**
	India	Uttar Pradesh	Lucknow and Kanpur
	(1)	(2)	(3)
A. Share of the population			
Male	0.52	0.53	0.53
Female	0.48	0.47	0.47
Children (0-6 y.o.)	0.12	0.14	0.09
Scheduled caste	0.20	0.22	0.45
B. Other characteristics			
Sex ratio (female to male)	1.08	1.12	1.12
Literacy rate	0.78	0.69	0.46

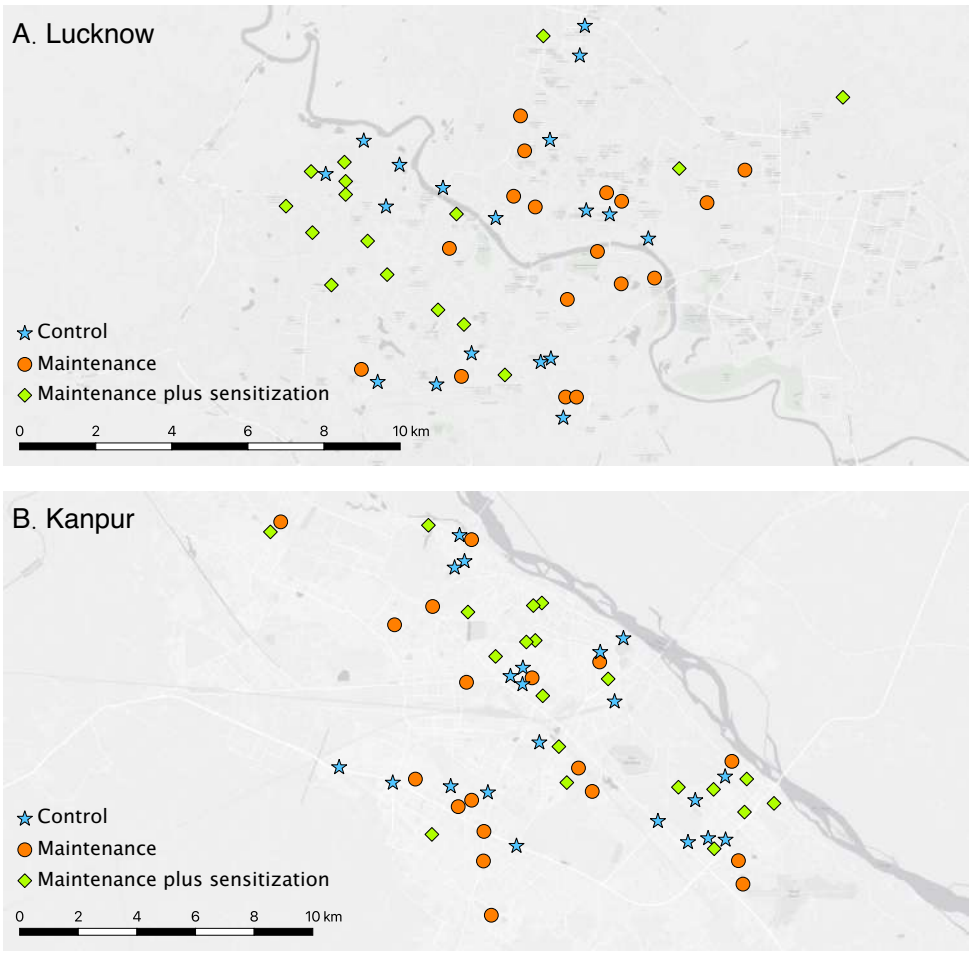
Note. The table provides descriptive statistics for the slum population in India in Column (1), for the slum population in Uttar Pradesh in Column (2), and for the study sample in Column (3). The source for Columns (1) and (2) is the 2011 Indian Slum Population Census (Government of India, 2011).

Table S.4.4: Selected CTs and households, by treatment arm and city

	Control		**T1**		**T2**		**Total**	
	N	%	N	%	N	%	N	%
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
A. CTs								
Lucknow	19	36.5	17	32.7	16	30.8	52	100
Kanpur	21	36.2	18	31.0	19	32.8	58	100
Total	40	36.4	35	31.8	35	31.8	110	100
B. Households								
Lucknow	255	35.5	225	31.3	239	33.2	719	100
Kanpur	321	37.5	262	30.6	273	31.9	856	100
Total	576	36.6	487	30.9	512	32.5	1,575	100

Note. The table presents the distribution of selected CTs (panel A) and households (panel B) by treatment arm and city. *T1* indicates the 'maintenance' treatment group, and *T2* indicates the'maintenance + sensitization' treatment group.

Figure S.4.3: Geographical distribution of CTs, by city and treatment group



Note. Panel A shows the geographical distribution of CTs selected for the study in the city of Lucknow. Panel B shows the geographical distribution of CTs selected for the study in the city of Kanpur. Details about the procedure to select CTs is provided in Supplementary Material S.4.1. Basemap source: Esri (see Supplementary Material S.4.9 for details and attributions).

S.4.2 CT surveys and observation

During regular unannounced visits to the CTs, we administered a questionnaire to the caretaker to collect data on cleaning practices, CT management and time allocation to different tasks. In addition to self-reported information from the caretakers, we also gathered information about the condition and cleanliness of CTs using observers. To provide uniform reports from observers, the data collection manual defined conditions for the visual evaluation of the state of the CT. Observers also recorded the number of users and the share of users who pay the fee for the duration of 1 hour using manual counters. We collected CT-level data in a sequence of six waves.

We collected data for all of the 110 selected CTs at the baseline, but only for 108 in follow-up 1, 109 in follow-up 2, 107 in follow-up 3, 105 in follow-up 4 and 106 in

follow-up 5, given that some CTs closed temporarily/permanently for refurbishment, and one slum was completely displaced after follow-up 2. In some cases, we were able to collect observations and bacteria swabs, while not being able to survey caretakers. In 92 percent of CTs, we surveyed caretakers in all five follow-up rounds. In addition, two new CTs opened very close to the study toilets, both allocated to the maintenance (T1) treatment arm. Because households in the catchment area also used these new CTs, we collected data from these toilets during follow-ups 2 to 5. These new CTs did not increase the number of clusters, since we consider them part of the same cluster as the old CTs given their close proximity. Table B4 shows that the number of CT observations and caretaker surveys and the addition of CTs in study catchment areas are orthogonal to treatment allocation.

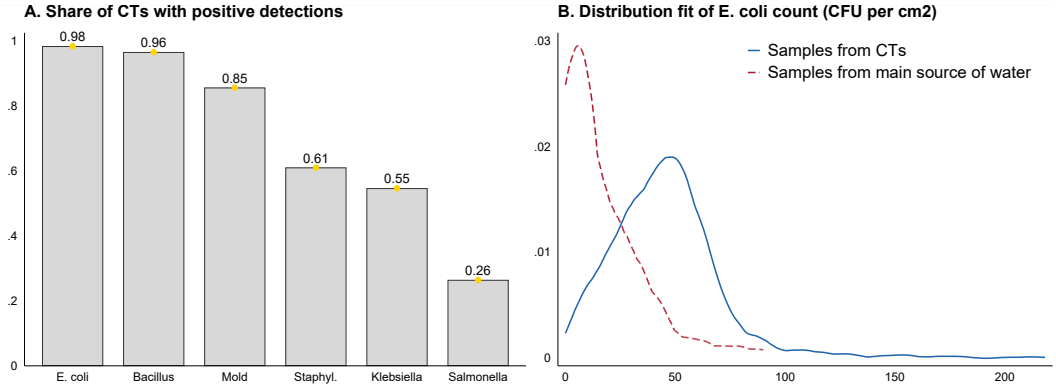
S.4.3 Laboratory tests

We collected data about bacteria and mold presence using samples analyzed in the laboratory. We first focus on the presence of the species *Escherichia coli* (*E. coli*) of genus *Escherichia*, an indicator of fecal contamination, measured as bacteria count (CFU per cm^2) using the arithmetic mean among all samples collected in a CT during a measurement round (see, e.g., WHO, 2017). We also compute the presence of potentially harmful bacteria of the genus *Bacillus*, genus *Staphylococcus*, genus *Klebsiella*, and genus *Salmonella*. For further information on the effect of bacteria on human health, see Jenkins and Maddocks (2019). In addition, we test for the presence of mold, which can cause allergic reactions and respiratory problems (Gent et al., 2002).

To implement these measurements, we prepared a protocol in conjunction with a laboratory based in Lucknow, which analyzed the samples. For each CT and during each survey round, three samples were collected using swabs in specific locations of the facility based on evidence about the microbial bio-geography in public toilets (Flores et al., 2011). CTs were first randomized into two groups: a *male* group, in which the swabs were collected in the male area of the CT throughout the study, and a *female* group, in which the same was performed in the female area of the CT. During each visit, the enumerator collected three samples. The first two samples were collected from the floor of the cubicles at the mid-point between the entrance wall and the latrine/water. Cubicles were randomly selected by the research team in each round to avoid the caretaker focusing on a specific point in the CT. A third sample, aimed at collecting information about the area where most people walk, was collected from the floor where one would take one's first step to enter the cubicle hallway.

At baseline, we also collected information about access to clean water. We collected and analyzed two samples of water for each catchment area. During the baseline survey, we asked households about their main source of water, and we then collected water samples from up to two randomly selected sources. Figure S.4.4 shows descriptive statistics at baseline for these measurements.

Figure S.4.4: Bacteria and mold detection at baseline



Note. Panel A presents the share of CTs where each bacteria type or mold was detected in at least one of the three samples. Panel B shows the distribution of the E. coli count from CT and water samples. The distribution fits are estimated non-parametrically using kernel density estimation assuming an Epanechnikov kernel function. Bandwidths are estimated by Silverman's rule of thumb (Silverman, 1986).

S.4.4 Slum resident survey

We collected household-level data in a sequence of four waves (refer to Figure S.2.1 for the timing and label of each wave). This was a standard household survey, collecting information on demographic and socio-economic characteristics, such as household composition, dwelling characteristics, assets, income and expenditure. This information was supplemented by a section on health and sanitation behavior, including attitudes and expectations associated with different sanitation practices. At baseline, we further collected information on child health for children under the age of 6, and on intra-household dynamics between spouses. The instrument, including all modules, has an average duration of one hour. All follow-up surveys aimed to recollect some of the same information collected at baseline. Some information collected at baseline was not collected during the follow-up surveys, such as detailed information on child health and childcare, and intra-household decision-making. Some new information was collected after the baseline survey, such as information related to exposure to the interventions.

The respondent is the main decision maker in the household. We select the respondent using the following rules: if the household head is present, then the respondent is the household head; if the household head is absent, then the respondent is the spouse of

the household head; if the household head and spouse are both absent, the household is revisited; if the household head and spouse are both absent during the revisit, then the respondent is the most senior member (over 18 years old) who is actively participates in the household's decision-making. At baseline, we also interviewed the spouse of the household head to gather information about intra-household decision-making, and the primary caregiver to collect information on child health for children aged 5 years or younger.

In total, we interviewed 1,575 households at baseline (an average of 12 households per cluster), 1,532 households during follow-up 1, 1,578 households at follow-up 3, and 1,772 households in the follow-up 5. Each baseline household was interviewed in 2.6 out of 3 follow-up measurements, with only 2 percent of baseline households that was never re-interviewed at follow-up. In terms of attrition from baseline to a specific follow-up surveys, the rate is 9 percent for follow-up 1, 19 percent for follow-up 3, and 14 percent for follow-up 5.[2] Columns (1)–(5) in Table B5 estimates the probability of attrition for each of these indicators as a function of the treatment status. Attrition does not differ between treatment and control groups for any of the attrition indicators. In order to maintain a comparable sample size in all follow-up surveys, we handled attrition with replacements at random using the sampling frame used for the baseline sampling. Column (6) tests whether the replacement was introduced differently across treatment arms. In total, 16 percent of follow-up observations are replacements, with no statistical difference across treatment arms.

The questionnaire for follow-up 5 was supplemented with a list randomization technique. Respondents were randomly allocated to one of four groups. Depending on the group, respondents received a different list of statements, and were asked to report how many of them were true. Table S.4.5 provides the list of statements. Group A received only a list of statements related to general behavior. Groups B–D received the same list and one extra statement capturing sensitive behavior.

Table S.4.5: Statements used for list randomization

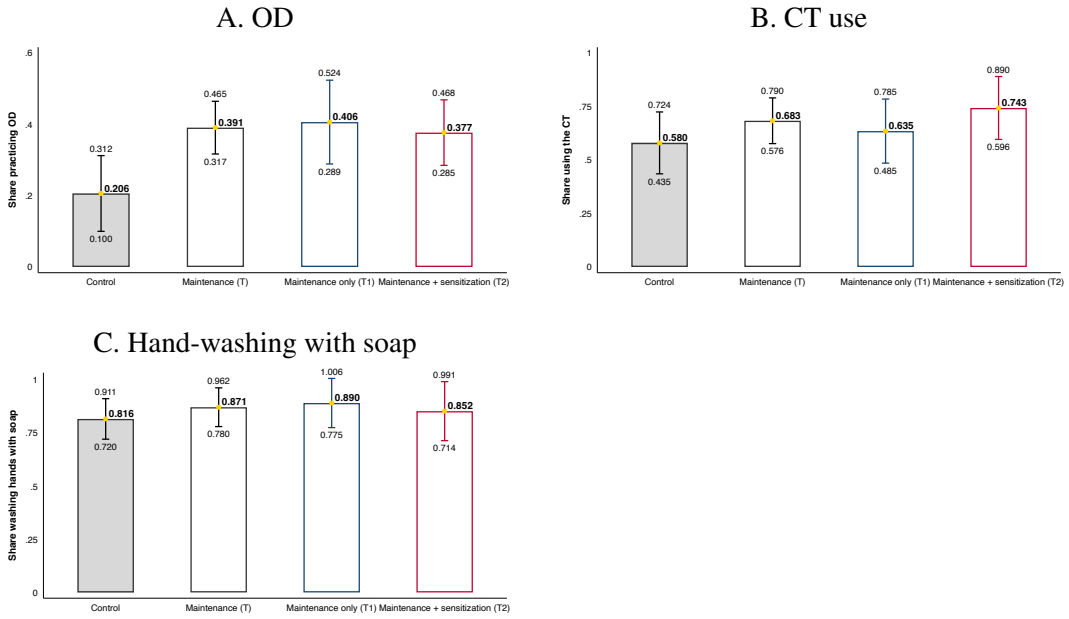
Group A	Group B	Group C	Group D
- I cooked yesterday	- I cooked yesterday	- I cooked yesterday	- I cooked yesterday
- I bought milk yesterday	- I bought milk yesterday	- I bought milk yesterday	- I bought milk yesterday
- I watched TV yesterday	- I watched TV yesterday	- I watched TV yesterday	- I watched TV yesterday
	- I defecated in the open yesterday	- I used the CT to defecate yesterday	- I washed my hands with soap yesterday

Note. Group A reports a list of statements related to general behavior. Groups B–D provide the same list, but adding one extra statement capturing sensitive behavior (OD, use of CT, or hand-washing).

[2]The attrition rate was the highest at follow-up 3 because the survey coincided with school vacations, a period when a share of slum residents goes back to their native villages.

Using the information from the list randomization questions, Figure S.4.5 shows the share of slum residents practicing OD, using the community toilet and hand-washing with soap, respectively, in the day previous to the interview. We observe a small insignificant increase in service use driven mostly by the sensitization campaign. Although the list randomization technique does not identify individual behavior, this result can be explained by a shift away from service use and into OD, balanced by a shift away from unimproved (private) sanitation and into CT use. In addition, users could have practiced both.

Figure S.4.5: Treatment effects on list randomization outcomes



Note. Confidence intervals are built using statistical significance at the 10 percent level and assuming errors are clustered at the level of the catchment area. Randomization of lists was performed at individual level, and data were collected during follow-up 5 only. Supplementary Material S.4 provides details about measurement.

Self-reported sanitation behavior was measured by asking survey respondents where each demographic group defecated the last two times. To prevent under-reporting of open defecation due to social stigma, we included the following prelude: 'I've been to many similar communities and I've seen that even people owning latrines and having nearby community toilets defecate in the open.'

S.4.5 Lab-in-the-field experiment: WTP for service use

WTP for service use is elicited to the respondent of the household survey and the spouse (up to two respondents per household), and is measured 4 times during the study in conjunction with the household survey. WTP is elicited using a standard incentivized version of the multiple price list (or take-it-or-leave-it) methodology. Participants were

prompted to choose between different amounts of cash (ranging from INR 0 to 60 with increases of INR 5) and a bundle of 10 tickets to use the CT in the catchment area where they live. During follow-up 5, following the introduction at the end of the study of monthly passes for families in a limited number of CTs, we also elicited the WTP for a monthly family pass for up to five members. In total, participants face 13 combinations. After all choices are made, one of the options is then randomly selected by drawing a numbered ball from a bag, and the decisions are realized. Following the realization of the game, in the case of the bundle of tickets being assigned, the respondent could allocate the 10 tickets or some of them to either male or female use. Before participating in the game, the participant was introduced to a practice round of the game using a bar of soap to facilitate familiarity with the rules. The exact explanation of the game read by the enumerator to the participant was as follows:

> *Now let us do the prize draw for 10 tickets to use the [CT name]. These tickets are being officially provided by [CT name] as a promotion to encourage people to use the CT. They can be used at any time in the next 2 months. You will be given the choice later to decide how many of the 10 tickets you would like to be for men and boys, and how many you would like to be for women and girls. We are going to ask you to make a series of choices between either receiving these 10 tickets or instead receiving amounts of cash. At the end of all of the choices, you will draw a ball from a bag to determine which one of these choices will be randomly selected for your lucky draw – you will get the tickets or the money, depending on what you chose. This means that any one of the choices that you make could be selected at the end. Therefore it is in your best interest just to answer your honest opinion about which option you would prefer in every single choice.*

In conjunction with the incentivized version of the WTP elicitation, we also collected information from participants about the price that female and male residents face to use the CT, and we asked about WTP for the use of a hypothetical higher-quality CT in a non-incentivized setting. For the latter, participants were asked directly whether they would purchase a ticket for different amounts of money, ranging from INR 0 to 10, with increases of INR 1. The exact question reads as follows:

> *Now I want to tell you an imaginary story. Imagine that starting from tomorrow, the owners of the nearest CT decided to change the price for using the defecation cubicles. At the same time, they would improve the quality of the CT to the highest standard, ensuring it was very clean, had good hand-washing facilities, and that all the cubicles had a light and a lock. Would you be willing to buy a ticket to use the defecation cubicles of the community toilet, if the price was...*

S.4.6 Lab-in-the-field experiment: adapted dictator game

To measure preference for maintenance among slum residents, we played an adapted dictator game in which participants are endowed with INR 50 and are given the option to donate all or part of it to a fund to purchase cleaning products for the CT. This component was administered to the respondent of the household survey and the spouse (up to two respondents per household), and measured in conjunction with each household survey. Having collected all the contributions to the cleanliness of the CT within each slum, the total amount was used to purchase cleaning products, which were then delivered to the caretaker. The exact setting reads as follows:

> *I would like to inform you that as an additional thank-you for participating in this study, you will receive an extra INR 50 in cash. We are asking all participants to choose between keeping some or all of this INR 50 for themselves, and donating some or all of this INR 50 for a special fund for cleaning products that we will deliver to the CT. How would you like to split the INR 50 between cash for yourself, and donation to the cleaning product fund for your CT?*

Similarly, to measure pro-social motivation for the cause among caretakers, we implemented an adapted dictator game in which the caretaker is endowed with INR 50 and is given the option to donate all or part of it to fund a sanitation project implemented by our partner, FINISH Society. Pro-social motivation among caretakers was measured during each CT survey. Having collected the contributions from all caretakers, the total amount was donated to the FINISH Society project. The exact setting faced by the caretakers reads as follows:

> *I would like to inform you that as a thank-you for participating in this study, you will receive INR 100 in cash. You can keep the full amount for yourself or you have the opportunity to donate some or all of it to FINISH Society to help with improving water access, sanitation and hygiene in disadvantaged areas of India. How would you like to split the INR 100 between cash for yourself and donation to charity?*

S.4.7 Lab-in-the-field experiment: public goods game (PGG)

We implemented a standard public goods game with the experiment participants. The game is based on the voluntary contribution mechanism, in which participants receive an endowment of INR 100, and they have to decide whether to keep the endowment or to invest part or all of it in a public pot. The contributions in the group are increased by a multiplier and shared equally among participants. The multiplier is randomly varied at catchment-area level to either double or triple the contributions. The game is designed so that while the total return to the investment in the pots is higher than the return

from keeping the endowment, there is no incentive to invest in the former because of the higher individual pay-off that can be obtained from keeping the endowment. The dominant strategy is therefore to not contribute at all, while the social optimum is to invest in the pot. We played simultaneously with three groups of equal size (ranging from four to six participants) in each community. Participants also received INR 20 as show-up fee. The instructions given to groups of six participants are the following (note that x is either 2 or 3):

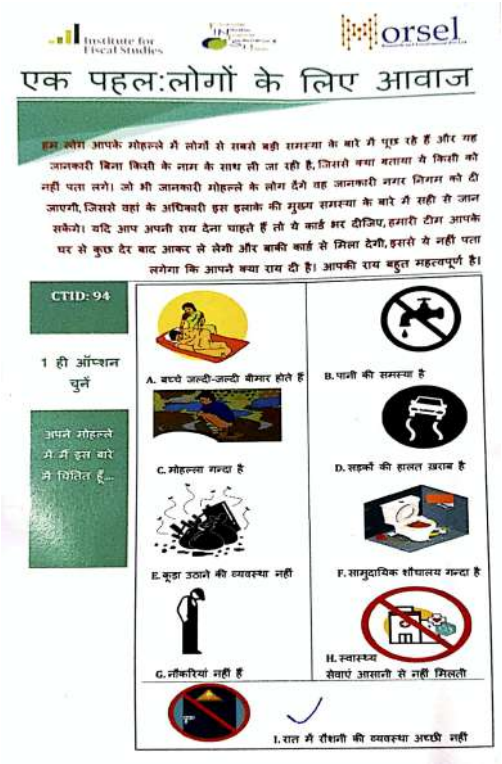
> *In this game, each player receives an endowment of INR 100 and you can choose to contribute (C) to the shared pot or keep (K) it. Out of the INR 100, you can decide how much to contribute and how much to keep. Secretly, you will put your donation amount in the pink envelope and the amount you want to keep in the blue envelope. All contributions will be summed and we will increase the total contribution by [x]. The final pot will be split equally among players. Let's look at some examples. If all 6 players contribute the INR 100, their individual payoffs would be equal to INR [$600 \cdot x/6$]; if one player contributes and other players keep the endowment, then the payoff of each player contributing is equal to INR [$pot \cdot x/6$], and the payoff of the player keeping is equal to INR [$100 + pot \cdot x/6$]; if all players keep, then their individual payoffs are INR 100.*

S.4.8 SCA: voice-to-the-people initiative

We distribute a card (Figure S.4.6) to participants asking about the most pressing issue in their community. Participants were invited to circle only one option among the following: children are frequently ill, water availability is limited, the community is dirty, the quality of roads is poor, there is no waste collection, the CT is dirty, jobs are missing, access to healthcare is limited, and lighting at night is poor. We used visual representations to facilitate selection of the issue among illiterate participants. Individual anonymized codes allowed the returned cards to be matched with the households in the sample. Two letters with the results for the corresponding cities were sent to the heads of sanitation and environment at the municipal corporations of Lucknow and Kanpur in October 2019. The specific instruction reads as follows:

> *We are collecting anonymous reports about the most pressing issue in your community. We will communicate this to the district municipal corporation to raise awareness among administrators. If you want to tell us your opinion, fill in the card and return to one of our team members. Your voice is important! In my community, I am most concerned about: [circle only one option]*

Figure S.4.6: Card distributed for the voice-to-the-people initiative



Note. The figure shows the card distributed to participants as part of the voice-to-the-people initiative. The options read in Hindi: children frequently ill (A); limited water availability (B); community is dirty (C); poor quality of roads (D); no trash collection (E); CT is dirty (F); no jobs (G); limited access to healthcare (H); poor lighting at night (I). In each household, up to two participants had the option to mark one of the issues.

S.4.9 Additional data sources

Basemaps throughout the paper were created using ArcGIS® software by Esri®. Basemaps are used in line with the Esri Master License Agreement, specifically for the inclusion of screen captures in academic publications. We use the *World Light Gray Base* (sources: Esri, HERE, Garmin, ®OpenStreetMap contributors, and the GIS User Community).

Appendix and Supplementary Material Bibliography

Andersen, S., Harrison, G. W., Lau, M. I., and Rutström, E. E. (2006). Elicitation using multiple price list formats. *Experimental Economics*, 9(4):383–405.

Armand, A., Augsburg, B., Bancalari, A., and Trivedi, B. (2020). Community toilet use in slums: Willingness-to-pay and the role of informational and supply side constraints. *3ie Impact Evaluation Report 113*.

Athey, S., Tibshirani, J., and Wager, S. (2019). Generalized random forests. *Annals of*

Statistics, 47(2):1179–1203.

Athey, S. and Wager, S. (2019). Estimating Treatment Effects with Causal Forests: An Application. *Observational Studies*, 5.

Basu, S., Kumbier, K., Brown, J. B., and Yu, B. (2018). Iterative random forests to discover predictive and stable high-order interactions. *Proceedings of the National Academy of Sciences of the United States of America*, 115(8):1943–1948.

Belloni, A., Chernozhukov, V., and Hansen, C. (2013). Inference on treatment effects after selection among high-dimensional controls. *Review of Economic Studies*, 81(2):608–650.

Chernozhukov, V., Demirer, M., Duflo, E., and Fernández-Val, I. (2017). Generic Machine Learning Inference on Heterogenous Treatment Effects in Randomized Experiments. Technical report, National Bureau of Economic Research.

Cortes, K. E., Fricke, H., Loeb, S., Song, D. S., and York, B. N. (2021). Too little or too much? Actionable advice in an early-childhood text messaging experiment. *Education Finance and Policy*, 16(2):209–232.

Flores, G. E., Bates, S. T., Knights, D., Lauber, C. L., Stombaugh, J., Knight, R., and Fierer, N. (2011). Microbial biogeography of public restroom surfaces. *PLoS ONE*, 6(11):1–7.

Gent, J. F., Ren, P., Belanger, K., Triche, E., Bracken, M. B., Holford, T. R., and Leaderer, B. P. (2002). Levels of household mold associated with respiratory symptoms in the first year of life in a cohort at risk for asthma. *Environmental Health Perspectives*, 110(12):A781–A786.

Gordon, D., Howe, L. D., Galobardes, B., Matijasevich, A., Johnston, D., Onwujekwe, O., Patel, R., Webb, E. A., Lawlor, D. A., and Hargreaves, J. R. (2012). Authors' response to: Alternatives to principal components analysis to derive asset-based indices to measure socio-economic position in low- and middle-income countries: The case for multiple correspondence analysis. *International Journal of Epidemiology*, 41(4):1209–1210.

Government of India (2011). Census of India 2011. Office of the Registrar General and Census Commissioner, Ministry of Home Affairs, Government of India, New Delhi, India.

Jenkins, R. and Maddocks, S. (2019). *Bacteriology methods for the study of infectious*

diseases. Academic Press.

Kline, T. (2014). *Psychological Testing: A Practical Approach to Design and Evaluation*. Sage Publications.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*, volume 26. CRC press.

Tibshirani, R. (1996). Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288.

WHO (2017). Guidelines for drinking-water quality, 4th edition: 1st addendum. Guidelines Review Committee – Water, Sanitation, Hygiene and Health, World Health Organization April 2017, World Health Organisation.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. MIT Press.