Alex Armand
Britta Augsburg
Antonella Bancalari
Bhartendu Trivedi

Community toilet use in Indian slums

Willingness-to-pay and the role of informational and supply side constraints

March 2020

Impact Evaluation Report 113

Water, sanitation and waste management

International Initiative for Impact Evaluation

About 3ie

The International Initiative for Impact Evaluation (3ie) promotes evidence-informed, equitable, inclusive and sustainable development. We support the generation and effective use of high-quality evidence to inform decision-making and improve the lives of people living in poverty in low- and middle-income countries. We provide guidance and support to produce, synthesise and quality assure evidence of what works, for whom, how, why and at what cost.

3ie impact evaluations

3ie-supported impact evaluations assess the difference a development intervention has made to social and economic outcomes. 3ie is committed to funding rigorous evaluations that include a theory-based design and that use the most appropriate mix of methods to capture outcomes and are useful in complex development contexts.

About this report

3ie accepted the final version of the report, *Community toilet use in Indian slums: willingness-to-pay and the role of informational and supply side constraints*, as partial fulfilment of requirements under grant DPW1.1105 awarded through Development Priorities Window 1. The report is technically sound and 3ie is making it available to the public in this final report version as it was received. No further work has been done.

The 3ie technical quality assurance team for this report comprises Francis Rathinam, Neeta Goel, Deeksha Ahuja, an anonymous external impact evaluation design expert reviewer and an anonymous external sector expert reviewer, with overall technical supervision by Marie Gaarder. The 3ie editorial production team for this report comprises Anushruti Ganguly and Akarsh Gupta.

All of the content is the sole responsibility of the authors and does not represent the opinions of 3ie, its donors or its board of commissioners. Any errors and omissions are also the sole responsibility of the authors. All affiliations of the authors listed in the title page are those that were in effect at the time the report was accepted. Please direct any comments or queries to the corresponding author, Alex Armand at: alex.armand@novasbe.pt.

Funding for this impact evaluation was provided by UK aid through the Department for International Development. A complete listing of all of 3ie's donors is available on the 3ie website.

Suggested citation: Armand, A, Augsburg, B, Bancalari A and Trivedi B, 2020. *Community toilet use in Indian slums: willingness-to-pay and the role of informational and supply side constraints*, 3ie Impact Evaluation Report 113. New Delhi: International Initiative for Impact Evaluation (3ie). Available at: https://doi.org/10.23846/DPW1IE113

Cover photo: Juggadery / Flickr

Community toilet use in Indian slums: willingness-to-pay and the role of informational and supply side constraints

Alex Armand
NOVA School of Business and Economics, Navarra Center for International Development and Institute for Fiscal Studies

Britta Augsburg
Institute for Fiscal Studies

Antonella Bancalari
London School of Economics and Institute for Fiscal Studies

Bhartendu Trivedi
Morsel India

Impact Evaluation Report 113

March 2020

International Initiative for Impact Evaluation

Acknowledgements

Many have contributed to the success of this report. We would like to acknowledge in particular the inspiring support of:

- All enumerators and supervisors that have spent countless hours working in the field;
- All collaborators of FINISH for their precious and highly-motivated assistance during the study design and during the implementation of the interventions.

Without the contribution of all of the above the success of this report would not have been possible.

Summary

While urbanization can bring benefits for economic, cultural and societal development, cities in low and middle-income countries are struggling to keep up with necessary infrastructure investment. In particular, in urban slums, the water, sanitation, and hygiene infrastructure is stressed beyond current capacity. Since inadequate sanitation is threatening to the general population, and in particular to early-life health, when coupled with high population density (Hathi et al. 2017), efforts to improve sanitary conditions in slums are of direct policy relevance. In the setting of urban slums, it is generally accepted that public and community toilets (CTs) will, for the foreseeable future, continue to be an important solution to improve sanitary conditions, given numerous constraints to increasing access to safely-managed private toilets. However, even where CTs are available, open defecation remains common behaviour among slum-dwellers (Indian Ministry of Home Affairs 2011).

This study implements a randomized-controlled trial (RCT) in the context of Indian slums in the cities of Lucknow and Kanpur (Uttar Pradesh, India). Slums in these districts are similar to other slums in India along a number of margins, including the demographic, literacy and caste composition. We select a total of 110 catchment areas of CTs in both cities and randomly allocated them to one of three experimental arms: Treatment 1 group, Treatment 2 group, and a control group. In the Treatment 1 (Supply-side intervention) group, we promote cleanliness and maintenance of CTs by introducing a *grant scheme* that offers standardized (i.e. same monetary value) packages to improve the quality of the CT, and by afterwards providing financial rewards aimed at incentivizing caretakers to keep the CT clean. In the Treatment 2 (Supply-side plus Information intervention) group**,** in addition to the intervention provided for Treatment 1, a household-level information campaign is provided. In the control group, participants do not receive supply-side nor information interventions. This design, combined with several rounds of data collection over a period of 18 months, both at the household and the CT level, allows us to address our three research objectives, which are (1) to document slum dwellers' willingness to pay (WTP) for CTs and its link with CT usage; (2) to identify the impact of releasing supply-side or both supply-side and informational constraints on WTP and usage; and (3) to determine the time horizon of such impacts (short-term, longer-term, or both).

The baseline report provided already an overview with respect to the first objective. Two key findings are worth reiterating. For one, we found that slum dwellers' WTP (using incentivized elicitation methods) was substantially below the market rate for CT usage, both for male and female dwellers. Second, we found that WTP for a hypothetical high-quality CT is substantially higher than for current standard CTs. acknowledging the drawbacks of un-incentivized elicitation (as it is the case of the hypothetical WTP), these results suggests potential for our treatments to increase WTP for CT usage. We discuss our findings with respect to impacts in this report, addressing objectives 2 and 3.

We conduct additional analysis with respect to measured WTP. A few findings stand out in particular with respect to gender differences: Men's and women's WTP for using CTs are determined differently. We find that men are willing to pay less for a hypothetical high-quality CT when local CTs are currently of bad quality, but this does not seem to affect women. When eliciting the WTP for using the current CT, we find that men have a higher WTP for CTs if a higher proportion of people in a 100m radius believe that

practicing OD is shameful. Yet, this social norm is not important for women, who instead attach a much higher importance to whether or not they have a toilet at home. We further find that both men and women are less willing to contribute to the CT's cleaning when it is already cleaned adequately. We also find evidence of *screening effects*: the share of households using CTs per catchment area is a function of the mean WTP of study households. The relationship is sharper when looking at the WTP of men than women, suggesting that men's WTP is a stronger determinant of household CT usage. Other strong determinants of CT usage - in line with existing literature - is distance from the household to the local CT and people's attitudes towards CTs. Moreover, expectations of health benefits from using CTs instead of practicing OD are negatively associated with OD practice. Notably, we observe that CTs which were adequately cleaned are paid for more often. When looking at the predictors of CT quality, we find that how adequately a CT is cleaned is a key predictor of the number of bacteria and observed cleanliness.

Randomization and interventions were implemented in line with the research protocol. Treatment groups are balanced on a vast array of characteristics, supporting the ex-ante comparability of CTs allocated to different groups. In addition, intervention uptake was very high. In terms of the supply intervention, 95% of CTs chose a grant scheme (of these, 41% repairs, 41% deep cleaning, and remaining 12% training and provision of cleaning agents and tools) and about 95% participated in the financial reward scheme (97% received the announcement and were available for the first payment, and 94% for the second and third payment). Similarly, the uptake of the information campaign was high: door-to-door information sessions and provision of leaflets was conducted in almost all households targeted across the three rounds, and households listened to an average of 7 voice messages (sent to mobiles) about health risks from open defecation. Further, all CTs allocated to the information treatment arm had informative posters on their walls, which users were hence exposed to. We demonstrate in this report that exposure to the information campaign resulted in significant effects on households reporting that they listened to voice messages about water, sanitation and hygiene (WASH) through communication technologies; that they saw posters in the community toilet; that they were visited by an agent talking about WASH; and, that they heard of, and for some attended, events organized in the community about WASH.

Despite the significant exposure, the treatments do not translate into an increase in the willingness to pay (WTP) for community toilet use during the whole duration of the study. If any, while the supply-side treatment leads to small (insignificant) increases in WTP, we do observe short-run negative effects on WTP driven by supplementing the supply-side intervention with the information campaign. The information campaign could have affected the extent to which households internalize the negative externalities from inadequate sanitation behaviour. At the same time, we find that the treatments increase the share of users that pay the fee to use the CT. This effect is driven by both male and female users, though the effect for males is more precisely estimated. We bring forward two possible explanations for this finding: On the one hand, the financial reward component of the supply intervention, which promised cash to the caretaker conditional on keeping the CT clean and having soap available, may have incentivized caretakers to collect resources (i.e. make (more) users pay) to achieve these goals. On the other hand, the information campaign may have reinforced the need to pay fees for households that were already regular users of the CT.

Looking at secondary outcomes, we find a short- and medium-run effect of the supply side treatment on hand-washing with soap– a potential result of the grant and incentive scheme. Impacts on other secondary outcomes, as well as remaining primary ones (in particular self-reported CT usage) are negligible.

Even though we observe small changes in individual behaviour, we document that the treatments increase the salience of CT cleanliness when reporting issues to policymakers. When given the opportunity to fill-in "voice-to-the-people" style of cards, in which they are given the possibility to report the most important concerns from a set of different topics to local public officers, more treated households report the dirtiness of the CT as their primary concern. This result highlights the role of policymakers in guaranteeing cleanliness of CTs by solving the potential coordination problem in slums.

Looking at CT-level outcomes, we find nil effects of the treatments on caretakers' behaviour, such as the percentage of time they report allocating to clean or supervising the cleaner and collecting fees; whether the caretaker reports the CT is cleaned at least twice per day; and, whether the caretaker reports adequate cleaning of the CT.

Contents

List of figures and tables

Acronyms

BL	Baseline
CT	Community Toilets
EL	Endline
FDGs	Focus Group Discussions
FU1	Follow-up 1
HH	Households
IFS	Institute for Fiscal Studies
ML	Midline
OD	Open defecation
pps	Percentage points
PGG	Public Good Game
SCA	Structured Community Activity
SS	Supply-side
SS+I	Supply-side + Information
ToC	Theory of Change
VCM	Voluntary Contribution Mechanism
WASH	Water, Sanitation and Hygiene
WTP	Willingness to pay

1. Introduction

While urbanization can bring benefits for economic, cultural and societal development, a rapid pace of urbanization can create enormous challenges. Cities in low and middle-income countries in particular are struggling to keep up with necessary infrastructure investment. Urban slums are a result of the unprecedented rate of urbanization that is leading 40 per cent of the world's urban population expansion to happen in areas with depressed private and public investment (Marx, Stoker, and Suri 2013). The water, sanitation, and hygiene infrastructure in particular is stressed beyond current capacity. Given that recent evidence suggests that inadequate sanitation is particularly threatening to early-life health when coupled with high population density (Hathi et al. 2017), efforts to improve sanitary conditions in slums are therefore of direct policy relevance.

It is generally accepted that public and community toilets (CTs) will, for the foreseeable future, continue to be an important solution to improve sanitary conditions in slums, given numerous constraints to increasing access to safely-managed private toilets.[1] The Joint Monitoring Program WHO-UNICEF has recently classified CTs in a higher step of the sanitation ladder than private on-site unimproved sanitation solutions (JMP WHO-UNICEF 2017). However, even where CTs are available, open defecation remains common behaviour among slum-dwellers. In India, CTs have been widely introduced in slums, but they are only used by 15% of slum-dwellers. Open defecation remains a common behaviour, practiced by more than half of Indian slum-dwellers without in-house sanitation facilities (Indian Ministry of Home Affairs 2011).

A likely explanation is that CTs are pay-to-use, which might exclude those with low ability to pay, including the poorest of the poor. Advocates of the "public health" approach argue that charging people for basic healthcare is unfair and only reaches the most affluent population (McNeil Jr. 2005; Benn 2006). Studies have shown that charging a small price for other preventative healthcare technologies such as insecticide-treated bed-nets dramatically decreases adoption (Cohen & Dupas, 2010). This is particularly concerning in cases such as toilet use, where negative externalities from non-adoption are severe.

However, charging a price for usage can affect usage through different channels, both positively and negatively. On the one hand, price is a potential carrier of information insofar as it is taken to signal quality or efficacy. For a newly-introduced product on the market, higher prices might signal higher quality and hence, higher prices could increase (or change) use through a higher perception of quality (Ashraf, Berry, and Shapiro 2010). In line with this, Dupas et al. (2016) found that, relative to free distribution of water disinfectant, a small non-monetary charge increases adoption. In addition, advocates of the "social marketing" approach argue that free goods are not valued and that pricing is a also a powerful tool to promote social beneficial causes and behaviour (Hastings 2007).

[1] A community toilet is typically located in or near a specific community area and used by a defined group of users or residents of that community area. Public toilets on the other hand are normally provided for the general public in places such as markets, train stations or other public areas, hence used by mostly undefined users. Shared toilets considered in this setting often combine these two aspects.

On the other hand, a potential effect of fees is to skew the composition of users towards individuals with a greater willingness to pay (WTP) for using the CTs (Roy 1951; Oster 1995). This is known as the “screening effect”. Empirical evidence on water and sanitation provides mix results. Two field experiments, one promoting latrine construction in rural Cambodia (BenYishay et al. 2016) and another promoting water filters in North Ghana (Berry, Fischer, and Guiteras 2012), found that adoption is not a function of WTP. Ashraf, Berry, and Shapiro (2010), however, found that adoption of water disinfectant increases with household’s WTP. Yet, screening may not be desirable in the case of a public good such as CTs, unless congestion is severe.[2]

To our knowledge, there is no study identifying the WTP for CTs and analysing its effect on usage (Evans et al. 2014). Overall, empirical evidence has shown that low usage rates have been accompanied by strikingly low willingness to pay (WTP) for in-house on-site sanitation solutions in areas where open defecation is prevalent (Van Minh et al. 2013; BenYishay et al. 2016) and for clean water in areas where water sources are faecally contaminated (Kremer et al. 2009; Ahuja, Kremer, and Peterson Zwane 2010; Jalan and Somanathan 2008; Null et al. 2012). In order to fully understand CTs’ use, it is important to investigate the drivers behind both low WTP and usage.

In this study, we focus on supply-side failures including poor quality of CTs, as well as potential lack of information on the health benefits compared to open defecation. Evidence suggests that poor sanitation infrastructure and quality of the facilities may alter valuation and propensity to use. A study in slums in Bhopal, India, found that 64% of CT users report dirt and smell as the most disliked infrastructure feature. This study further found a positive correlation between facility cleanliness and upkeep and facility usage (Biran et al. 2011). Similarly, a study in slums in Kumasi, Ghana, found that cleanliness is a key determinant for the use of CTs (Mazeau et al. 2014). In line with these observations, a systematic review of 24 studies that examined the association between structural and design characteristics of sanitation facilities and facility use reported that better maintenance, cleanliness, newer latrines and better hygiene access were all frequently associated with higher use, whereas poorer sanitation conditions were all associated with lower usage (Garn et al. 2016). For instance, a number of studies on the use of sanitation facilities in schools found that dirtiness, bad smell, the presence of faeces, absence of door and lack of lock significantly reduce usage (Garn et al. 2014; Caruso et al. 2014; Ness 2015).

It is widely accepted that information provision, for example on the health risks of faecal contamination, may improve perception and behaviour regarding sanitation, water and hygiene. Several studies found a positive impact of the provision of information on sanitation, water and hygiene outcomes. Particularly relevant for this paper is Gertler et al.’s (2015) randomized experiment, which found a positive impact of health promotion campaigns aimed to eliminate open defecation –popularly known as Community-Led Total Sanitation Campaigns (CLTS)— on the use of shared sanitation facilities in rural Indonesia and Tanzania. Likewise, there is evidence that CLTS increases WTP, construction and

[2] In an alternative mechanism, low WTP and low use can be both reflecting the underlying preferences associated with the use of CTs. In our setting, if a CT is present, it is assumed (and supported by the data) that individuals gain utility from its use. For this reason, we select for the study only households in high proximity to a CT.

adoption of private latrines and hygienic behaviours (Pickering et al. 2015; Van Minh et al. 2013; Jalan and Somanathan 2008; Luoto, Levine, and Albert 2011; Spears 2014; Cameron, Shah, and Olivia 2013), even sustainable over time (Cairncross et al. 2005; Wilson and Chandler 1993). Notably, Guiteras, Levinsohn, and Mobarak (2015) found evidence that when poverty and supply failures are binding constraints, information has an impact when coupled with subsidies and supply-side interventions.

The report is organized as follows. Section 2 presents the intervention in detail, while section 3 discusses the evaluation methods and design. Section 4 presents the main findings of the study, while section 5 discusses the cost of interventions. Section 6 presents a discussion of the findings and the lessons learned from the study, while section 7 concludes and provides policy recommendations.

2. Intervention

2.1 Description

The interventions have been implemented in the cities of Lucknow and Kanpur, in the state of Uttar Pradesh, in northern India. Lucknow is the capital city of state, the eleventh-most populous city and the twelfth-most populous urban agglomeration of India, with a total population of 2,817,105 inhabitants (Census of India, 2011). Similarly, Kanpur is the 12th most populous city in India, with a total population of 2,767,348 inhabitants (Census of India, 2011).

These two cities are comparable to other growing cities in India and Southeast Asia. In terms of share of population living in slums, Lucknow reaches 13% of its population, while in Kanpur 15% of the population lives in slums (Census of India, 2011). This is relatively lower to other larger cities in India, such as Mumbai (41%), Kolkata (30%) and Chennai (29%), and to other major cities in Asia, such as Dhaka (40%), Colombo (40-50%) or Hyderabad in Pakistan (33%).

India shares many key characteristics with other countries that faces similar challenges of urbanization and low sanitation access. In 2014 for example, GDP per capita in India (1595 USD) is close to that of Sub-Saharan Africa (1773 USD) and Southeast Asia (1515 USD). More specifically, Uttar Pradesh has a comparable poverty rate to the rest of India (28% versus 22%). In terms of population growth there are similarities: in 2014, population grew by 1.2% in India, compared to 1.4% in Southeast Asia, and 2.7% in Sub-Saharan Africa. Life expectancy presents also many similarities among the selected areas and broader regions. Life expectancy in India (66.5 years) is in line with Southeast Asia (66.9), while Ghana (61.1) and Kenya (61.7) are slightly better off than the average in Sub-Saharan Africa (56.8).

We summarize in the next section the main components of the intervention.[3]

[3] During the interventions implemented by the research team, activities of the Swachh Bharat Mission (SBM) were simultaneously implemented. This is a nation-wide campaign in India with the objective to eliminate open defecation through the construction of household-owned and community-owned toilets and establishing an accountable mechanism of monitoring toilet use. The research interventions are randomized, while SBM activities were not randomized, which guarantee homogeneous coverage of these activities by treatment arm. We are not aware of any other WASH interventions ongoing at the time of the intervention.

2.1.1 Supply-side intervention

The aim of the supply-side intervention is to improve the quality of community toilets by releasing resource constraints and incentivising the performance of the CT management. In order to understand how the supply-side intervention was implemented, it is important to explain how the CT management works. There are two main types of management modalities. In one management modality, the CT is managed by the municipality, who hires private contractors to operate a set of community toilets. Caretakers are hired by these private contractors to collect the fees and clean the CT in exchange of a fixed salary and accommodation (they live inside the CT compound). In the other management modality, the municipality concedes the operation and maintenance of a set of CTs to a social service organization through a public-private partnership lasting multiple decades. These social service organizations have a city manager and several mid-line managers that supervise the performance of CTs in a specific fraction of the city. CTs are operated by caretakers, who are in charge of collecting the fee from users and supervise the cleanliness of the community toilet. The management hires cleaners and provides the necessary tools and products to clean each CT. Notably, the greater the revenue generated by a CT, the greater the resources allocated to it. Caretakers have a monthly fee goal and keep the remaining collected amount as part of their salary. On average, caretakers in both modalities earn Rs. 6,000 per month.

We implemented the supply-side intervention at the caretaker level because of three main reasons. First, caretakers are able to improve the cleanliness of CTs by exerting more effort when cleaning or supervising the cleaner. Second, caretakers can stop users if they refuse to pay for the fee to use the CT. Third, caretakers work only in one CT, contrary to mid-level managers that supervise more than one CT in the study. For sample size and contamination concerns, we do not treat mid-line managers, though we engage the management at the highest level in order to make the implementations of the supply-side intervention feasible. To prevent control caretakers from learning about the intervention, we provide a list to the city managers of all CTs in the study without revealing which CT is allocated to which treatment arm. City managers are informed about the intervention, but are encouraged to not discuss with contractors, mid-line managers and caretakers the different interventions in order to prevent complaints.

We executed different components of the supply-side intervention in a phased-in design in order to disentangle differential effects. We started by implementing a Grant Scheme, followed by the announcement and payment of a Financial Reward Scheme.

CT Grant Scheme

The supply-side intervention started with the Grant Scheme aimed at improving the quality of the CTs. We hypothesize that resource constraints prevent caretakers from improving cleanliness and maintenance of CTs, and hence, an initial push could allow them to take the CT quality to a higher level. The idea is to break a vicious cycle in which low quality of CTs generates low WTP and usage and this in turn generates low revenue and resources to operate and maintain the CT.

We conducted two visits right after the baseline survey. In the first visit, the CT caretaker(s) had the opportunity to choose a type of aid according to their needs and expectations of increasing user traffic (hence more resources for the CT). The grant scheme offered three standardized (i.e. same monetary value) packages from which the caretaker could select one: (i) deep cleaning (i.e. septic tank sewage removal, unclogging latrines and sewerage pipes and cleaning walls, floors and inside toilets), (ii) sanitation/water connection repairs and/or infrastructure refurbishment or (iii) cleaning tools and agents and cleaning training. For CTs that selected 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 for the works to be conducted was set. Based on this information, our partner FINISH arranged the work with an external contractor. To keep consistency in the works, we used the same contractor in all CTs.

During the second visit, the contractor developed the work under the supervision of FINISH. Pictures of the improved areas were taken after the work is done. Figure 1 shows an example before and after a deep cleaning of a CT in Lucknow. For CTs that selected training and cleaning tools, the second visit consisted of two sessions: theory and practice. In the theory session, FINISH explained why the CT should be frequently cleaned, which frequency, at which hours, using which tools and products and how the cleaning should be done. For this, FINISH used flipcharts with cartoons and relevant messages. In the practical session, the caretakers and cleaners had to accomplish cleaning tasks under the supervision of a trainer that provided feedback. In addition, FINISH endowed the CTs with several tools and cleaning products, including: 4 hand-gloves, 5 floor cleaners, 4 toilet disinfectants, 5 liquid soap, 4 toilet cleaning brush, 2 wipes, 4 nose masks, 2 brooms, 2 buckets and mug set, 3 surf-ghadi, 2 hand-washing dispensers, 2 dustpans and 2 dustbins.

Figure 1: Example of before and after a deep cleaning of a CT in Lucknow

Source: Picture taken by the team as part of the tracking application.

CT Financial Reward Scheme

We introduced a Financial Reward Scheme in order to promote a sustained maintenance and cleanliness of CTs. Two months after baseline and after the completion of the Grant Scheme works, we announced the Financial Reward Scheme aimed at incentivizing caretakers to maintain the CT clean. Caretakers could receive a financial reward conditional on achieving the following:

- availability of soap in hand-washing facilities for both genders;
- cubicles free of visible faeces (inside and outside the latrines);
- bacteria count in defecation cubicles kept to a minimum standard (computed as below the median of the demeaned baseline distribution by city).

Achieving (i) and (ii) paid Rs. 500 and achieving (iii) paid Rs. 1000, so the maximum amount a caretaker could receive was Rs. 2000, representing a third of a monthly salary. We provided a greater pay-off for reducing bacteria because this is a harder task to comply. Caretakers were informed that an external agent was going to return to measure each condition on a random day and time within the following two months. We informed caretakers that we would pay the financial reward depending on what the external agent measured. In CTs with more than one caretaker, the financial reward was split among them.

At this point, we also informed CT caretakers of their baseline cleanliness performance to help them estimate the additional effort they had to exert in order to achieve the conditions. The literature highlights the fact that individuals may go through a variety of seemingly irrational behaviours when informed about their past performance in order to cope with failure –from handicapping their own future performance to practicing self-deception through selective memory or awareness (Benabou and Tirole 2002). Yet, this is more prevalent when external rewards are based on rankings (Ashraf, Bandiera, and Lee 2014). Furthermore, if financial rewards are based on improvements from past performance, individuals may behave strategically (Muralidharan and Sundararaman 2011), performing badly in initial rounds in order to achieve greater pay-offs in the future. Therefore, our financial reward scheme is linked to individual performance, not relative to others' performance, and works for each round independently.

After two months and every two months from then on, we measured each of the conditions and paid accordingly. In each round, we announced that a new round will take place and we reminded the caretaker(s) of the conditions of the Financial Reward Scheme. In total, we paid four rounds of financial rewards.

The incentive scheme may create some multi-tasking problems[4] (Neal 2011) and may crowd out intrinsic motivation (Besley and Ghatak 2005; Benabou and Tirole 2003). To measure the extent of a multi-tasking problem, we collect detailed data on time allocation and cleaning practices. In addition, we measure CT cleanliness relying on observations as well as a precise measure of bacteria counts in randomly selected areas of the toilet compound and cubicles. To measure the extent of a crowding-out in intrinsic motivation, we measured pro-social preferences through several rounds of a dictator game in which caretakers can donate to a sanitation project.

2.1.2 Information campaign

For selected CTs, the supply side intervention is complemented by an information campaign. We hypothesize that information and supply-side nudges can interact in a positive way to increase the WTP and use of CTs and the demand for CT cleanliness. The primary aim of our information campaign is to create awareness of health risks of open defecation to their family (private) and the neighbours of their community (public). We provided information on the pathways (i.e. 4 Fs –fingers, flies, fluids and fields)

[4] Assuming that agents engage in two types of tasks, one that follows best practices and another aimed to increase the imperfect performance observed by the principal, the introduction of incentives may create distortions in the allocation of effort. Caretakers may, for example, clean the same number of hours, but throw water everywhere to make the CT look clean, instead of using disinfectant.

through which open defecation can cause infectious diseases, stunting and even death in family members and neighbours. Furthermore, we provided information on the relevance of ensuring females' safety and the importance of male household members to accompany females when using CTs, particularly at night.[5] Moreover, we provided information on the benefits of using CTs and the importance of guaranteeing that everybody pays the CT fee to achieve a sustainable operation and maintenance. In addition, we emphasized the rights of users when paying fees and encourage users to demand the CT management clean and well-maintained toilets.

We targeted all household members, especially household heads and spouse. We paid particular attention in providing information that can be processed by participants with low levels of literacy. We provided information through four means:

(i) Door-to-door information campaign using a flipchart with cartoons and messages targeted to all household members, especially household heads and spouse. This information session covered the following sections: how open defecation affects your community; how open defecation affects your family; benefits of using community toilets; what can you and your family do to make the CT better; your rights when you pay the fee for using the CT. Figure 2 captured the moment in which one of our enumerators provided the door-to-door information campaign to a family in Lucknow;

(ii) Circulation of leaflets with a summary of the flipchart;

(iii) 5 posters (3 medium size and 2 large) placed in each CT using catchy phrases and Bollywood memorable scenes, highlighting messages provided during the door-to-door campaign;

(iv) Monthly reminders in the form of voice messages sent to mobile phones.

Figure 2: Example of Information campaign provided to a family in Lucknow

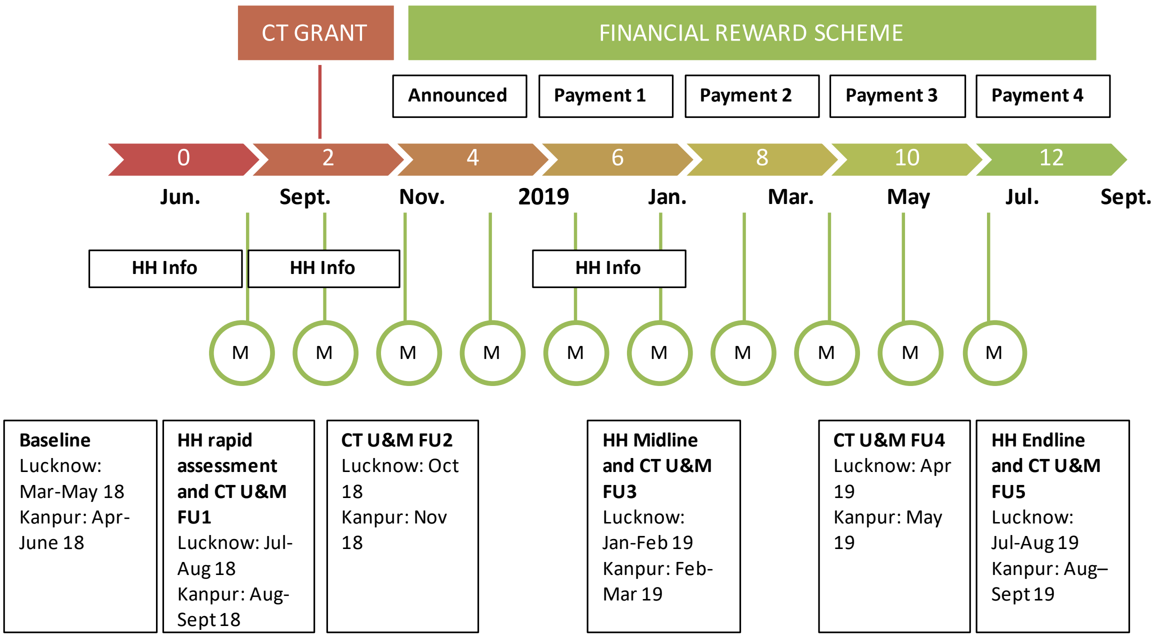
Source: picture taken by the team as part of the supervision of pilot activities.

---

[5] Qualitative evidence conducted by the research team before the baseline highlighted the constraints faced by household members in relation to sanitation practices. There were very gender specific. In-house/dwelling (makeshift) toilets are used by female members, thereby reducing their need to leave the house when it is dark. Clearly, women might still decide to move for other reasons.

Figure 3 shows the timeline of our intervention. In the upper side, we present the timeline of the supply-side intervention.

Figure 3:Timeline of intervention and survey rounds



The CT Grant Scheme took place during the first two months of the study – between June and August 2018. Two months after, we announced the Financial Reward Scheme. From month 6 and every two months until the end of the study, we paid the financial reward to those CTs that achieve the conditionality. The last payment of the financial reward (payment 4) was done after the end-line survey, at the end of the study – August and September 2019. The lower side of the figure shows the timeline of the information campaign. In months 0, 2 and 6 we conducted the door-to-door campaign, using flipcharts and leaving summary leaflets to household members. Between month 1 and 11 of the study, we sent monthly voice messages –a total of 10 rounds of voice message. Figure 3 also shows the timeline of our data collection (please refer to the boxes at the bottom). Section 3.4.2 explains in detail the timeline of the collection of primary data.

2.2 Theory of change

This study is based on the following three status quo considerations, which represents the assumptions behind the presented theory of change:

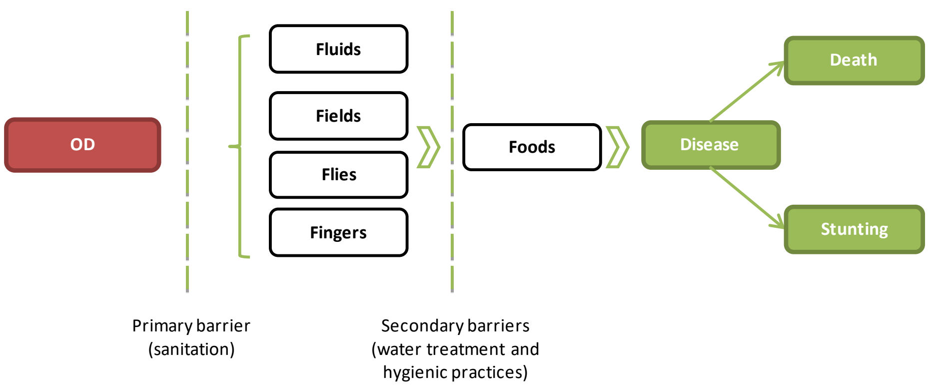
(i) Open defecation has significant health implications. Health benefits from improved sanitation and the eradication of OD are estimated to be substantial (Prüss et al., 2002). Epidemiological research has found that chronic environmental exposure to faecal germs is one of the main causes of the prevalence of infectious diseases in low- and middle-income countries (LMICs). The pathogens that cause these diseases are transmitted through several pathways summarized in Wagner and Lanoix's F-diagram (1958) presented in Figure 4. Pathogens are transmitted from faeces to the host via fluids (drinking water), fields (soil), flies (anthropods) and fingers (hands), who enter orally directly or through the ingestion of food. Primary barriers, such as safe disposal of faeces through adequate sanitation facilities, and secondary barriers, such as drinking water and adequate hygienic practices, allows the isolation of feaces so that infectious agents cannot possibly get to a new host. The F-diagram predicts that if the feacal pathogens do not infect new hosts, child death and stunting in survivor children can be prevented.

Drinking water is one of the main pathways though which open defecation affects the health of slum dwellers in India. Two main facts support this argument. Firstly, water contamination at point-of-use is significantly higher than at source, implying that a large degree of contamination happens after water is taken from the tap. In fact, while about 70% of Indian slum dwellers have access to tap water (Census 2011), only 46% have access to tap water within the premises they live in. 28% have to walk 100 meters or more to collect water from a tapped source. Secondly, as a result of the poorly maintained water infrastructure and supply shortfalls, water pipes with cracks suck in faecal matter when resuming water supply. In fact, no Indian city has continuous piped water. A Service Level Benchmarking Program, initiated by the Ministry of Urban Development (MoUD) in 2006, found that the average duration of water supply in 28 Indian cities was 3.3 hours per day.

(ii) The slum-dwellers' WTP for CT use and demand for cleanliness is low, and this explains low rates of CT use among potential users, leading to open defecation as the only alternative for many slum dwellers. The government of India, along with NGO's and the private sector, has expanded the coverage of CTs during the last decade. However, even where CTs are available at a relatively close distance, open defecation remains common behaviour among slum-dwellers. This phenomenon is likely explained by a low willingness-to-pay for CTs that are pay-to-use. The fee –standardized at 5 rupees per usage— may be screening out those with low willingness to pay and those with low ability to pay, including the poorest of the poor. This is known as the "screening effect" (Ashraf, Berry, and Shapiro 2010).

(iii) Public, non-profit and private organizations working on sanitation access are essential actors in creating and implementing solutions that pursue SDG #6: Ensure availability and sustainable management of water and sanitation for all. Therein, we assume that appropriate guidance on what constraints are important to address sanitation challenges and which different intervention approaches are effective, can increase short and long-term wellbeing of their beneficiaries.

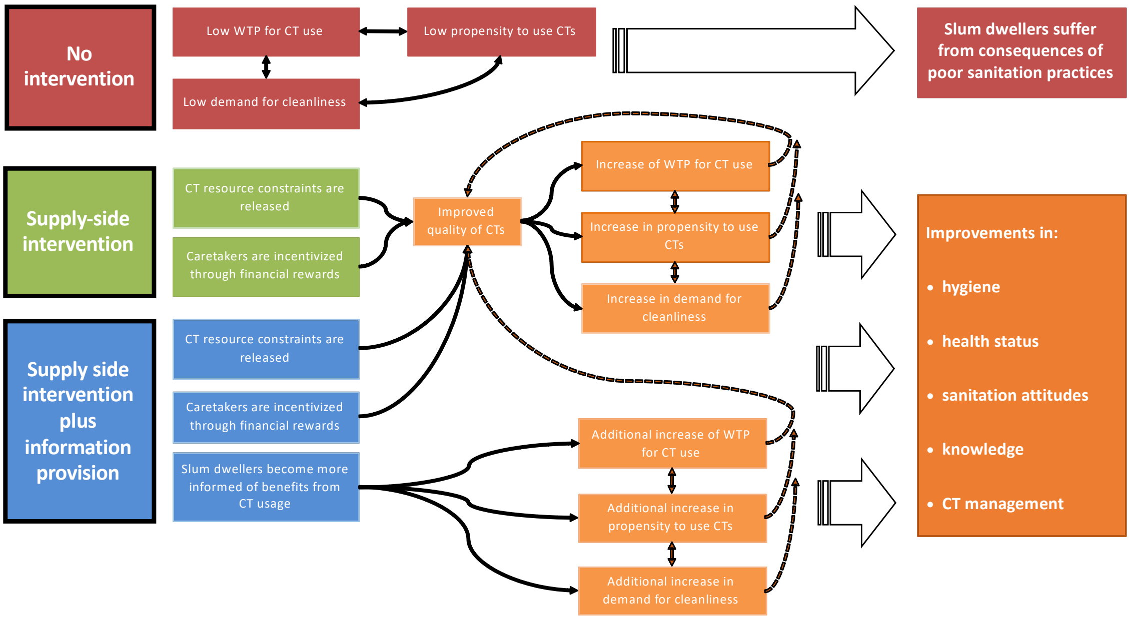
Figure 4: How open defecation affects health



Source: adaptation from Wagner & Lanoix (1958).

The theory of change is summarized in Figure 5. We have identified interventions that seek to improve sanitation behaviour in urban slum settlements, by improving the supply side – more specifically by improving the quality of existing CTs – accompanied by providing information about the risks associated with open defecation and the benefits and rights when paying for using CTs.

Figure 5: Theory of Change



The assumption that we can affect usage behaviour and related WTP through (i) the improvement of the quality of community toilets, and (ii) the provision of information, is based on the following rationale:

(i) ***The improvement of the quality of community toilets***. A study by the Water and Sanitation Program of the World Bank (2016) identifies the lack of cleanliness and poor up-keep and lack of proper disposal facilities as the main constraints to use. Although an observed deficiency may not be a key deterrent for use, other studies identify cleanliness as an important determinant of usage. Biran et al. (2011) and a related project report by Mahon (2010) for example analyse patterns and determinants of communal latrine usage in urban poverty pockets in Bhopal, India. They find that by far most disliked feature of CTs was dirt and smell, as mentioned by 64% of users. The second most staged reason was queuing (19%), followed by lack of water (11%). In line, they find a positive association of the proxy variable for CT cleanliness/upkeep and usage. This generates the following three hypotheses:

a. *Releasing resource constraints improve the quality of CTs.* In general, the quality of CTs is precarious and this is due to a vicious cycle. Most CTs face resource constraints to generate substantial improvements because usage and related revenue are not high enough. In addition, the centrally managed system that redistributes all resources suffers from inefficiencies. Caretakers are not endowed on time with the necessary tools and agents or aid for repairs, even after repeated requests to the management. Releasing resources constraints to improve cleanliness and maintenance should result in a visible improvement in quality.

b. *Providing financial rewards to caretakers improve the quality of CTs.* Sustaining the improvement obtained by releasing resource constraints depends on the effort exerted by caretakers to operate and maintain the CT clean. Therefore, we further postulate that extrinsic incentives can play an important role in determining the investment in the quality of the CT by caretakers. There is evidence that extrinsic rewards may affect positively performance on pro-socially motivated tasks, as soon as they are announced and specifically when incentive schemes are linked to own performance (Ashraf, Bandiera, and Jack 2014; Ashraf, Bandiera, and Lee 2014). We thus expect financial rewards to increase effort and this to be translated into noticeable improvements in cleanliness.
c. *Supply-side improvements increase WTP for CT use, the propensity to use CTs and the demand for cleanliness in the short term.* Improvements in the quality of the CT can increase the slum-dwellers' value of the CT and the demand for such standard. Yet, beyond the conscious channel, quality improvements can serve as nudges that positively affect sanitation practices. There is evidence that people make biased assessments of values, which are affected by the context in which they make decisions. Changes to sanitation habits such as usage of community toilets can be automatically triggered by context cues and small nudges, including changes in physical setting (Wood, Tam, and Witt 2005). Therefore, we expect supply-side improvements such as cleanliness and better infrastructure maintenance to increase the WTP for and usage of CTs and the demand for CT cleanliness. Nevertheless, evidence also suggests that supply-side improvements could also lead to unexpected results, especially in presence of behavioural responses in their target population. In fact, public investments can discourage (or crowds-out) private investments, therefore leading to negative impacts of public investments (see, for instance, Peltzman, 1973, and Cutler and Gruber, 1996). For developing countries, evidence of this mechanism has been observed for investments in education (Das et al., 2011), in sanitary behaviour (Bennett, 2012), and anti-malaria behaviour (Armand et al., 2017).

(ii) ***The provision of information***. The assumption that WTP and usage can be moved through the provision of information is based on several studies related to water, hygiene and sanitation. For one, there is robust evidence that information provision about the quality of water significantly impacts behaviour in terms of purifying water (Jalan and Somanathan, 2008; Madajewicz et al., 2007; Luoto; 2009). Second, an experiment in rural India shows that deliberation about the usefulness and price of soap increases price sensitivity significantly (Spears, 2014). And third, more specific to sanitation, are studies that demonstrate that knowledge about the link between poor sanitation and disease is a significant predictor of willingness-to-pay for household-level toilet facilities (Minh et al, 2013). We bring forward the following hypothesis. Combining supply-side improvements with the provision of information increases WTP for CT use, the propensity to use CTs and the demand for cleanliness in the short term by more than supply side improvements on their own. Based on evidence that information provision can shift WTP for health investments such as soap and water (Jalan and Somanathan, 2008; Madajewicz et al., 2007; Luoto; 2009; Spears, 2014) and adoption of shared sanitation facilities (Gertler et al. 2015), we postulate that information and supply-side nudges can interact in a positive way to increase the WTP and use of CTs and the demand for CT cleanliness.

Our metric for defining these interventions as successful considers short and long-run impacts separately. In the *short-run*, we classify the interventions as effective is we observe increases/improvements in our primary and secondary outcomes. In particular, we define the following *primary outcomes:* WTP for CT usage; CT usage (i.e. improved sanitation behaviour, accompanied by a decrease in open defecation); demand for cleanliness; and CT quality. Our *secondary outcomes are*: hygiene; health status; sanitation attitudes, expectations and knowledge; and CT management. Detailed information on the primary and secondary outcomes is provided in the Results section within 4.2.2. Research analysis.

Supply-side improvements alone and with the provision of information can result in *long-run* effects on WTP for CT use, the propensity to use CTs and the demand for cleanliness. In order to reach long-run effects, we assume that short-run effects can be sustained, i.e. that participants will not revert to the status quo WTP and valuation of improved sanitation in the short-term, and second, we assume that intervention participants will modify not only their own behaviour, but will also encourage their household, neighbours and kin to adapt. This latter point is particularly important when considering the role of parents in determining their children's exposure to risk.

A possible negative externality of such a success would be congestion, and possible difficulties in keeping improved standards due to higher usage. This aspect was monitored throughout the study, so that implementing partners have relevant information to act upon. A potential unsuccessful outcome would include a decrease in investment or protective practices as a result of information that ambiguously portrays risk and potential health costs, or perhaps decreases expected returns from current investments.

2.3 Monitoring plan

We conducted efforts on the field to monitor two components of the project: (i) the implementation of the intervention; and (ii) the data collection.

In order to monitor the implementation of the intervention, a series of mobile phone apps were designed purposely for the implementation and supervision of the intervention. The apps were designed in conjunction with the research team, Morsel and FINISH. These tracking apps allow us to identify treatment fidelity.

The CT tracking app is used by FINISH and captures whether the intervention was provided, the date, details of the participants that were exposed to the intervention and the contact details of the respondent signing the consent form. Furthermore, we captured extra details from the caretaker (e.g. choice, availability, satisfaction) and pictures from the community toilet to visualize outputs of the intervention. For example, we cross-checked with pictures collected in the tracking app if the deep cleaning was actually conducted. Before and after pictures of the same area cleaned allowed us to monitor the work. Furthermore, we include a confirmation form that needs to be signed by the caretaker indicating if the CT received the intervention. In addition, we took pictures of the invoice signed by contractors and a representative of FINISH for each work done at a CT. Finally, the app captures if the CT has each of the 5 posters, in which condition they are and if replacement is needed.

The implementation of the information campaign is captured in a separate app used by Morsel. The app is pre-populated to indicate if a household should receive the information campaign and records whether the flip-chart was explained and the attitude of the participants while the information was provided. Furthermore, it captures whether the leaflet was provided to the household.

Finally, the implementation of the voice message was recorded in an especially designed tracking app. This tracking app is pre-populated with all mobile phones and message categories to be sent in each round. The tracking app records whether the voice call was answered and the minutes it was answered before the call finished.

For the monitoring of the data collection on the field, we had 4 layers of supervision: (i) 3 field supervisors of enumerators; (ii) 1 main supervisor of the field supervisors; (iii) 1 project manager; and (iv) the London base team. After training was conducted, the data quality was checked regularly from the piloting until the final collection By the project manager based in Lucknow and the research team based in London. Identified errors were rectified on the spot in the data collection application. If necessary, enumerators went back to re-collect data or collect data missing by mistake.

3. Evaluation

3.1 Primary and secondary questions

The *primary research* questions are:

- What is slum-dwellers' WTP for CT usage? And does WTP affect CT usage (evidence of a screening effect)?
- Can the releasing of resource constraints and provision of financial incentives to caretakers improve CT quality?
- Can improvements in CT quality increase slum-dwellers' WTP, usage and demand for cleanliness?
- Is the effect of CT quality improvements on WTP, usage and demand for cleanliness greater when also providing slum-dwellers with information about health risks of open defecation, as well as benefits of CTs and users' duties and rights when paying the CT fee?
- Do the interventions have longer-run effects on WTP, usage and demand for cleanliness?

The *secondary research* questions are:

- Can the release of resource constraints and provision of financial incentives to caretakers improve the management of CTs?
- Can improvements in CT quality improve attitudes towards adequate sanitation practices as well as knowledge and expectations of health benefits? Are these effects greater when providing information about health risks of open defecation, as well as benefits of CTs and users' duties and rights when paying the CT fee?
- Can improvements in CT quality improve the hygiene practices and health status of slum dwellers? Are these effects greater when providing information about health risks of open defecation, as well as benefits of CTs and users' duties and rights when paying the CT fee?

3.2 Design and methods

We address primary and secondary research questions in the context of Indian slums in the cities of Lucknow and Kanpur (Uttar Pradesh, India) using a randomized-controlled trial (RCT) design.[6] Figure 6 shows the study area.

Figure 6: Study location

Note. Lucknow and Kanpur are the two selected cities for the experiment; both are based in the Indian State of Uttar Pradesh. Base-map source: ESRI.

We randomly selected 110 catchment areas of CTs in both cities and we allocated them to one of three experimental arms. As discussed in Section 3.4.1, we have explicitly removed CTs that could also be interpreted as public toilets, which led to a total of 110 CTs in the population used for randomization. Randomization into the treatment arms was carried out at the cluster level, specifically the *CT catchment area level*. We provide more detail on the definition in the next section. Randomizing at the cluster level has the advantage to limit contamination of the control group, especially considering the possible spread of information. To allocate clusters to treatment arms, we stratified the sampled clusters by the main organization managing CTs in our study area (versus other organizations) and by city of study (Lucknow and Kanpur). We then build blocks of 3 CTs using m-distance (Mahalanobis) relative proximity. To construct m-distances, we used the rich census information we collected, including CT and slum-dweller characteristics. After forming blocks of similar clusters (CTs), we randomly allocated each CT in a block to either treatment 1, treatment 2 or the control group. Each one of the three possibilities has the same probability. The statistical software Stata, and specifically the random number generator, is used to apply this procedure.

---

[6] To identify slums, we make use of recognized slums according to the Census of India.

Catchment areas were allocated to one of these three groups:

A. **Control**. Participants in this group did not receive supply-side nor information interventions.
B. **Treatment 1 (Supply-side intervention).** In this treatment group, cleanliness and maintenance of CTs was promoted by introducing a *Grant Scheme* offering standardized (i.e. same monetary value) packages including either (i) deep cleaning (i.e. septic tank sewage removal, unclogging latrines and sewerage pipes and cleaning walls, floors and inside toilets), (ii) sanitation/water connection repairs and/or infrastructure refurbishment or (iii) cleaning tools and agents and cleaning training, from which CT caretakers can choose according to their needs and their expectations to increase usage. In addition, two months after baseline, we announced a *Financial Reward Scheme* aimed at incentivizing caretakers to keep the CT clean. At this point, CT caretakers were informed of their baseline cleanliness performance. After two months (and every two months from then on), caretakers received a financial reward conditional on achieving the following: (i) availability of soap in hand-washing facilities; (ii) latrines free of visible faeces; and (iii) bacteria count in defecation cubicles kept to a minimum standard (to be computed from the baseline distribution).
C. **Treatment 2 (Supply side intervention plus information provision).** In this treatment arm, the supply side intervention (as described under point B) was complemented with a household-level information provision component. Participants in this group received information about private (family) and public (community) health risks of open defecation. Furthermore, we provided information on the relevance of ensuring females' safety and encouraged male household members to accompany females when using CTs, particularly at night. In addition, we provided information on the benefits of using CTs and the importance of ensuring that everybody pays the CT fee. We put emphasis on providing information on the rights of users when paying fees, including clean and well-maintained toilets, as well as promoting users to demand maintenance and cleanliness of CTs. This component targeted all household members, especially household heads and spouse, and was designed such that information could be processed by participants with low literacy. We provided information in four forms. Firstly, a door-to-door information campaign using a flipchart with cartoons and messages. Secondly, by leaving a leaflet with a summary of the flipchart to households visited during the surveys. Thirdly, by placing posters in the CT, highlighting messages provided during the door-to-door campaign. Fourthly, by sending monthly reminders in the form of voice messages to all mobile phones collected in the household. To disentangle the effects of receiving voice messages, all treatment arms received voice messages without new information (specifically stating that the community toilet is open from early morning until late evening). Households allocated to the supply-side treatment additionally received voice messages informing that the community toilet has been granted aid to improve its service.

3.3 Ethics

The study has obtained ethics approval from the University College London Research Ethics Committee (REC) on March 23rd 2017 until December 31st 2020 (ethics application code is 2168/012). Changes to the design and data collection were approved in May 2018. The trial and the pre-analysis plan have also been registered with the AEA RCT Registry. The registry number is AEARCTR-0003087.

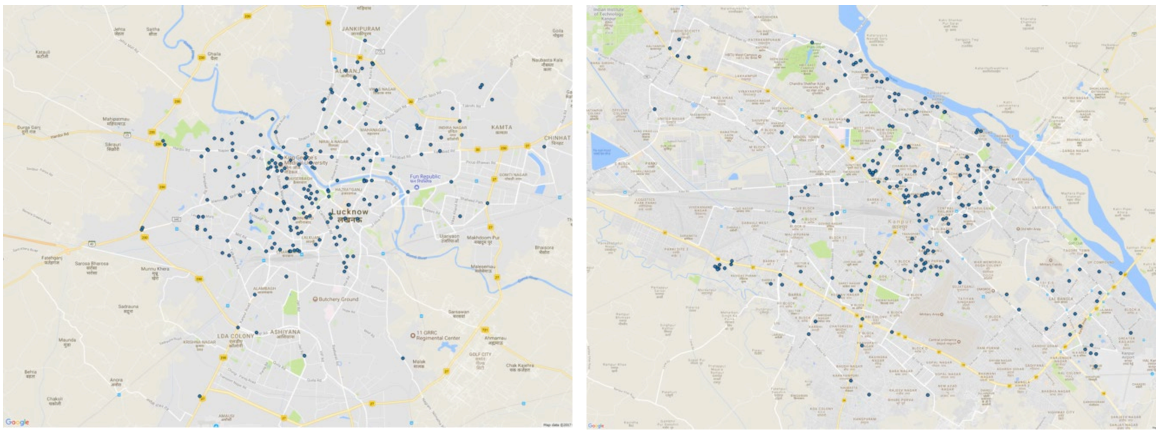
3.4 Sampling and data collection

3.4.1 Sampling

This study performs a two-level randomization design and therefore, we are interested in two different level of the sampling frame. The **first level** of the sampling frame is defined by all community toilets (CTs) in the cities of Lucknow and Kanpur. In order to obtain this sampling frame, we performed a CT census in both study cities. This census collected GPS coordinates and information on ownership, management, location, problems (i.e. dirty, stinky, observed faeces and flies, unsafe, non-functional), service (i.e. opening times and lightning availability), infrastructure characteristics (i.e. year of construction, number of toilets by gender, doors and locks functionality, hand-washing and showering facilities, quality of walls and floors, type of sanitation facility) maintenance activities (number of staff, cleaning frequency, cleaning tools), users (type and main users), catchment area (distance users come from) and fees (methods and fees). We identified at total of 409 CT. Figure 7 presents the distribution of CTs in Lucknow (left map) and Kanpur (right map).

Figure 7: CT Sampling Frame

Lucknow **Kanpur**



Note: The dots denote the geo-location of all community toilets identified in Lucknow and Kanpur through our Community Toilet Census. Base-map source: ESRI.

The **second-level sampling frame** is characterized by all households in the catchment areas of selected community toilets (which represents a cluster). Out of the 409 CTs identified in the first stage, we chose a subset of CTs to become part of the study, based on the following criteria: the CT has to be pay-to-use; the CT has to be located close to a residential area (slum) and used by residents. We drop CTs for which the distance to another CT is below a certain threshold. In particular, there should be sufficient distance between two CTs to avoid users switching between CTs (possibly driven by their treatment status). 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 drop CTs in whose catchment areas fewer than eight eligible households are living. A household is considered eligible if the following conditions are respected: the household lives in the catchment area of a selected CT, which is broadly defined as slum area within 250 meters from the CT building. Households are linked to CTs based on geo-coordinates collected during the census;[7] at least one household member uses a CT or shared toilets (i.e. neighbours, makeshift, work, school), or practices open defecation; the household must have reported during the census interview not to intend to migrate during the following 18 months (i.e. until the planned study endline survey). This selection criteria result in 110 CTs selected for the study. Table 1 shows the distribution of CT catchment areas by treatment arm and city.

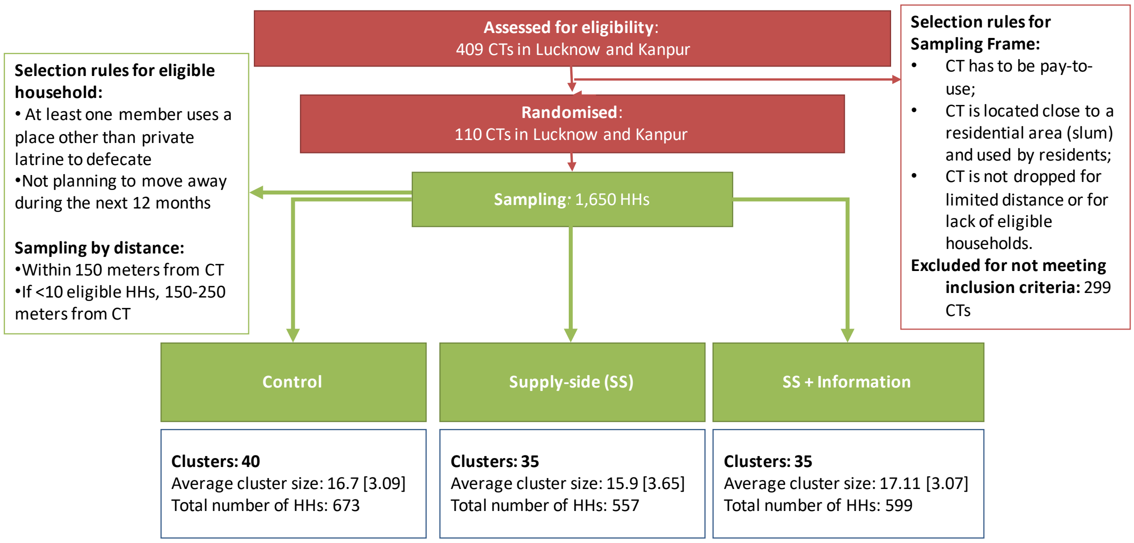
Table 1: Number of CT catchment areas per treatment arm and city

	Control	Supply-side (SS)	SS + Information	Total
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%)

Figure 8 shows the flowchart for the CT and household sampling procedure. Within each of the 110 CTs and their catchment areas, we sampled up to 17 eligible households. Given that distance is a major determinant of CT usage, we focused on eligible households living closer to the CT (within 150 meters). Since some CTs have more dispersed populations, we conducted a two-step sampling procedure. First, in large-population catchment areas (where 10 or more eligible households are available within 150 meters), we sampled only from eligible households that are located within this bound. Second, in small-population CTs (where less than 10 eligible households are available within 150 meters), we first sampled all eligible households within 150 meters and then randomly selected the remaining households from those that are located between 150 and 250 meters from the CT. In total, we sampled for interview 1,650 households in 110 randomization units (catchment areas of CTs) and one CT caretaker per randomization unit.

[7] A small percentage of households with inaccurate geo-coordinates (missing geocodes or distance to closest community toilet greater than the 400 meters buffer area set for household census) were linked to the community toilet located in the slum (0.7 percent in Lucknow and 1.6 percent in Kanpur). In these cases, we imputed the distance to the community toilet using the median of eligible households in the slum.

Figure 8: Flowchart for the sampling procedure



Note: The flowchart presents the procedure followed for the sampling of community toilets and its catchment area.

3.4.2 Data collection

We distinguish between quantitative instruments, behavioural measurements and laboratory tests. shows the timeline of the study, including the data rounds.

Quantitative instruments

The quantitative data sources used in the study are the following:

1. **CT census.** The purpose is to obtain a mapping of CTs in Lucknow and Kanpur, together with the distribution of slums close to these CTs; understand how many there are, and how they are distributed, who uses the CTs (from the community, passers-by, workers, etc.), their characteristics, management structure, payment, among others. The scope is of 409 CTs identified in Lucknow (201) and Kanpur (208). The instrument was administered to CT caretakers and/or supervisors. The census includes collection of GPS coordinates. Timing was March-July 2017.
2. **Household census.** The main purpose is to identify our sample of eligible household respondents. The household census covered households located within slum borders and 400 meters from each of the 144 pre-selected CTs (i.e. pay-to-use, used by residents and not close to another CT within 200 meters). The questionnaire collects household demographics (number of members and gender primarily) and dwelling characteristics as well as sanitation practices (focusing on toilet ownership, CT usage and open defecation practice). The household census also collects GPS coordinates. Timing was August 2017-March 2018.
3. **Baseline (BL) survey.** This is a standard baseline survey, collecting information on socio-demographics as well as dwelling characteristics, assets, income and expenditure. In addition, we collected baseline information on health status and sanitation and hygienic behaviour and priors, attitudes and health expectations linked to open defecation and CTs. Sanitation behaviour was measured by

asking where each demographic group went to defecate during the last 2 times.[8] 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."* We also asked about the sanitation behaviour of the most intimate neighbour in order to identified the extent of response bias when asked to disclose the behaviour of their own household (Yeatman and Trinitapoli 2011). We further collected information on child health for children under the age of six, as well child hygienic practices and childcare knowledge. To understand intra-household dynamics across spouses, we collected data on household decision-making. Finally, the interviewer conducted observations of the household environment and the respondent. The instrument, including all modules, has an average duration of one hour. The timing was March-June 2018.

The main respondent for the household survey is an adult household member – in most cases the household head, sometimes the spouse or other knowledgeable household member, always falling in the age range of 18-64 years. We also interviewed the spouse of the household head, aiming to collect information on decision-making dynamics from one male and one female household member. In households with children aged 5 years or younger, we further interviewed the primary caregiver to collect information on child health and sanitation practices.

The scope of the baseline survey was of 110 catchment areas and a total 1,650 sampled households (16 households per catchment area, on average). The non-response rate was 20% and the percentage of households that became non-eligible (i.e. using only their private latrine or planning to move during the next year) was 12%. Figure 11 shows that the non-response and non-eligibility rates are balanced across treatment arms. To keep a sample size in line with our power calculations, we interviewed replacement households randomly sampled and in a random order. In total, we interviewed 1575 households, with an average cluster size of 12 households. Table 2 shows the distribution of study households (eligible and interviewed at baseline), by treatment arm and city. The percentage of study households in the control group is slightly larger (5 ppts) due to a larger number of catchment areas. Also, the percentage of study households is slightly larger in Kanpur than in Lucknow, given the greater population density in the former

Table 2: Distribution of study households, by treatment arm and city

	Control	**Supply-side (SS)**	**SS + Information**	**Total**
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 (31.0%)	512 (32.5%)	1,575 (100%)

[8] The demographic groups are HH head, Spouse, Other males aged above 14, Other females aged above 14, Males aged 6-14, Females aged 6-14, Neighbour. Refer to Appendix B for further details about measurements.

4. **Rapid-assessment**. Since the intervention had a different component in the first two months (see section 0), we aimed at measuring the effect of the first component of the supply side intervention with a short-run follow-up. The purpose is to update information on household demographics and re-collect information on sanitation behaviour, priors, attitudes and health expectations linked to open defecation and CTs. We also collected information on the health status of household members and hygienic behaviour. Furthermore, we collected information on exposure to information campaigns and the use of tickets for the CT. Finally, the interviewer conducted observations of the household environment and the respondent. Timing was July-September 2018. The attrition rate was 8.12%. Figure 9 shows that attrition is slightly higher in the supply-side treatment arm, but it is balanced across treatment arms. To keep a sample size in line with our power calculations, we interviewed replacement households randomly sampled and in a random order. In total, we interviewed 1,532 households.
5. **Midline survey.** The purpose is to update information on household demographics and re-collect information on sanitation behaviour, priors, attitudes and health expectations linked to open defecation and CTs. We also collected information on the health status of household members, hygienic behaviour and access to drinking water. Furthermore, we collected information on exposure to information campaigns and the use of tickets for the CT. In addition, we collected information on expenditure on hygienic goods and sanitation facilities. In this round, we also collected data on household decision-making from one male and one female senior[9] member. Finally, the interviewer conducted observations of the household environment and the respondent. Timing was January-March 2019. We aimed to survey all households eligible and interviewed at baseline, as well as all replacement households interviewed during the rapid assessment (1,660 households). The attrition rate was 18.8%, slightly higher than during the rapid assessment because our timing coincided with school vacations and many families go back to their villages. Figure 9 shows that attrition is slightly higher in the control arm due to the displacement of a whole community (20%), but attrition is still balanced across treatment arms. To keep the sample size in line with our power calculations, we interviewed replacement households randomly sampled and in a random order. In total, we interviewed 1,578 households.
6. **Endline survey.** The purpose is to re-collect information on sanitation behaviour, priors, attitudes and health expectations linked to open defecation and CTs. We also collected information on the health status of household members, hygienic behaviour and access to drinking water. Furthermore, we collected information on exposure to information campaigns and the use of tickets for the CT. In addition, we collected information on expenditure on hygienic goods and sanitation facilities. In this round, we also collected data on household decision-making from one male and one female senior member. Moreover, we update information on

[9] We define seniority with the following guidelines: Household head; If the household head is absent, then the respondent should be the spouse of the household head; If household head and spouse are absent, revisit household; If household head and spouse are absent during revisit, then the respondent should be the most senior and knowledgeable member (above the age of 18) in terms of actively participating in the household's decision-making.

household demographics. In households with children aged 5 years or younger, we further interviewed the primary caregiver to collect information on child health and sanitation practices. Finally, the interviewer conducted observations of the household environment and the respondent. Timing was July-September 2019. We aimed to survey all households eligible and interviewed at baseline, as well as all replacement households interviewed during the rapid assessment and midline (1,896 households). The attrition rate was 13.13%. Figure 9 shows that attrition is slightly higher in the control arm due to the displacement of a whole community (14.3%), but attrition is still balanced across treatment arms. To keep the sample size in line with our power calculations, we interviewed replacement households randomly sampled and in a random order. In total, we interviewed 1,772 households.

We include below a table comparing demographics (collected in all data rounds) of households in the study the whole time, those that were lost after baseline and the replacement households. We can observe that the attrition households had fewer adult household members and more children. They were also more likely to have a male household head. Although the selection of replacement households was random, the replacement households were smaller and less likely to have a male household head.

Table 3: Comparison of demographics by attrition status

	All Mean	**Always Mean**	**Lost Mean**	**Replacement Mean**	**(3)-(2) Diff.**	**(4)-(2) Diff.**
	(1)	(2)	(3)	(4)	(5)	(6)
Number of HH members	5.00	5.11	4.66	4.87	-0.44^{***}	-0.24^{***}
Number of children	0.48	0.48	0.54	0.40	0.07^{***}	-0.08^{***}
Number of adult members	4.66	4.77	4.21	4.71	-0.55^{***}	-0.06
HH head male	0.73	0.73	0.78	0.67	0.05^{***}	-0.06^{***}
Observations	6465	4456	1175	834	5631	5290

7. **WTP survey**. The WTP survey is conducted to the most senior male and the most senior female in terms of decision-making from each household. Seniority is defined in terms of decision-making rather than age. We start by asking the price that females and males face if they want to use the CT. Next, we present a hypothetical scenario of a CT improved to its highest standard and provide a hypothetical choice between varying amounts of cash and a bundle of 10 tickets to use the hypothetical CT. We then play an incentivized WTP game, in which we offer the choice between varying amounts of cash and a bundle of 10 tickets to use the CT. Respondents could select tickets for male use and/or tickets for female use. The purpose is to measure WTP for CT use, to examine gender differences in willingness to pay and to evaluate whether exposure to the treatments affects CT WTP over time. At Endline, we also offered respondents the choice between varying amounts of cash and a monthly family pass for up to 5 members. The aim of this last section is to measure WTP for a family pass that was introduced by the Government at the same time as our study. In addition, a dictator/donation game was played, in which we provided participants Rs. 50 and gave them the choice to donate any amount to a fund to keep clean the CT. The aim of this dictator game is to measure demand for cleanliness of the CT. The

WTP survey is administered to all households who also respond to the household survey (two respondents per household). This survey was administered at the following time: baseline (March-June 2018), rapid-assessment (July—September 2018), midline (January-March 2019) and end-line (July-September 2019).

8. **CT usage survey and monitoring.** We implement a series of CT usage surveys and monitoring in order to gather information about CT outcomes for the duration of the project. We focus on three different types of indicators: 1) number of users; 2) observed cleanliness and maintenance and 3) bacteria count, to be implemented in all 110 study CTs. 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 proxy effort. At that time, the interviewer additionally observed and recorded the condition of toilets. User count is measured by observation of the interviewer during 1 hour using manual counters. This observation is recorded at dawn, when it was reported during the CT Census that traffic is the highest. Cleanliness is measured with both observation and with bacterial counts. Bacterial counts are measured with bacterial swabs and lab analysis. The detailed procedure adopted for the measurement of the bacteria count is presented in the next section. Timing was at baseline (March - June 2018), 2 months after BL/Follow-up 1 (July – September 2018), 4 months after BL/Follow-up 2 (October – November 2018), 6 months after BL/Midline (January – March 2019), 8 months after baseline/ Follow-up 4(April – May 2019) and 12 months after BL/Endline (June – August 2019). We aimed to survey all 110 community toilets eligible for the study. Figure 10 shows that attrition is slightly higher in the control and supply-side arm due to the fact that a few CTs closed or were abandoned without a caretaker (hence became free-to-use). In addition, 2 new CTs opened very close to CTs allocated to the supply treatment arm. Because slum dwellers also used these new CT, we conducted the CT usage survey and monitoring in these additional CTs during Follow-up 4 (April – May 2019) and Endline (June – August 2019). These new CTs did not increase the number of clusters since we consider them part of the same cluster as the old CTs due to their close proximity.

Behavioural measurements

This study includes a set of behavioural measurements that we describe in this section.

1. **WTP game**. We elicit WTP to use a hypothetical high-standard CT and the closest CT to the household. First, our hypothetical measure of WTP is elicited by describing a hypothetical scenario in which a nearby CT is improved to the highest standard (described in detail). We offer different hypothetical choices between cash and a bundle of 10 tickets for this high-standard CT. We elicit WTP for high-quality CTs, even above the current market price, albeit with the usual caveats of unincentivized elicitation. Second, our incentivized measure of WTP is elicited using a multiple price list and random draw. This methodology bounds the WTP for a good or lottery by prompting the individual to choose between different amounts of money or a bundle of tickets to use the CTs, before randomly drawing one allocation to be paid. Before the random draw, the respondent is asked to allocate the tickets they could win between tickets for male use and tickets for female use. At Endline, we also offer different choices between cash and a family

pass for unlimited use of the CT by 5 members within 1 month. Because the choice that will be paid at the end is randomly selected, each choice is independent. We identify the WTP for CT use as the point in which the participant is willing to take the tickets/pass as opposed to cash. We piloted the design of the experiment in great detail in order to apply the most effective methodology in our setting, both in terms of question framing and in terms of actual and expected price points. We measure the WTP of the household head and the most senior female in the household four times during the study: baseline (March-June 2018), rapid-assessment (July—September 2018), midline (January-March 2019) and end-line (July-September 2019).

2. **Dictator game.** We conduct a "dictator game/donation game" type of behavioural measurement with both slum-dwellers and caretakers. First, we measure slum-dwellers' willingness to contribute financially to the cleanliness of their CT in a game in which they can choose to donate part of an endowment to the improvement of cleanliness of the CT. Depending on the level of contribution, we provide a number of air-freshener packages to the CT. We are aware that this methodology may not be only capturing demand for cleanliness, but also trust and/or altruism towards its community and the CT management. We directly ask whether they trust their community and caretaker, as well as the level they expect their community to contribute. Secondly, we measure the caretaker's intrinsic motivation in a game in which they can choose to allocate part of an endowment to a sanitation project of our NGO partner FINISH Society. We measure the slum dwellers' demand for CT cleanliness and caretaker's intrinsic motivation four times during the study: baseline (March-June 2018), rapid-assessment (July—September 2018), midline (January-March 2019) and end-line (July-September 2019).
3. **Voice-to-the-people cards.** We provide households with the opportunity to fill-in "voice-to-the-people" style of cards, in which they are given the possibility to report the most important concerns from a set of different topics (including CT cleanliness) to local public officers. Respondents are informed that the content of these cards will be summarized and provided to local public officers. The aim of this game is to measure the political salience of CT cleanliness and WASH concerns among slum dwellers. In addition, we distribute the cards to a set of randomly selected neighbours to understand if concerns are different for households exposed to the study. We conduct this SCA at midline (January-March 2019).
4. **Community game.** We conduct two behavioural games with several participants[10] in the community to measure social capital and community preference for CT cleanliness. The former is measured applying a Public Good Game (PGG) and the latter, applying a modified version of the standard PGG. The version of the PGG we use is called a voluntary contribution mechanism (VCM). The VCM experiment captures trust and the willingness to cooperate among the members of a specific group by choosing whether to keep the endowment or to invest part or all of it in a group pot where the benefits for all members increase by a multiplier with the level of contributions. The modified

[10] We play simultaneously with three groups in each community. Participants in each group are always even and range between 4 and 6.

version gives the option to keep the endowment or to invest part or all of it in a pot to buy cleaning products for the CT, which also increases by a multiplier with the level of contributions. The games are 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 to not contribute at all, while the social optimum is to invest in the pot. The incentive to invest in the pot is given by individual motivations concerning the group's well-being. Deviations from the dominant strategy are thus considered measures of social capital in the group, determined by attributes such as altruism, trust, social distance, reciprocity, sympathy, among others (Attanasio, Pellerano, and Reyes 2009). We expect both games to capture social capital; hence the difference between both games determines community preference for CT cleanliness. The outcome of each game is individually determined; we play these two games sequentially (in a random order) and the pay-off from one of the two games is randomly selected. In addition, we randomly vary the multiplier (double or triple the total contribution) across groups to determine how an external contribution (such as a government intervention) can affect the willingness to contribute to a public good. Furthermore, we play with three types of groups within each community: only females, only males and mixed. This allows us to observe differences in social capital and preference for CT cleanliness by gender. Finally, we collect follow-up questions on motivations and social network to understand the mechanisms behind the outcome.[11]

The last two behavioural measures involve structured community activities (SCA). SCA have the advantage of focusing on concrete, real-world scenarios that allow unobtrusive measurement of leader and community decision-making, more objectively than lab experiments, hypothetical vignettes or surveys (Casey, Glennerster, and Miguel 2012).

Laboratory tests
Among quantitative data, we collected information from laboratory tests. The specific source is defined here:

1. **CT bacteria count.** In order to collect measures of cleanliness at the level of CT, we collect three samples of bacteria swabs for each round of CT usage survey. To obtain an average measure of bacteria count in a CT, we follow this procedure. We first randomize all CTs in two groups: a "male" and a "female" group. In the "male" group the bacteria swabs are collected in the male area of the CT throughout the study, while in the "female" group the same is done in the female area of the CT. Secondly, during each visit the enumerator makes for the CT usage survey, three samples of bacteria swabs are collected in different points of the CT. The first two samples are collected from two randomly-selected cubicles, the location of which are provided by the research team and which varies in each round of data collection in order to avoid the caretaker to focus on cleanliness of a specific position in the CT. The exact position for the bacteria swab is the mid-point between the entrance wall of the cubicle and the

[11] The measurement is currently underway; we therefore do not include any result relative to this behaviour game in this document.

latrine/water. The randomly-selected cubicles are provided to the enumerator in the form of a number x. To collect swab from cubicle x, the enumerator is instructed to face the cubicles and count from right to left. If there are two rows of cubicles, the enumerator is instructed to face the row that is first observed while entering the cubicle hallway and count from right to left. A third sample is collected in the position of the first step into the cubicle hallways, defined as the place where the enumerator walks first to enter the cubicle hallways and where most people walk by. This sample aims at collecting a bacteria sample in the area where most people walk. Bacteria swabs are collected at baseline (March - June 2018), 2 months after BL (July – September 2018), 4 months after BL (October – November 2018), 6 months after BL (January – March 2019), 8 months after baseline (April – May 2019) and 12 months after BL (June – August 2019).

2. **Household water bacteria counts**. In order to measure the quality of water that slum-dwellers have access to, we collect two samples of water from each catchment area. To obtain an average measure of bacteria count in drinking water, we follow this procedure. We first identify the source of water used by the majority of eligible households for drinking purposes. From the sampled households for the study (around 15), we randomly select two households from which water samples should be collected. Water samples are collected from the source, rather than the collected water, to prevent confounding hygienic practices with the quality of water. Whenever the source of water is shared by all eligible households (e.g. public tap), we only take one sample from that source. Water bacteria samples are collected at Baseline (March - June 2018).

Figure 9: Consort Flow Chart for household surveys

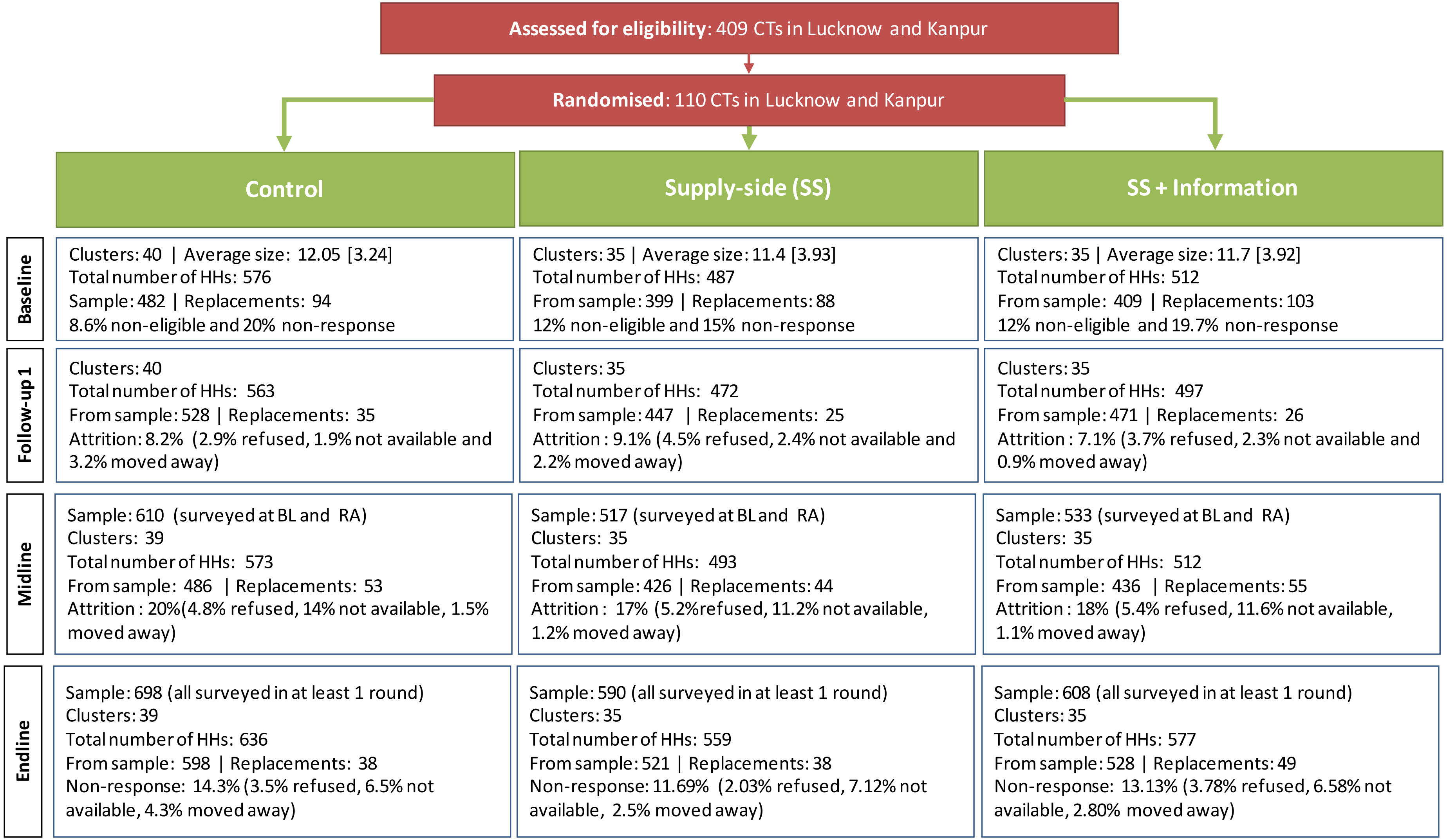
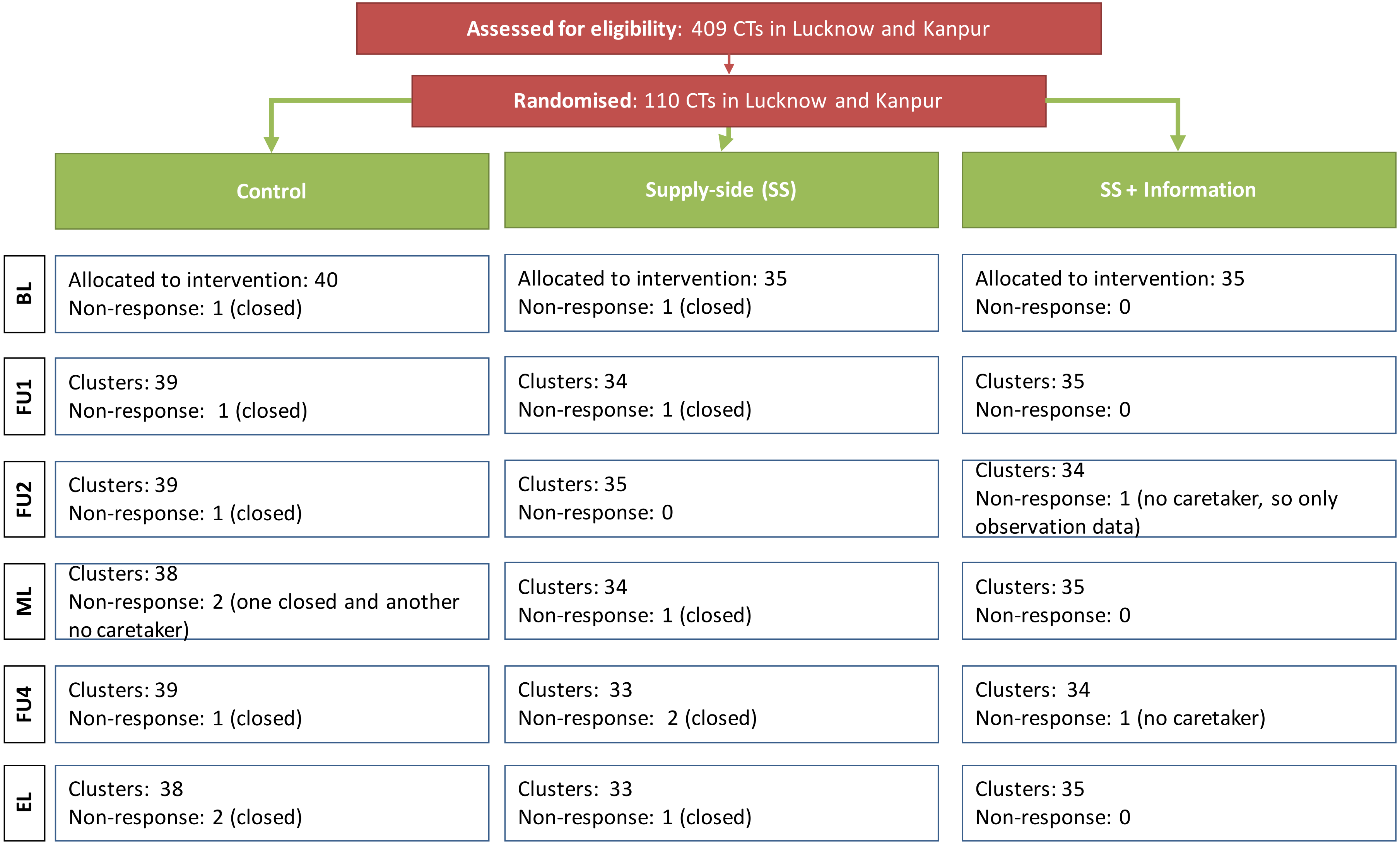


Figure 10: Consort Flow Chart for CT surveys



4. Findings

4.1 Intervention implementation fidelity

4.1.1 Supply intervention (Treatment groups 1 & 2)

CT Grant Scheme. Overall, the uptake of the grant scheme was 95 percent. The grant scheme was not provided to 3 CTs allocated to the supply-side treatment arm – 2 in Lucknow and 1 in Kanpur – because they were closed for refurbishment.[12] *:* Table 4 shows the percentage of CTs that selected each grant choice by city: around 80% of CTs selected either repairs or deep cleaning. Although all CTs allocated to the supply-side intervention selected the optimal grant choice, the take-up of the grant was not universal. From the 68 CTs allocated to the supply-side intervention, we were only able to provide the grant to 95.7%. We were not able to carry out the works in 3 CTs because they were temporarily closed.

Table 4: Selection of grant choice and uptake

	Repairs	Deep cleaning	Training + cleaning agents and tools	Grant scheme uptake
Lucknow	14	9	9	30 (93.8%)
Kanpur	14	19	3	35 (97.2%)
Total	28 (41.2%)	28 (41.2%)	12 (17.7%)	65 (95.7%)

We asked caretakers at endline if they recall whether the CT where they work received a grant scheme to improve cleanliness and quality. Table 5 shows the percentage of caretakers that reply "No" or "Yes" to the question by city and treatment arm. In total, almost 1 year after the grant scheme, 89% of caretakers allocated to either the supply or supply plus information treatment arms were able to recall the grant scheme. The percentage of caretakers recalling the grant scheme is larger in Kanpur than in Lucknow (95 percent vs. 81 percent). We also asked caretakers whether they knew of any other CT receiving a grant scheme from us. We find that only 5 percent of caretakers heard of another CT receiving a grant scheme.

Table 5: Caretaker's recall of grant scheme at Endline

		Supply	Information + Supply	Total
Lucknow	No	1 (5.3%)	3 (18.8%)	4 (11.4%)
	Yes	18 (94.7%)	13 (81.3%)	31 (88.6%)
Kanpur	No	3 (16.7%)	1 (5.3%)	35 (100%)
	Yes	15 (83.3%)	18 (94.7%)	33 (89.2%)
Total	No	4 (10.8%)	4 (11.4%)	8 (11.1%)
	Yes	33 (89.2%)	31 (88.6%)	64 (88.9%)

CT Financial Reward Scheme. Two months after the CT Grant Scheme, we announced the Financial Reward Scheme. From month 6 and every two months until the end of the study, our partner FINISH paid the financial reward to those CTs that achieve the conditionality, which was determined based on CT observations we collected. The last payment of the financial reward (payment 4) was done after the study's end-line survey.

[12] We could still consider these CTs as treated because of the works conducted on them.

Table 5 shows the uptake of the Financial Reward Scheme, including the announcement and all payment rounds. In general, the uptake is almost perfect. 97% of caretakers (100% in Kanpur and 94% in Lucknow) were available and/or willing to listen to the initial announcement of the scheme. The uptake of the first payment is on average 97% and 95% during the following rounds, again with higher uptake in Kanpur than in Lucknow throughout. The main reason for the imperfect uptake is the change of caretakers. When caretakers were leaving the CT[13], FINISH was announcing the scheme to the new caretaker and paying the incentive in the following round. There was no case in which a caretaker moved from a treated CT to work to a control CT. In a few cases, the CT was temporarily close and/or no caretaker was assigned to work on it.

Table 6: Financial Reward Scheme uptake

	Announcement	Payment 1	Payment 2	Payment 3
Lucknow	31 (93.9%)	31 (93.9%)	31 (93.9%)	31 (93.9%)
Kanpur	37 (100%)	37 (100%)	35 (94.6%)	35 (94.6%)
Total	68 (97.1%)	68 (97.1%)	66 (94.3%)	66 (94.3%)

4.1.2 Information campaign (Treatment group 2 only)

Door-to-door campaign and provision of leaflets. The door-to-door campaign and provision of leaflets was conducted three times: at the end of the baseline, rapid assessment and mid-line surveys. We were able to conduct the information session and leave a leaflet in almost all households surveyed across the three rounds and allocated to the supply plus information arm: provided to 432 HHs (99.5% of eligible were treated) after the baseline survey, 441 HHs (99.1% of eligible were treated) after the rapid assessment and 507 HHs (99% of eligible were treated) after the midline survey. Due to an error in the field, 12 HHs allocated to control (2.1% of HHs in this arm) and 6 HHs (1.2% of HHs in this arm) received the information campaign after the rapid assessment. We are confident that this small contamination rate does not threat the internal validity of our estimates. No contamination happened at later stages.

Posters. All CTs allocated to the supply-side plus information treatment arm always had the informative posters on their walls. Twice per month, we visited all CTs to corroborate that all 5 posters were properly pasted in the CT walls. In case a poster fell or was in bad shape, we replaced it with a new one.

Table 7 shows evidence of households in the study recalling being exposed to the information campaign. Households allocated to the information campaign were 9 ppts (16% over control mean) more likely to report hearing a message about water, sanitation and hygiene (WASH) through communication technologies; 16 ppts (48% over control mean) more likely to report seeing these messages in the community toilet; 7 ppts (28% over control mean) more likely to report being visited by an agent talking about WASH or the construction of toilets; and 6 ppts (18% over control mean) more likely to report hearing of events organized in the community about WASH.

[13] We paid the financial reward to old caretakers, as long as they were reachable and willing to arrange the payment.

Table 7: Exposure to information campaign

	Control mean	Supply-side		Supply-side + Info		Obs.
		Coef.	SE	Coef.	SE	
	(1)	(2)	(3)	(4)	(5)	(6)
Heard messages	0.57	0.02	0.03	0.09***	0.02	9780
CT posters	0.33	0.02	0.03	0.16***	0.03	6716
Visited	0.25	0.01	0.02	0.07***	0.02	9780
Community activity	0.32	0.02	0.02	0.06***	0.02	9780

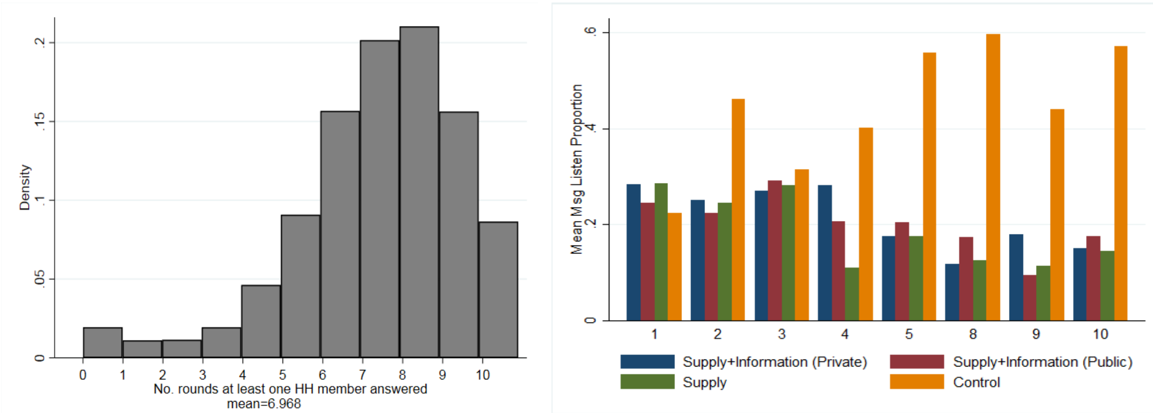
Note. Sample includes all data rounds. Exposure to posters in the CT was only collected at midline and endline rounds. All specifications include round dummies and strata variables as control: (i) city and (ii) managed by main provider. Standard errors clustered by CT catchment area. Statistical significance denoted by * p<0.1, ** p<0.05 and *** p<0.01.

Control mean are provided because households randomly allocated to “Control” or “Supply-side only” could have been exposed to information about water, sanitation and hygiene (WASH) provided by other agents (other NGOs or government). For example, as part of the Swatch Bharat Mission, there are some CTs that already had posters highlighting the importance of adequate sanitation.

In light of the significant effort conducted by the Government of India to provide information on WASH through the Swachh Bharat mission, as reflected by the high report of exposure in control and supply-side only areas (columns 1 and 2), it is reassuring to see that our intervention had an impact over and above their efforts. Households randomly allocated to our information campaign, were more likely to recall receiving information about WASH through all means (voice messages, CT posters, household visits and community activities).

Voice messages. We sent voice messages to all mobile phones provided at baseline for each household allocated to the supply-side plus information treatment arm. We measure if messages where listened and for how long. The program we use to systematically send sends voice messages via calls to all mobile phones in the sample provides data on: (i) whether the mobile number was valid or not; (ii) whether the call was answered or not; and (iii) how many minutes was the call answered. The uptake of this component is surprisingly high. The left panel in Figure 11 shows that households listened on average to 7 of the 10 monthly rounds of messages. The right panel in Figure 11 shows that households in general listened to a good proportion of the message. More than 20% of the private and public health information messages, as well as supply-side messages, were heard. This percentage drops after round 4, exactly when we start repeating the sequence of messages. The finding that households in control areas listen to a larger proportion of messages sent is mechanical, as their messages were shorter. Using the average time listened of each message, we can determine the exact information that participants heard. Doing so is useful for 2 reasons. For one, we know which part of the information households heard and second, we can check whether treated households were exposed to the message sent to supply only and control households. If they were, we could be certain that any observed differences in behaviour was due to the key message delivered only to the supply + information treated households. Unfortunately, they did not listen long enough to be exposed to the information provided to supply only and control (as that part of the message was added in the end). Therefore, we cannot rule out that control and supply-side participants changed their knowledge, attitudes and behaviour as a result of greater salience of the opening hours of the CTs (control message) and learning that the CT received aid (supply-side).

Figure 11: Descriptive statistics about voice messages



Note: The left figure indicates the number of rounds at least one HH member answered message by HH. The right figure shows the mean proportion of message listened across message categories and rounds.

4.2 Impact analysis

This section presents the main effects of the intervention. For a more detailed definition of primary and secondary outcomes, refer to the pre-analysis plan for the definition (see Appendix B).

4.2.1 Descriptive statistics and balance tables

This section provides descriptive statistics (i.e. means and standard deviations) at baseline, as well as balance tests for primary and secondary outcomes, also performed using pre-intervention data. Please refer to Appendix B – Pre-Analysis Plan for a detailed description of how we collect and measure each of the outcome variables.

Table 8 uses data from the CT census, providing important characteristics of community toilets in Lucknow and Kanpur. The analysis is separated by selection into the study (pay-to-use, located close to a residential area, used by residents and not another pay-to-use community toilet within 300 meters), or not selected. More than 60% of CTs selected for the study are managed by the main provider of CT services in the region (73% in Lucknow and 59 % in Kanpur). Selected CTs are all located in a residential area, around 30% are surrounded by market area and street vendors, though in both cities these shares are lower for selected than not selected CTs. Approximately 80% of CTs are near a road and 20% near a government office. 12% of the selected community toilets are in a purely residential area. In line with our selection strategy, selected CTs are more disperse than non-selected CTs – the former are on average located more than 500 meters away from the next closest pay-to-use CT.

Selected CTs are less frequently cleaned than non-selected CTs and a lower percentage of selected CTs in Kanpur are frequently cleaned than Lucknow – 37% are cleaned more than twice per day in Lucknow and only 17% in Kanpur. CTs are slightly smaller in

Lucknow than in Kanpur in terms of defecation cubicles. Selected CTs have on average 13 defecation cubicles, slightly higher than that of non-selected CTs (~10 cubicles). On average, 40% of cubicles in selected CTs are female only, slightly higher than non-selected CTs (~10%). The percentage of females' cubicles in selected CTs in Lucknow and Kanpur is similar. Around 35% of selected CTs have an area nearby used for open defecation, compared to only around 20% in non-selected CTs. CTs are on average open for 18 hours and employ two staff. In around 85% of cases, the caretaker lives on the ground, implying that these CTs have always a worker present. CTs tend to rely on an improved type of sanitation facility –more than 65% of CTs are pour-flush to sewerage. Residents are the main users in a higher percentage of selected CTs compared to non-selected CTs and in a higher percentage of selected CTs in Kanpur compared to Lucknow –98 and 58%, respectively. While more than 80% of selected CTs in Kanpur offer fee exemptions for females and children, only 17% of selected CTs in Lucknow do so. On average, a low percentage of CTs offer a family pass – 4% in selected CTs in Lucknow and 10% in Kanpur.

Table 8: Descriptive statistics of CT sample population, by city and study participation

	Total CT population		**Lucknow**		**Kanpur**	
	Selected	**Not selected**	**Selected**	**Not selected**	**Selected**	**Not selected**
	(1)	(2)	(3)	(4)	(5)	(6)
Toilet is pay-to-use	1.00	0.85	1.00	0.85	1.00	0.85
	(0.0)	(0.4)	(0.0)	(0.4)	(0.0)	(0.4)
Total users	231.41	246.38	308.04	264.40	157.61	118.57
	(324.9)	(294.3)	(399.7)	(309.4)	(209.8)	(61.3)
Manager: Main provider	0.66	0.58	0.73	0.60	0.59	0.56
	(0.5)	(0.5)	(0.4)	(0.5)	(0.5)	(0.5)
Surrounding: Residential	0.99	0.81	0.98	0.65	1.00	0.99
	(0.1)	(0.4)	(0.1)	(0.5)	(0.0)	(0.1)
Surrounding: Market	0.33	0.42	0.33	0.57	0.33	0.25
	(0.5)	(0.5)	(0.5)	(0.5)	(0.5)	(0.4)
Surrounding: Street Vendors	0.33	0.49	0.38	0.66	0.28	0.31
	(0.5)	(0.5)	(0.5)	(0.5)	(0.5)	(0.5)
Surrounding: Road	0.84	0.69	0.85	0.81	0.83	0.56
	(0.4)	(0.5)	(0.4)	(0.4)	(0.4)	(0.5)
Surrounding: Government office	0.25	0.36	0.29	0.55	0.21	0.15
	(0.4)	(0.5)	(0.5)	(0.5)	(0.4)	(0.4)
Surrounding: Only residential	0.12	0.21	0.12	0.10	0.12	0.33
	(0.3)	(0.4)	(0.3)	(0.3)	(0.3)	(0.5)
Distance to closest CT	0.54	0.30	0.55	0.39	0.53	0.20
	(0.4)	(0.4)	(0.6)	(0.6)	(0.3)	(0.1)
Cleaned more than twice per day	0.26	0.77	0.37	0.64	0.17	0.90
	(0.4)	(0.4)	(0.5)	(0.5)	(0.4)	(0.3)
Total toilets	13.00	10.64	12.52	10.02	13.45	14.10
	(5.6)	(6.2)	(4.8)	(5.6)	(6.3)	(8.4)
Female toilets (pct. total)	0.40	0.36	0.39	0.35	0.42	0.41
	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
OD area close	0.35	0.19	0.35	0.18	0.36	0.24

	Total CT population		Lucknow		Kanpur	
	Selected	Not selected	Selected	Not selected	Selected	Not selected
	(1)	(2)	(3)	(4)	(5)	(6)
	(0.5)	(0.4)	(0.5)	(0.4)	(0.5)	(0.4)
Total hours open	17.77	18.29	18.42	18.44	17.16	17.48
	(1.5)	(1.6)	(1.6)	(1.6)	(1.1)	(1.4)
Employees: Staff	1.85	2.28	1.85	2.30	1.86	2.19
	(0.7)	(1.0)	(0.8)	(1.0)	(0.7)	(1.0)
Employee present: Always	0.86	0.74	0.85	0.72	0.88	0.86
	(0.3)	(0.4)	(0.4)	(0.5)	(0.3)	(0.4)
Sanitation: pour-flush to sewerage	0.69	0.65	0.67	0.63	0.71	0.76
	(0.5)	(0.5)	(0.5)	(0.5)	(0.5)	(0.4)
Main users: Residents	0.76	0.56	0.52	0.24	0.98	0.85
	(0.4)	(0.5)	(0.5)	(0.4)	(0.1)	(0.4)
Fee exemption: children and females	0.53	0.20	0.17	0.13	0.86	0.62
	(0.5)	(0.4)	(0.4)	(0.3)	(0.4)	(0.5)
Fee charged as monthly family pass	0.07	0.07	0.04	0.01	0.10	0.13

Note. Standard deviations in parenthesis. Because this table presents descriptive statistics from the CT census, the number of total users is based on the caretaker's report.

Table 9 uses the slum census data and presents descriptive statistics (i.e. means and standard deviations) of important socio-economic and demographic characteristics of the full sample population of slum dwellers in Lucknow and Kanpur, by eligibility for the study. Around 20% and 27% of households living in the cities' slums are eligible to be part of our study in Lucknow and Kanpur, respectively. The final sample of eligible households to be interviewed as part of our study is a random draw of these total eligible households by catchment area, implying that our sample is representative of this group. By construction, eligible households are those that use CTs or any other sanitation solution besides private latrines and who are not planning to migrate during the following 18 months (the duration of the study). In 72% of eligible households in Lucknow at least one member uses CTs, and this is the case in 63% of eligible households in Kanpur. In 34% of eligible households in Lucknow at least one member practices open defecation compared to 37% of eligible households in Kanpur. While around 10% of eligible households own a toilet, almost all non-eligible households own a private toilet. Eligible households live on average 20 meters closer to the CT than non-eligible households, who live on average 160 meters away from the closest CT.

Table 9: Descriptive statistics of HH sample population, by city and eligibility

	Total population		Lucknow		Kanpur	
	Eligible	Non-eligible	Eligible	Non-eligible	Eligible	Non-eligible
	(1)	(2)	(3)	(4)	(5)	(6)
HH uses CT	0.66	0.01	0.72	0.01	0.63	0.01
	(0.5)	(0.1)	(0.4)	(0.1)	(0.5)	(0.1)
HH practices OD	0.36	0.01	0.34	0.01	0.37	0.01
	(0.5)	(0.1)	(0.5)	(0.1)	(0.5)	(0.1)
Owns a toilet	0.12	0.99	0.10	0.99	0.14	0.99
	(0.3)	(0.1)	(0.3)	(0.1)	(0.3)	(0.1)
Distance to closest CT	141.32	163.38	143.99	169.33	139.93	159.24
	(94.2)	(98.0)	(98.6)	(103.8)	(91.8)	(93.6)
HH members: Total	4.88	5.10	4.91	5.22	4.87	5.02
	(2.1)	(2.3)	(2.1)	(2.4)	(2.1)	(2.2)
HH members: No children	0.69	0.73	0.66	0.71	0.71	0.74
<5y	(0.5)	(0.4)	(0.5)	(0.5)	(0.5)	(0.4)
Household head male	0.79	0.80	0.77	0.77	0.81	0.83
	(0.4)	(0.4)	(0.4)	(0.4)	(0.4)	(0.4)
Age of the household head	44.56	47.62	43.29	47.30	45.21	47.84
	(12.5)	(12.8)	(12.3)	(12.5)	(12.5)	(13.0)
Religion: Hindu	0.83	0.74	0.75	0.69	0.87	0.79
	(0.4)	(0.4)	(0.4)	(0.5)	(0.3)	(0.4)
Cast: General	0.08	0.20	0.08	0.22	0.08	0.19
	(0.3)	(0.4)	(0.3)	(0.4)	(0.3)	(0.4)
Ration card: No	0.39	0.39	0.38	0.32	0.39	0.44
	(0.5)	(0.5)	(0.5)	(0.5)	(0.5)	(0.5)
Own mobile: HH heads or oldest HH members	0.65	0.72	0.68	0.73	0.63	0.71
	(0.5)	(0.4)	(0.5)	(0.4)	(0.5)	(0.5)
Ownership: Own	0.68	0.80	0.64	0.85	0.71	0.78
	(0.5)	(0.4)	(0.5)	(0.4)	(0.5)	(0.4)
Dwelling from strong	0.57	0.86	0.60	0.87	0.56	0.86
material	(0.5)	(0.3)	(0.5)	(0.3)	(0.5)	(0.3)
Water: Piped	0.68	0.71	0.84	0.93	0.60	0.55
	(0.5)	(0.5)	(0.4)	(0.3)	(0.5)	(0.5)
Observations	5,554	22,493	1,891	9,249	3,663	13,244

Note. Standard deviation in parenthesis.

On average, both eligible and non-eligible households have five members, around 70% do not have children and in near 80% the household head is a man. Eligible and non-eligible households differ in a number of socio-demographic characteristics. Eligible households have younger heads (4 year younger in Lucknow and 3 in Kanpur), a higher prevalence of household heads that are Hindu (6 and 8 pp higher in Lucknow and Kanpur, respectively) and a lower prevalence of household heads belonging to the general caste (14 and 11 pp lower in Lucknow and Kanpur, respectively). While in Lucknow a higher percentage of eligible households do not own a ration card (38% versus 32%), this is the opposite in Kanpur (39% versus 44%). Several characteristics suggest that eligible households are less well off. While in 65% of households the head or the oldest members own a mobile, this is the case for 70% of non-eligible households (similar across both cities). Similarly, a lower percentage of eligible households own the dwelling they are living in (21 and 7 pp difference in Lucknow and Kanpur, respectively) and have a dwelling of strong material (27 and 30 pp difference in Lucknow and Kanpur, respectively). While in Lucknow a lower percentage of eligible households have access to piped water (84% versus 93%), this is the opposite in Kanpur (60% versus 55%).

Table 10 presents balance test for primary outcomes at the household level, Table 11 presents balance test for secondary outcomes at the household level, and Table 12 presents balance test for primary and secondary outcomes at CT level. Each table presents in column (1) the mean and standard deviation of the variable when sample is restricted to the control group only. In column (2), for each individual and household outcomes, we begin by testing for balance on the joint treatment groups versus the control group, while in columns (3) and (4), we look at balance specifically within each treatment group. In column (5), we then test for jointly-significance of differences estimated in column (3) and (4). A similar procedure is followed for catchment-area-level outcomes. Overall, results show that randomization was successful in creating observationally equivalent groups.

Table 10: Balance test: primary outcomes at household level

	(1) **Control**	(2) **Any treatment**	(3) **Treatment1: SS**	(4) **Treatment 2: SS + Info**	(5) **Joint test 1-2**
	Mean [sd]	Diff. (se)	Diff (se)	Diff (se)	p-value [N]
WTP CT use - male	1.482	-0.054	0.081	-0.190	0.517
	[2.048]	(0.168)	(0.197)	(0.213)	[1304]
WTP CT use - female	1.367	-0.061	-0.166	0.040	0.521
	[2.018]	(0.153)	(0.165)	(0.208)	[1381]
Hypothetical WTP - male	5.608	-0.151	-0.158	-0.145	0.834
	[3.046]	(0.251)	(0.303)	(0.308)	[1304]
Hypothetical WTP - female	5.401	-0.195	-0.175	-0.214	0.732
	[3.066]	(0.257)	(0.310)	(0.277)	[1381]
Members used at least once CT (any group)	0.819	0.013	-0.006	0.032	0.648
	[0.385]	(0.045)	(0.053)	(0.047)	[1572]
All members always use CT	0.224	-0.008	-0.010	-0.007	0.957
	[0.418]	(0.028)	(0.034)	(0.031)	[1572]
Members practice at least once OD (any group)	0.322	-0.033	0.014	-0.078	0.181
	[0.468]	(0.050)	(0.061)	(0.053)	[1572]
Number of CT users within 1 hour	37.959	-2.569	-1.698	-3.396	0.547
	[13.614]	(2.811)	(3.360)	(3.091)	[1551]
% users paid within 1 hour	0.645	-0.010	-0.052	0.031	0.440
	[0.306]	(0.057)	(0.069)	(0.062)	[1531]
WTP CT cleanliness - male	12.718	-0.513	-1.270	0.248	0.220
	[11.935]	(0.752)	(0.902)	(0.859)	[1304]
WTP CT cleanliness - female	12.044	-0.848	-0.958	-0.742	0.434
	[10.521]	(0.662)	(0.785)	(0.743)	[1381]

Note. Column (1) reports sample mean and standard deviation in brackets for the control group. Column (2) 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 (3)-(4) report the difference with the control group for each treatment group. Standard errors clustered at community toilet level are reported in parentheses. Column (5) present a joint test of significance of the coefficients for each treatment dummy. Significance level indicated by: *** p<0.01, ** p<0.05, * p<0.1.

Similarly, secondary outcomes are also generally well balanced. We only find three minor deviations. For one, the indicator whether the respondent washes his/her hands after defecating is slightly higher in the Supply only treatment groups, significant only at the 10 percent level and not jointly significant. The household dirtiness index is slightly lower when pooling the two treatments and compare to control. Finally, the only imbalance that holds through when conducting the joint F-test is in the variable capturing knowledge of private health benefits from sanitation where we find that the supply only treatment group has a 0.037 points lower index value compared to control. The magnitude of the difference, which is significant at the five percent level, is hence not very big. We conclude, that we have overall a very well-balanced sample considering the primary and secondary outcomes of our study.

Table 11: Balance test: secondary outcomes at household level

	(1) **Control** Mean [sd]	(2) **Any treatment** Diff. (se)	(3) **Treatment1 : SS** Diff (se)	(4) **Treatment 2: SS + Info** Diff (se)	(5) **Joint test 1-2** p-value [N]
Respondent and spouse wash hands with soap after defecating	0.871	0.027	0.034*	0.020	0.245
	[0.335]	(0.018)	(0.020)	(0.021)	[1572]
HH dirty index	0.510	-0.042*	-0.042	-0.043	0.207
	[0.290]	(0.024)	(0.030)	(0.027)	[1574]
At least 1 member with diarrhoea during the previous 15 days	0.097	0.016	0.001	0.031	0.265
	[0.297]	(0.019)	(0.021)	(0.022)	[1574]
At least 1 member with fever during last 15 days	0.214	-0.013	-0.011	-0.014	0.883
	[0.410]	(0.026)	(0.030)	(0.029)	[1574]
Positive CT attitudes index	0.575	0.002	-0.011	0.015	0.207
	[0.152]	(0.013)	(0.016)	(0.014)	[1574]
Diff in expected adult illness likelihood between OD and CT	0.626	0.013	0.008	0.017	0.811
	[0.365]	(0.023)	(0.030)	(0.026)	[1574]
Diff in expected adult illness likelihood between dirty and clean CT	0.723	-0.016	-0.014	-0.017	0.744
	[0.334]	(0.020)	(0.025)	(0.024)	[1574]
Diff in expected female unsafe incidents likelihood between OD and CT	0.767	0.040	0.044	0.035	0.321
	[0.367]	(0.026)	(0.031)	(0.029)	[1574]
Adequate knowledge of private health index	0.907	-0.020	-0.037**	-0.002	0.097*
	[0.196]	(0.013)	(0.018)	(0.013)	[1574]
Adequate knowledge of public health index	0.846	-0.008	-0.028	0.012	0.192
	[0.216]	(0.016)	(0.021)	(0.017)	[1574]

Note. Column (1) reports sample mean and standard deviation in brackets for the control group. Column (2) 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 (3)-(4) report the difference with the control group for each treatment group. Standard errors clustered at community toilet level are reported in parentheses. Column (5) present a joint test of significance of the coefficients for each treatment dummy. Significance level indicated by: *** p<0.01, ** p<0.05, * p<0.1.

Table 12 finally shows balancedness in CT level outcomes and again, we see that our study sample is also nicely balanced at this level. The only imbalance we observe is for the percentage of time that a caretaker spends on cleaning, which is five pp lower in the supply only treatment arm, significant at the 5% level. This difference comes through when pooling the treatments but is not jointly significant when estimating the balancedness equation with both treatment indicators.

Table 12: Balance test: primary and secondary outcomes at CT level

	(1) **Control** Mean [sd]	(2) **Any treatment** Diff. (se)	(3) **Treatment1 : SS** Diff (se)	(4) **Treatment 2: SS + Info** Diff (se)	(5) **Joint test 1-2** p-value [N]
Total number of users	37.410	-0.391	-1.361	0.552	0.868
	[13.331]	(3.065)	(3.604)	(3.575)	[108]
% users that pay	0.672	-0.041	-0.085	0.001	0.293
	[0.304]	(0.053)	(0.061)	(0.061)	[108]
Dirtiness index	0.312	0.002	0.026	-0.021	0.727
	[0.279]	(0.049)	(0.058)	(0.057)	[108]
Bad quality index	0.394	0.013	0.034	-0.008	0.620
	[0.194]	(0.038)	(0.044)	(0.044)	[108]
Bacteria found: E-coli	1.000	0.000	0.000	0.000	.
	[0.000]	(0.000)	(0.000)	(0.000)	[108]
Bacteria count: E-coli	64.987	2.154	5.148	-0.753	0.737
	[46.492]	(6.832)	(8.023)	(7.959)	[108]
Bacteria found: Bacillus	0.974	0.011	-0.005	0.026	0.581
	[0.160]	(0.027)	(0.032)	(0.031)	[108]
Bacteria count: Bacillus	12.004	0.783	0.838	0.730	0.812
	[7.692]	(1.215)	(1.430)	(1.419)	[108]
Bacteria found: Mould	0.872	-0.003	-0.051	0.044	0.476
	[0.339]	(0.065)	(0.076)	(0.076)	[108]
Bacteria found: Staphyl	0.641	-0.034	-0.090	0.020	0.459
	[0.486]	(0.078)	(0.091)	(0.090)	[108]
Bacteria found: Salmonella	0.308	-0.061	0.016	-0.136	0.278
	[0.468]	(0.088)	(0.102)	(0.101)	[108]
Bacteria found: Klebsiella	0.564	-0.011	-0.027	0.003	0.875
	[0.502]	(0.053)	(0.063)	(0.062)	[108]
% time allocated to clean/supervise cleaner	0.483	-0.041*	-0.051*	-0.031	0.161
	[0.136]	(0.023)	(0.027)	(0.027)	[108]
% time allocated to collect fees	0.356	-0.009	-0.018	-0.001	0.639
	[0.108]	(0.017)	(0.020)	(0.020)	[108]
CT cleaned at least twice per day	0.868	-0.028	-0.015	-0.040	0.890
	[0.343]	(0.073)	(0.085)	(0.085)	[107]
CT adequate cleaning index	0.735	-0.003	-0.013	0.007	0.870
	[0.187]	(0.033)	(0.038)	(0.038)	[108]

Note. Column (1) reports sample mean and standard deviation in brackets for the control group. Column (2) 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 (3)-(4) report the difference with the control group for each treatment group. Standard errors clustered at community toilet level are reported in parentheses. Column (5) present a joint test of significance of the coefficients for each treatment dummy. Significance level indicated by: *** p<0.01, ** p<0.05, * p<0.1.

4.2.2 Research analyses

The evaluation design for the comparison of different interventions examines differences in outcomes across individuals and households living in catchment areas (clusters, sometimes referred to as communities) assigned to different treatment groups. Since these areas were allocated at random to different treatment groups, they are expected to be identical on average on all their other characteristics, observed or unobserved. A simple comparison of households across clusters gives us the impact on household-level outcomes of implementing one versus another intervention. Similarly, a comparison of cluster-level outcomes across clusters assigned to different treatment groups identifies the effect for cluster-level outcomes.

We start by focusing on the general effect of improving the quality of supply – without (group 1) or with providing information to the households in study clusters (group 2). For individual and household outcomes, let $T1_{im}$ and $T2_{im}$ be indicator variables that takes value 1 if respondent (or household) *i* lives in a catchment area *m* assigned to groups 1 or 2 respectively, and equal to 0 if respondent *i* lives in a catchment area m that is part of the catchment areas assigned to the control group. In order to estimate the effect of the interventions on the outcome Y_{imt} at time t, we estimate the following model:

$$Y_{imt} = \alpha + \beta_1 T1_{im} + \beta_2 T2_{im} + \boldsymbol{X}'_{imt}\gamma + \theta_m + \tau_t + \varepsilon_{imt} \qquad (1)$$

where $\boldsymbol{X}'_{imt}$ is a vector of controls, including household and individual characteristics, and geographical control variables. θ_m are strata dummy variables capturing the dimensions along which the randomization was stratified. τ_t are dummy variables capturing time-specific effects. ε_{im} is a residual idiosyncratic error term picking up unobserved determinants of the outcome of interest, clustered at the catchment area level since this was the unit of randomisation.

The impact on outcome Y_{imt} of improving the supply provision is given by β_1 and the impact of improving supply combined with providing information about the importance of improved sanitation practices is given by β_2.

Similarly, for **catchment-area-level outcomes**, we can estimate the effect of providing information comparing outcomes among catchment areas allocated to different treatment groups. We estimate the effect of interventions using the following equation:

$$Y_{mt} = \alpha + \beta_1 T1_m + \beta_2 T2_m + \boldsymbol{X}'_{mt}\gamma + \theta_m + \tau_t + \varepsilon_{mt} \quad (2)$$

where $\boldsymbol{X}'_{mt}$ is a vector of community characteristics and ε_{mt} is an error term idiosyncratic to the catchment area. The interpretation of coefficients is similar to the one presented for the individual and household outcomes, and the same set of p-values with be provided.

Following work by (Mckenzie 2011) for outcomes with high autocorrelation, we run an ANCOVA specification as a robustness check, where we account for the baseline value of the outcomes considered, namely $Y_{im,t0}$. Denoting t > 0 as the follow-up periods and t0 as the baseline period, the model is:

$$Y_{im,t} = \alpha + \beta_1 T1_{im,t} + \beta_2 T2_{im,t} + \boldsymbol{X}'_{im,t0}\gamma + \delta Y_{im,t0} + \theta_m + \tau_t + \varepsilon_{im,t} \qquad (3)$$

We also study heterogeneous treatment effects. To this purpose, for each sub-group k in the variable for which we want to study heterogeneity in the effect, we define an indicator d_{ik} that takes value 1 if household/individual *i* belong to sub-group k and 0 otherwise. For binary indicators, the sub-group definition is straightforward. For non-binary dimensions of interest, we split into sub-groups based on the median of the distribution.

For individual and household outcomes, similarly to the previous section we then estimate the following model, where the impact on outcome Y_{imt} of each treatment in sub-group k as opposed to the main effect is given by β_{1k}, β_{2k}:

$$Y_{imt} = \alpha + \beta_1 T1_{im} + \beta_2 T2_{im} + \sum_k \beta_{1k} d_{ik} T1_{imt} + \sum_k \beta_{2k} d_{ik} T2_{imt} + \sum_k \eta_k d_{ik} +$$

$$+\boldsymbol{X}'_{imt}\gamma + \theta_m + \tau_t + \varepsilon_{imt} \qquad (4)$$

where $\boldsymbol{X}'_{imt}$ is a vector of controls and ε_{imt} is a residual idiosyncratic error term picking up unobserved determinants of the outcome of interest. Similarly, to the main specification, we estimate this specification both for individual post-baseline measurements (allowing estimates to change not only with respect to k, but also with respect to t), and for all post-baseline measurements (t > 0). As for our main specification, we use the ANCOVA specification as our preferred model. We present p-values of the null hypotheses that $\beta_{1k} = 0$, $\beta_{2k} = 0$, suggesting no heterogeneous effects are present.

For **community-level outcomes**, for each sub-group k in the variable for which we want to control for heterogeneity in the effect, we define an indicator d_k that takes value 1 if community m belongs to sub-group k and 0 otherwise. We can then estimate the effect of providing information using the following equation:

$$Y_{mt} = \alpha + \beta_1 T1_m + \beta_2 T2_m + \sum_k \beta_{1k} d_k T1_{mt} + \sum_k \beta_{2k} d_k T2_{mt} + \sum_k \eta_k d_k + \boldsymbol{X}'_{mt}\gamma + \theta_m + \tau_t + \varepsilon_{mt} \qquad (5)$$

where $\boldsymbol{X}'_{mt}$ is a vector of community characteristics and ε_{mt} is an error term idiosyncratic to the community. For all household/individual level models, we cluster standard errors at the catchment area level to correct for serial correlation within clusters.

This methodology allows identifying the effects of the interventions on selected outcomes, guarantee internal validity of estimates. In terms of external validity, implementing the intervention in two cities only could reduce the heterogeneity of the sample studied. However, as discussed in Section 2.1, the setting shares many common features with other cities in Asia and South-East Asia, guaranteeing a good degree of external validity.

4.2.3 Results

In this section we present the main results of our study, focusing on the main impacts on the primary and secondary outcomes detailed in the Pre-Analysis Plan. In summary, we do not find a sustained impact on the willingness-to-pay and usage of community toilets from our intervention. We do find a positive impact on the expectation of illness from unsafe sanitation practices, awareness of health risks from open defecation and the report of hand washing with soap.

We first show results pooling together all rounds of follow-up data (rapid assessment, midline and endline) and then show dynamics in impacts on primary outcomes by showing results by follow-up. We show results from our pre-specified estimation model, which controls for gender of the respondent, round dummies and strata variables (city and managed by main CT provider). Columns (1) to (4) show estimates of a linear regression model and column (6)-(9) shows estimates of an ANCOVA model for robustness. Columns (1)-(2) and (6)-(7) shows the estimates of the effect of the supply-side treatment and columns (3)-(4) and (8)-(9) show estimates of the effect of the supply-side plus information treatment.

Table 13 presents the average impacts of the intervention on primary outcomes at the household level. Please refer to Appendix B – Pre-Analysis Plan for a detailed description of how we collect and measure each of the outcome variables. The primary outcomes are: willingness to pay (*WTP*) for using the closest community toilet; willingness to pay for a hypothetical community toilet of high-standard (*Hypo WTP*); the amount donated for cleanliness of the closest community toilet (*CT clean donate*); indicator capturing whether all members of the household always use the community toilet (*Always use*); indicator capturing whether at least one member of the household uses the community toilet (*Use CT*); and, an indicator capturing whether at least one member of the household practices open defecation (*OD*). Estimating the effects both without and with an ANCOVA model, we find that, on average, the supply-side intervention alone and complemented by the information campaign had no effect on primary outcomes over the period of study.

When estimating the impact of the treatments on the main primary outcome willingness to pay separately by data collection round, we observe that the estimated coefficients vary over time. Figure 12 plots the estimated coefficients of a linear regression of willingness to pay on the supply-side treatment alone and complemented by the information campaign at baseline (BL), rapid assessment (FU1), midline (ML) and endline (EL). We include confidence intervals at the 90 percent level. We find a short-term negative effect on WTP. On average, the treatment lowered the WTP for CT usage by Rs 0.2 (17 percent higher than the control mean). This negative effect is more precisely estimated for the supply-side treatment complemented by information and remains negative for a longer period of time in this treatment arm. Notably, the negative effect is not sustained over the longer-term in either treatment arm. We hypothesize this short-term effect to the grant scheme, which was implemented right before the rapid assessment.

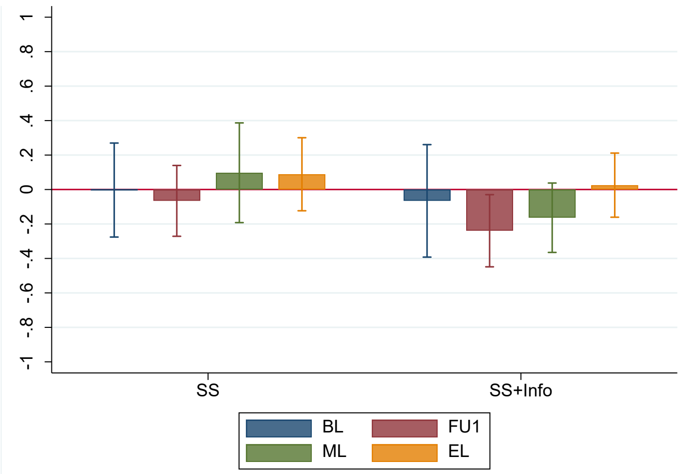
Table 13: Impact on primary outcomes

	No ANCOVA					ANCOVA				
	SS		**SS + Info**		**Obs.**	**SS**		**SS + Info**		**Obs.**
	Coef.	SE	Coef.	SE		Coef.	SE	Coef.	SE	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
WTP	0.05	0.09	-0.12	0.09	8808	0.01	0.09	-0.12	0.09	6464
Hypo WTP	-0.00	0.14	0.01	0.13	8812	-0.14	0.14	-0.04	0.12	6468
CT clean donate	-0.46	0.40	-0.07	0.42	8808	-0.44	0.41	0.17	0.43	6464
Always use CT	-0.02	0.04	-0.04	0.04	9392	-0.02	0.04	-0.04	0.03	7796
Use CT	-0.02	0.05	-0.02	0.05	9392	-0.02	0.03	-0.04	0.03	7796
OD	0.02	0.05	-0.04	0.04	9392	0.02	0.04	-0.02	0.03	7796

Note. All specifications include gender of the respondent, round dummies and strata variables as control: (i) city and (ii) managed by main provider. WTP and information about CT fees collected at the individual level for the most senior men and women present in the household. Sanitation behaviour measured for each gender group in the household. Sample includes all data rounds. Columns (5) and (10) show the number of gender-household-round observations. Sample size varies due to availability of the most senior women and men when eliciting WTP. Standard errors clustered by CT catchment area. Statistical significance denoted by * p<0.1, ** p<0.05 and *** p<0.01. See Appendix B for details about variables and measurement.

In line with a crowding-out hypothesis, slum dwellers may be less willing to privately contribute to a public good that received external aid for its operation and maintenance. The information campaign could have affected negatively the extent to which households internalize the negative externalities from inadequate sanitation behaviour. The messages provided could have increased the awareness of community toilets as a public good. For example, households may be more aware that even if they use the CTs, but their neighbour does not, the environmental quality of their community will not improve. Individuals are not able to enforce adequate sanitation behaviour of their neighbours in such setting. The awareness of such coordination problem may decrease even further their private willingness to pay and use of CTs and increase their demand for public intervention.

Figure 12: Impact on WTP by data collection round



Note. All specifications include gender of the respondent and strata variables as control: (i) city and (ii) managed by main provider. Standard errors clustered by CT catchment area. Confident intervals at the 90 percent level included.

Table 14 presents the main impact of the intervention on primary outcomes when pooling together the treatment arms, thereby increasing power. Estimating the effects both without and with an ANCOVA model while we find a significant negative short-term impact (not shown) but no effect on primary outcomes over the period of study.

Table 14: Impact on primary outcomes, pooled treatment

	No ANCOVA			**ANCOVA**		
	Coef.	SE	Obs.	Coef.	SE	Obs.
	(1)	(2)	(3)	(4)	(5)	(6)
WTP	-0.04	0.07	8808	-0.05	0.08	6464
Hypo WTP	0.00	0.12	8812	-0.09	0.11	6468
CT clean donate	-0.26	0.37	8808	-0.13	0.38	6464
Always use CT	-0.03	0.03	9392	-0.03	0.03	7796
Use CT	-0.02	0.04	9392	-0.03	0.02	7796
OD	-0.01	0.04	9392	0.00	0.03	7796

Note. All specifications include gender of the respondent, round dummies and strata variables as control: (i) city and (ii) managed by main provider. WTP and information about CT fees collected at the individual level for the most senior men and women present in the household. Sanitation behaviour measured for each gender group in the household. Sample includes all data rounds. Columns (5) and (10) show the number of gender-household-round observations. Sample size varies due to availability of the most senior women and men when eliciting WTP. Standard errors clustered by CT catchment area. Statistical significance denoted by * p<0.1, ** p<0.05 and *** p<0.01. See Appendix B for details about variables and measurement.

Table 15 presents the main impact of the intervention on secondary outcomes at the household level. The secondary outcomes include: an indicator of whether the respondent and spouse washed their hands with soap after defecating the last time (Hand-wash soap); an index capturing observed dirtiness in the household (HH dirty); an indicator of whether at least one household member had diarrhoea during the 2 weeks prior to the survey (Diarrhoea); the amount spent on CT fees during the month prior to the survey (CT spent); an index capturing positive attitudes towards CTs (CT + attitude); the difference in expected adult and child illness incidence between a community practicing open defecation and one practicing CT (Adult ill and Child ill OD-CT); the difference in expected adult and child illness incidence between a community using a dirty and a clean CT (Adult ill and Child ill dirty-clean); and an index capturing knowledge of private and public health risks from open defecation (OD health risk).

Table 15: Impact on HH secondary outcomes

	No ANCOVA					ANCOVA				
	SS		**SS + Info**		**Obs.**	**SS**		**SS + Info**		**Obs.**
	Coef.	SE	Coef.	SE		Coef.	SE	Coef.	SE	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Hand-wash soap	0.02*	0.01	0.01	0.01	4890	0.02**	0.01	0.02**	0.01	4056
HH dirty	-0.03	0.02	0.01	0.02	4890	-0.03	0.02	0.01	0.02	4056
Diarrhoea	0.02	0.01	-0.01	0.01	4890	0.02	0.02	-0.01	0.01	4056
CT spent (log)	-0.15	0.22	0.17	0.19	4485	-0.20	0.18	0.08	0.15	3677
CT + attitude	-0.01	0.02	0.01	0.01	4890	-0.01	0.01	0.01	0.01	4056
Adult ill OD-CT	-0.00	0.01	0.02	0.01	4890	-0.00	0.02	0.02	0.01	4056
Child ill OD-CT	0.07	0.04	0.03	0.03	4890	0.08	0.05	0.04	0.04	4056
Adult ill CT dirty-clean	-0.04	0.03	-0.01	0.02	4890	-0.04	0.03	-0.01	0.02	4056
Child ill CT dirty-clean	0.04	0.03	0.05**	0.02	4890	0.04	0.03	0.06**	0.03	4056
OD private health risk	0.01	0.02	0.05**	0.02	4890	0.00	0.02	0.04*	0.02	4056
OD public health risk	0.01	0.02	-0.03	0.03	4890	-0.00	0.03	-0.03	0.03	4056

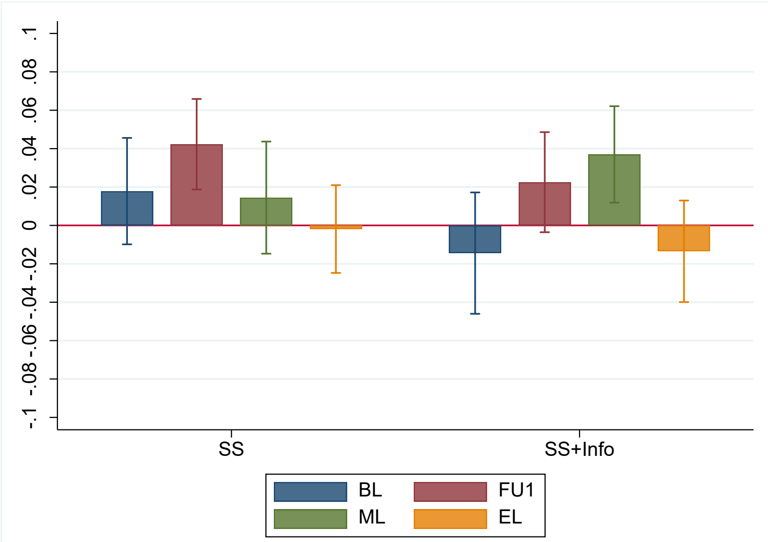
Note. All specifications include round dummies and strata variables as control: (i) city and (ii) managed by main provider. Sample includes all data rounds. Columns (5) and (10) show the number household-round observations. Standard errors clustered by CT catchment area. Statistical significance denoted by * p<0.1, ** p<0.05 and *** p<0.01. See Appendix B for details about variables and measurement.

We find that both treatment arms increased hand-washing with soap by 2 pps, translated into a 2% increase from an already high control mean (95%). The supply-side plus information treatment arm increased by 6 pps (7% higher than control mean) the difference in expected child illness between a community using a dirty CT and a community using a clean CT. The supply-side plus information treatment arm also increased by 4 pps (6% higher than control mean) the index capturing knowledge of private health risks from open defecation.

We find that the impact of the intervention on hand-washing with soap varies over time. We estimate the impact by data collection round using the same approach as discussed for our primary measure WTP. As shown in Figure 13, we find that the effect is positive and statistically significant during the rapid assessment and midline rounds, but nil at endline. The effect of the supply-side treatment arm is not statistically different from the supply-side plus information treatment arm. We therefore attribute this effect to the grant-scheme, which distributed soap and dispensers in community toilets. The positive effect at midline could also be attributed to the incentive scheme, which rewarded caretakers if their community toilets had soap available. However, the nil effect at endline suggests that any potential influence of the incentive scheme was not sustained over time.

Table 16 presents the impact of the intervention on secondary outcomes at the household level pooling together the supply-side and supply-side plus information treatment arms. We find that the intervention increased hand-washing with soap by 2 pps and increased by 5 pps the difference in expected child illness between a community using a dirty CT and a community using a clean CT. The positive impact on knowledge of private health risks from open defecation is not statistically significant anymore; evidence of the impact being driven by the information campaign.

Figure 13: Impact on hand-washing by data collection round



Note. All specifications include strata variables as control: (i) city and (ii) managed by main provider. Standard errors clustered by CT catchment area. Confident intervals at the 90 percent level included.

Table 16: Impact on secondary outcomes, pooled treatment

	No ANCOVA			**ANCOVA**		
	Coef.	SE	Obs.	Coef.	SE	Obs.
	(1)	(2)	(3)	(4)	(5)	(6)
Hand-wash soap	0.02*	0.01	4890	0.02**	0.01	4056
HH dirty	-0.01	0.02	4890	-0.01	0.02	4056
Diarrhoea	0.01	0.01	4890	0.01	0.01	4056
CT spent (log)	0.01	0.17	4485	-0.05	0.14	3677
CT + attitude	-0.00	0.01	4890	0.00	0.01	4056
Adult ill OD-CT	0.01	0.01	4890	0.01	0.01	4056
Child ill OD-CT	0.05	0.03	4890	0.06	0.04	4056
Adult ill CT dirty-clean	-0.02	0.02	4890	-0.03	0.02	4056
Child ill CT dirty-clean	0.04**	0.02	4890	0.05**	0.02	4056
OD private health risk	0.03	0.02	4890	0.02	0.02	4056
OD public health risk	-0.01	0.02	4890	-0.01	0.02	4056

Note. All specifications include gender of the respondent, round dummies and strata variables as control: (i) city and (ii) managed by main provider. WTP and information about CT fees collected at the individual level for the most senior men and women present in the household. Sanitation behaviour measured for each gender group in the household. Sample includes all data rounds. Columns (5) and (10) show the number of gender-household-round observations. Sample size varies due to availability of the most senior women and men when eliciting WTP. Standard errors clustered by CT catchment area. Statistical significance denoted by * p<0.1, ** p<0.05 and *** p<0.01. See Appendix B for details about variables and measurement.

We finally provide households with the opportunity to fill-in "voice-to-the-people" style of cards, in which they are given the possibility to report the most important concerns from a set of different topics to local public officers. Table 17 shows the impact of the treatment arms on this measure of political salience. Each outcome measures whether the household reported that specific issue as the most salient in their community. We can see that both treatment arms decreased the likelihood of a household reporting that the community is dirty as their primary concern by 7pps (18 percent lower than the control mean) and increased the likelihood of reporting that the community toilet is dirty as their primary concern by 5 pps (60 percent higher than the control mean). Pooling both treatment arms increases the precision of the estimated effects to a 5 percent significance level. Notably, there are no effects on the political salience of other issues that slum dwellers face unrelated to the intervention.

Table 17: Impact on political salience, pooled treatment arms

	Coef.	se	Obs.
	(1)	(2)	(5)
Children ill	0.01	0.02	1580
Limited water	-0.04	0.03	1580
Community dirty	-0.07**	0.04	1580
Poor roads	0.03	0.02	1580
Waste	-0.00	0.03	1580
CT dirty	0.05**	0.03	1580
No jobs	0.00	0.03	1580
Poor healthcare	-0.01	0.01	1580
Poor lighting	0.00	0.01	1580

Note. All specifications include strata variables as control: (i) city and (ii) managed by main provider. Data collected only at midline. Column (5) shows the number of households. Standard errors clustered by CT catchment area. Statistical significance denoted by * p<0.1, ** p<0.05 and *** p<0.01. See Appendix B for details about variables and measurement.

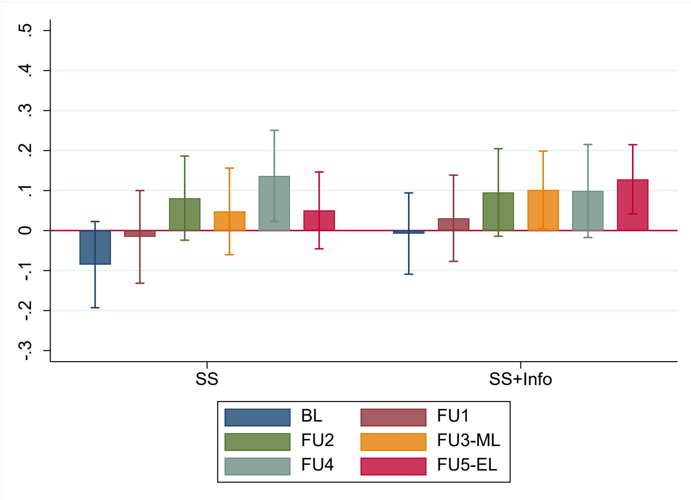
We next turn to impacts of the interventions at the community toilet level. Table 18 presents impacts on primary outcomes, including: the number of users entering the defecation cubicles in the lapse of 1 hour and the percentage that pay the fee; an index capturing observed dirtiness in the compound area and in the cubicles; an index capturing observed bad quality of the CT infrastructure; and, the count of E-coli colonies and number of bacteria types found in the CT compound and cubicles. We find that the supply-side plus information arm increased the percentage of users that pay by 9 pps. This result remains robust using the ANCOVA model. When disaggregating by gender, we find that both male and female users drive the effect, though the effect in the percentage of males paying is more precisely estimated. When estimating the impact of the intervention on the percentage of users that pay to use the CT separately by data collection round, we observe an interesting pattern, shown in Figure 14. We find nil effects during the first follow-up, which happened right after the grant scheme, but a positive effect once the incentive scheme was announced and implemented. The positive effect is sustained throughout the duration of the incentive scheme, a finding more consistent for the supply plus information treatment arm.

Table 18: Impact on CT primary outcomes

	No ANCOVA					ANCOVA				
	SS		**SS + Info**		**Obs.**	**SS**		**SS + Info**		**Obs.**
	Coef.	SE	Coef.	SE		Coef.	SE	Coef.	SE	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Users	-3.01	2.09	-0.24	2.25	542	-3.03	2.15	-0.43	2.15	528
% users pay	0.06	0.05	0.09*	0.05	542	0.04	0.05	0.09*	0.05	528
Dirty compound	-0.03	0.02	-0.02	0.02	542	-0.01	0.02	-0.02	0.02	528
Dirty cubicles	-0.04	0.03	-0.02	0.03	539	-0.01	0.02	-0.01	0.02	520
Bad quality infrastructure	-0.03	0.04	-0.04	0.04	541	-0.02	0.04	-0.05	0.04	527
E-coli count	-0.37	1.58	0.23	1.63	543	-0.61	1.57	0.25	1.64	535
Bacteria type	-0.04	0.05	0.02	0.06	543	-0.05	0.05	0.01	0.06	535

Note. Sample including all data collection rounds. All specifications include round dummies as well as strata variables as control: (i) city and (ii) managed by main provider. Standard errors clustered by CT catchment area. Statistical significance denoted by * p<0.1, ** p<0.05 and *** p<0.01. See Appendix B for details about variables and measurement.

Figure 14: Impact on % users that pay by data collection round



Note. All specifications include strata variables as control: (i) city and (ii) managed by main provider. Standard errors clustered by CT catchment area. Confident intervals at the 90 percent level included.

Based on our Theory of Change, one could expect the treatments to increase the valuation of CTs, hence the willingness to pay and the use of CTs while paying the fee. Because we do not find an increase in the willingness to pay for CTs, the rationale of our Theory of Change is not at play. On the one hand, the positive impact on the percentage of people that pay may be a result of the decrease in the number of users (denominator), though this latter effect is no statistically significant. On the other hand, the positive impact on the percentage of people that pay may be a result of an increase in the actual number of people that pay when using the CT. While the financial reward component may have incentivized caretakers to collect fees to achieve the conditionality criteria, our data available does not reveal any effects of the treatments on caretaker's behaviour. An alternative explanation for an increase in the number of users that pay is that the information campaign reinforced the need to pay fees for households that were already regular users of the CT.

Focusing on the first follow-up round, right after the grant scheme, we also find that the treatment decreased by 0.24 the number of bacteria species. Restricting the sample to the second follow-up, right after the announcement of the incentive scheme, we find that the treatment decreased the observed dirtiness of the CT compound. Perhaps the expectation of receiving an extrinsic reward improved the caretakers' effort to keep the CT clean. This was better achieved in the compound area in which the caretaker has better control over the cleanliness, since the cubicles are constantly in use and its cleanliness depends more on the users' behaviour.

Table 19 presents the impact of the intervention on secondary outcomes measured at the community toilet level, including: the percentage of time that caretakers report allocating to cleaning or supervising the cleaner and collecting fees; an indicator of whether the caretaker reports the CT is cleaned at least twice per day; an index capturing reported adequate cleaning of the CT; and, the amount donated by the caretaker to a sanitation project as a proxy for pro-social motivation. We find nil effects of the intervention on these secondary outcomes. Focusing on the first follow-up round, right after the grant scheme, we find that the treatments increased by 6pps the percentage of time that caretakers report allocating to cleaning or supervising the cleaner and collecting fees. Yet, restricting the sample to the second follow-up, right after the announcement of the incentive scheme, we find that the treatments decreased by 6pps the percentage of time that caretakers report allocating to cleaning or supervising the cleaner. The magnitude and statistical significance is higher for the estimated effect of the supply plus information treatment.

Table 19: Impact on CT secondary outcomes

	No ANCOVA					**ANCOVA**				
	SS		**SS + Info**		**Obs.**	**SS**		**SS + Info**		**Obs.**
	Coef.	SE	Coef.	SE		Coef.	SE	Coef.	SE	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
% cleaning	-0.03	0.02	-0.02	0.02	541	-0.02	0.02	-0.02	0.02	527
% collecting fees	0.00	0.02	0.02	0.02	541	0.00	0.03	0.02	0.02	527
Clean >= 2x	-0.03	0.08	-0.02	0.07	542	-0.05	0.08	-0.02	0.07	528
Adequate cleaning	-0.01	0.01	0.01	0.01	542	-0.01	0.01	0.00	0.01	528
Donation to sanitation	-2.06	3.05	-2.81	2.86	542	-3.57	2.80	-3.16	2.77	528

Note. Sample including all data collection rounds. All specifications include round dummies as well as strata variables as control: (i) city and (ii) managed by main provider. Standard errors clustered by CT catchment area. Statistical significance denoted by * p<0.1, ** p<0.05 and *** p<0.01. See Appendix B for details about variables and measurement.

When pooling together the supply-side alone and supply-side plus information treatment arms, the positive effect on the proportion of users that pay in the ANCOVA model becomes more precisely estimated (Table 20). We find nil effects on other primary and secondary outcomes.

Table 20: Impact on secondary outcomes, pooled treatment

	No ANCOVA			**ANCOVA**		
	Coef.	SE	Obs.	Coef.	SE	Obs.
	(1)	(2)	(3)	(4)	(5)	(6)
Users	-1.63	1.89	542	-1.48	1.91	528
% users pay	0.08*	0.04	542	0.08**	0.04	528
Dirty compound	-0.02	0.02	542	-0.02	0.02	528
Dirty cubicles	-0.03	0.02	539	-0.01	0.02	516
Bad quality infrastructure	-0.04	0.03	541	-0.03	0.03	527
E-coli count	-0.07	1.40	543	-0.17	1.40	535
Bacteria type	-0.01	0.05	543	-0.02	0.05	535
% cleaning	-0.02	0.01	541	-0.02	0.01	527
% collecting fees	0.01	0.02	541	0.01	0.02	527
Clean >= 2x	-0.03	0.07	542	-0.03	0.07	528
Adequate cleaning	0.00	0.01	542	-0.00	0.01	528

Note. Sample including all data collection rounds. All specifications include round dummies as well as strata variables as control: (i) city and (ii) managed by main provider. Standard errors clustered by CT catchment area. Statistical significance denoted by * p<0.1, ** p<0.05 and *** p<0.01. See Appendix B for details about variables and measurement.

4.2.4 Heterogeneity of impacts

In this section we present evidence of heterogeneous effects on primary outcomes following dimensions detailed in the Pre-Analysis Plan and pooling together all follow-up survey rounds. All heterogeneity dimensions are balanced at baseline. There is a small imbalance in community trust, but it disappears when pooling together the supply-side and supply plus info treatment arms. [14] Individual-level dimensions include gender and baseline willingness to pay. Household-level dimensions include: baseline open defecation practice, knowledge of health risks linked to open defecation, assets and trust on the community to keep the CT clean. CT-level dimensions include: baseline cleanliness and quality. Tables 14 to 22 show the estimated heterogeneous effects by stratifying the sample and column (9)-(10) show the p-value of the estimated interaction between the supply-side and supply-side plus information treatment arms, respectively, with each dimension.

Table 21 shows heterogeneous effects along the gender dimension. Columns (1) to (4) show the estimated effect of the treatment for women and column (5) to (8) for men. We can observe that both treatments decreased the likelihood of women reporting that they use the community toilet. Excluding the Endline survey round, we find that both treatments decreased men's WTP for CT usage in the short-run – the magnitude is larger during the rapid assessment. There are nil effects on women's WTP for CT usage in every follow-up round.

[14] We measure trust by asking the question: I'd like to ask you how much you trust people from your community to keep the community clean. Could you tell me whether you trust people from your community completely, somewhat, not very much or not at all to keep the community clean?

Table 21: Heterogenous impact by gender of respondent

	Women				Men				Interaction	
	SS		SS + Info		SS		SS + Info		SS	SS + I
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE	p-value	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
WTP	0.08	0.10	-0.14	0.09	-0.06	0.10	-0.09	0.12	0.14	0.62
Hypo WTP	-0.07	0.16	0.05	0.14	-0.21	0.16	-0.15	0.13	0.27	0.08
CT clean donate	-0.36	0.39	0.37	0.43	-0.52	0.53	-0.07	0.54	0.67	0.33
Always use CT	-0.02	0.04	-0.04	0.03	-0.02	0.04	-0.04	0.03	0.82	0.98
Use CT	-0.04	0.03	-0.05*	0.03	-0.01	0.04	-0.03	0.03	0.04	0.07
OD	0.02	0.04	-0.02	0.03	0.02	0.04	-0.02	0.03	0.78	0.99

Note. Sample includes all data rounds. All specifications include round dummies and strata variables as control: (i) city and (ii) managed by main provider. Standard errors clustered by CT catchment area. Roman and Wolf adjusted p-values will be available in an updated version of the report. See Appendix B for details about variables and measurement.

Table 22 explores heterogeneous effects by baseline WTP for CT use. Columns (1) to (4) show the estimated effect of the treatment for individuals with baseline WTP below the median and column (5) to (8) for individuals with WTP above the median. There is no heterogeneity along this dimension. Excluding the Endline survey round, we find a negative effect of the supply-side plus information treatment on WTP only on individuals with a low baseline WTP.

Table 22: Heterogenous impact by WTP

	Low WTP				High WTP				Interaction	
	SS		SS + Info		SS		SS + Info		SS	SS + I
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE	p-value	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
WTP	0.03	0.10	-0.13	0.10	-0.02	0.14	-0.09	0.14	0.73	0.71
Hypo WTP	-0.14	0.18	-0.02	0.13	-0.13	0.16	-0.06	0.16	0.98	0.80
CT clean donate	-0.56	0.44	0.41	0.51	-0.19	0.58	-0.30	0.49	0.51	0.23
Always use CT	-0.04	0.04	-0.04	0.04	-0.01	0.05	-0.04	0.04	0.70	0.94
Use CT	-0.02	0.04	-0.04	0.03	-0.05	0.05	-0.04	0.03	0.55	0.86
OD	0.01	0.04	-0.02	0.03	0.05	0.05	0.01	0.03	0.41	0.25
Private latrine	-0.03	0.02	0.01	0.02	-0.00	0.02	-0.00	0.02	0.34	0.54

Note. Sample includes all data rounds. All specifications include gender of the respondents, round dummies and strata variables as control: (i) city and (ii) managed by main provider. Standard errors clustered by CT catchment area. Roman and Wolf adjusted p-values will be available in an updated version of the report. See Appendix B for details about variables and measurement.

Table 23 explores heterogeneous effects by baseline practice of open defecation. Columns (1) to (4) show the estimated effect of the treatment for households not practicing OD and columns (5) to (8) for households practising OD. For households not practicing OD at baseline, the supply-side treatment decreased the WTP for a hypothetical high-quality CT. Notably, the supply-side plus information treatment arm decreased the likelihood of reporting using the CT (every household member and at least one member), but increased the likelihood of reporting having a private latrine. For households practicing OD at baseline, the supply-side plus information treatment arm decreased their WTP, though the estimated interaction term is not statistically significant (column 10).

Table 23: Heterogenous impact by OD

	No OD				OD				Interaction	
	SS		SS + Info		SS		SS + Info		SS	SS + I
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE	p-value	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
WTP	-0.00	0.10	-0.05	0.11	0.05	0.12	-0.30**	0.12	0.68	0.16
Hypo WTP	-0.34**	0.16	-0.11	0.14	0.29	0.22	0.02	0.19	0.03	0.56
CT clean donate	-0.46	0.50	-0.01	0.50	-0.17	0.45	0.28	0.54	0.63	0.50
Always use CT	-0.03	0.04	-0.07**	0.03	0.01	0.05	0.00	0.06	0.48	0.22
Use CT	-0.01	0.03	-0.04*	0.02	-0.03	0.05	-0.05	0.07	0.77	0.99
OD	0.00	0.03	-0.01	0.02	0.04	0.08	-0.03	0.09	0.55	0.86
Private latrine	-0.01	0.01	0.03*	0.02	-0.04*	0.02	-0.04	0.03	0.25	0.05

Note. Sample includes all data rounds. All specifications include gender of the respondents, round dummies and strata variables as control: (i) city and (ii) managed by main provider. Standard errors clustered by CT catchment area. Coefficients estimated through ANCOVA models. Roman and Wolf adjusted p-values will be available in an updated version of the report. See Appendix B for details about variables and measurement.

Table 24: Heterogenous impact by knowledge of OD health risks

	Not perfect knowledge				Perfect knowledge				Interaction	
	SS		SS + Info		SS		SS + Info		SS	SS + I
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE	p-value	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
WTP	-0.04	0.11	-0.15	0.11	0.06	0.10	-0.09	0.11	0.36	0.61
Hypo WTP	-0.34**	0.17	-0.15	0.16	0.10	0.20	0.09	0.16	0.07	0.26
CT clean donate	-0.34	0.53	0.07	0.51	-0.53	0.54	0.28	0.54	0.76	0.74
Always use CT	-0.02	0.04	-0.04	0.04	-0.03	0.05	-0.05	0.03	0.91	0.88
Use CT	0.01	0.04	-0.02	0.03	-0.07	0.04	-0.06**	0.03	0.05	0.31
OD	0.01	0.04	-0.03	0.03	0.03	0.05	0.00	0.03	0.46	0.19
Private latrine	-0.01	0.02	0.02	0.03	-0.02	0.02	0.01	0.02	0.58	0.81

Note. Sample includes all data rounds. All specifications include gender of the respondents, round dummies and strata variables as control: (i) city and (ii) managed by main provider. Standard errors clustered by CT catchment area. Coefficients estimated through ANCOVA models. Roman and Wolf adjusted p-values will be available in an updated version of the report. See Appendix B for details about variables and measurement.

Table 24 explores heterogeneous effects by baseline knowledge of the health risks linked to practicing open defecation. Columns (1) to (4) show the estimated effect of the treatment for households with knowledge below the median and columns (5) to (8) for households with knowledge above the median of the distribution. For households with low knowledge of OD risks, the supply-side treatment decreased the WTP for a hypothetical high-quality CT. For households with high knowledge of OD risks, the treatments decrease their likelihood of reporting using the CT (at least one household member).

Table 25 explores heterogeneous effects by baseline wealth. Columns (1) to (4) show the estimated effect of the treatment for households with an asset index below the median of the distribution and columns (5) to (8) for households with asset index above the median. For wealthier households, the supply-side plus information treatment arm decreased the likelihood of reporting using the CT (every household member and at least one member), but increased the likelihood of reporting having a private latrine. Perhaps the information campaign drove an increase in construction and adoption of private latrines in household not facing a liquidity constraint.

Table 25: Heterogenous impact by wealth

	Low asset				High asset				Interaction	
	SS		SS + Info		SS		SS + Info		SS	SS + I
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE	p-value	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
WTP	0.05	0.11	-0.08	0.12	-0.05	0.11	-0.18	0.11	0.51	0.50
Hypo WTP	-0.01	0.19	0.03	0.16	-0.30*	0.17	-0.09	0.19	0.23	0.76
CT clean donate	-0.62	0.47	0.11	0.52	-0.15	0.55	0.35	0.52	0.54	0.78
Always use CT	-0.01	0.04	0.01	0.04	-0.04	0.04	-0.12***	0.03	0.40	0.00
Use CT	-0.01	0.04	-0.01	0.03	-0.04	0.04	-0.08**	0.03	0.63	0.15
OD	0.02	0.04	-0.04	0.04	0.01	0.04	0.00	0.02	0.77	0.19
Private latrine	-0.03*	0.02	-0.02	0.02	-0.00	0.02	0.06**	0.03	0.25	0.01

Note. Sample includes all data rounds. All specifications include gender of the respondents, round dummies and strata variables as control: (i) city and (ii) managed by main provider. Standard errors clustered by CT catchment area. Coefficients estimated through ANCOVA models. Roman and Wolf adjusted p-values will be available in an updated version of the report. See Appendix B for details about variables and measurement.

Table 26: Heterogenous impact by community trust

	No trust				Trust				Interaction	
	SS		SS + Info		SS		SS + Info		SS	SS + I
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE	p-value	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
WTP	0.03	0.14	0.03	0.16	-0.04	0.09	-0.19*	0.10	0.59	0.19
Hypo WTP	-0.39*	0.20	-0.17	0.16	-0.03	0.16	0.00	0.14	0.07	0.41
CT clean donate	-0.76	0.61	0.41	0.75	-0.16	0.45	0.07	0.45	0.31	0.71
Always use CT	-0.05	0.05	-0.01	0.04	-0.01	0.04	-0.06*	0.03	0.56	0.19
Use CT	-0.06	0.06	0.02	0.04	-0.01	0.03	-0.06**	0.03	0.43	0.02
OD	0.03	0.06	-0.06	0.04	0.01	0.03	0.00	0.03	0.91	0.03
Private latrine	-0.01	0.02	-0.03	0.02	-0.02	0.02	0.02	0.02	0.58	0.04

Note. Sample includes all data rounds. All specifications include gender of the respondents, round dummies and strata variables as control: (i) city and (ii) managed by main provider. Standard errors clustered by CT catchment area. Coefficients estimated through ANCOVA models. Roman and Wolf adjusted p-values will be available in an updated version of the report. See Appendix B for details about variables and measurement.

Table 26 explores heterogeneous effects by baseline trust on the community to keep it clean. Columns (1) to (4) show the estimated effect of the treatment for households with no trust in the community and columns (5) to (8) for households with trust in the community. For households with low trust in their community, the treatments decreased their WTP for a hypothetical high-quality CT. For households with trust in the community, the supply-side plus information treatment arm decreased their WTP for CT usage and the likelihood of reporting using the CT (every household member and at least one member), but increased the likelihood of reporting having a private latrine.

Table 27 explores heterogeneous effects by cleanliness of the community toilet. Columns (1) to (4) show the estimated effect of the treatments for households located close to a CT with a dirtiness index below the median and columns (5) to (8) for households located close to a CT with a dirtiness index above the median of the distribution. For households close to a clean CT, the treatments decreased their likelihood of reporting using the CT (every household member and at least one member) and increased the likelihood of reporting practising open defecation.

Table 27: Heterogenous impact by CT cleanliness

	Clean CT				**Dirty CT**				**Interaction**	
	SS		**SS + Info**		**SS**		**SS + Info**		**SS**	**SS + I**
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE	p-value	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
WTP	0.15	0.14	-0.15	0.13	-0.15	0.09	-0.11	0.13	0.11	0.77
Hypo WTP	-0.28	0.20	-0.14	0.15	-0.16	0.20	-0.02	0.16	0.58	0.39
CT clean donate	-0.65	0.45	0.67	0.58	-0.20	0.66	-0.29	0.63	0.51	0.27
Always use CT	-0.10*	0.06	-0.09*	0.04	0.04	0.05	-0.01	0.04	0.06	0.15
Use CT	-0.10*	0.06	-0.09**	0.04	0.03	0.04	-0.00	0.03	0.05	0.06
OD	0.11*	0.06	0.03	0.04	-0.05	0.04	-0.05	0.03	0.02	0.10
Private latrine	-0.03	0.02	0.01	0.03	-0.01	0.02	0.01	0.02	0.76	0.98

Note. Sample includes all data rounds. All specifications include gender of the respondents, round dummies and strata variables as control: (i) city and (ii) managed by main provider. Standard errors clustered by CT catchment area. Coefficients estimated through ANCOVA models. Roman and Wolf adjusted p-values will be available in an updated version of the report. See Appendix B for details about variables and measurement.

Table 28: Heterogenous impact by CT quality

	Good quality CT				**Bad quality CT**				**Interaction**	
	SS		**SS + Info**		**SS**		**SS + Info**		**SS**	**SS + I**
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE	p-value	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
WTP	0.06	0.16	-0.13	0.11	-0.05	0.11	-0.07	0.13	0.55	0.89
Hypo WTP	-0.37	0.22	-0.25	0.18	0.09	0.16	0.23	0.15	0.13	0.09
CT clean donate	-0.37	0.54	0.59	0.57	-0.12	0.52	-0.15	0.56	0.89	0.28
Always use CT	-0.06	0.05	-0.06	0.04	0.01	0.05	-0.03	0.04	0.34	0.62
Use CT	-0.05	0.04	-0.03	0.04	0.00	0.05	-0.05	0.04	0.40	0.83
OD	0.07	0.05	0.02	0.05	-0.04	0.05	-0.05	0.04	0.12	0.31
Private latrine	-0.05**	0.02	-0.03	0.02	0.00	0.02	0.05*	0.03	0.08	0.04

Note. Sample includes all data rounds. All specifications include gender of the respondents, round dummies and strata variables as control: (i) city and (ii) managed by main provider. Standard errors clustered by CT catchment area. Coefficients estimated through ANCOVA models. Roman and Wolf adjusted p-values will be available in an updated version of the report. See Appendix B for details about variables and measurement.

Table 28 explores heterogeneous effects by infrastructure quality of the community toilet. Columns (1) to (4) show the estimated effect of the treatments for households located close to a CT with a bad quality index below the median and columns (5) to (8) for households located close to a CT with a bad quality index above the median of the distribution. There is no major heterogeneity along this dimension. For households close to a CT of good quality, the supply-side treatment arm decreases their likelihood of having a private latrine. In the same line, for households close to a CT of bad quality, the supply-side plus information treatment arm increases the likelihood of having a private latrine.

5. Cost analysis

The interventions are implemented by our partners FINISH and Morsel. We present in Table 29 a summary of the costing associated with each activity by providing a breakdown of costs associated with the implementation of both the supply side intervention and the information campaign. Apart from total intervention cost, we are providing unitary cost at the CT level since the supply side intervention is targeting 70 CTs, while the information campaign is targeting only 35 CTs. In addition, we provide unitary cost per eligible household and person (see eligibility criteria in Section 2), and per household and person living in proximity of the CT. In each slum, there are on average 49 eligible households (240 people), and a total of 249 households (1,261 people) living in proximity of the CT.

Table 29: Cost analysis

Intervention	Total expenditure (INR)	Cost (US$)				
		Per CT	Per eligible hh	Per eligible person	Per hh	Per person
Supply side intervention (70 CTs)						
Grant scheme	728,000.00	145.60	2.97	0.61	0.58	0.12
Incentives	269,000.00	53.80	1.10	0.22	0.22	0.04
Information campaign intervention (35 CTs)						
Design and Printing of Information Material	50,000.00	20.00	0.41	0.08	0.08	0.02
Door-to-door campaign	440,770.00	176.31	3.60	0.73	0.71	0.14
Text messages	21,661.82	8.66	0.18	0.04	0.03	0.01
Intervention management (70 CTs)						
Human resources	638,500.00	127.70	2.61	0.53	0.51	0.10
Total cost of interventions	**2,147,931.82**	**532.07**	**10.86**	**2.22**	**2.14**	**0.42**

Note. For conversion of Indian Rupees into US Dollars, we assume an exchange rate of 0.014 INR/US$. Door-to-door information campaign includes transportation costs.

6. Discussion

This study makes several contributions to the existing literature. First, it provides the first documentation of WTP for CT access, thereby complementing research on the effect of prices on adoption of preventive health technologies, by expanding the evidence base to CTs. Second, it contributes to the literature on the determinants of WTP for and adoption of sanitation technologies, by providing evidence on the role of supply-side factors or supply-side factors plus information in influencing WTP. Beyond its academic contributions, this paper informs pricing policy, helping to determine the feasibility of projects and guiding the magnitude and targeting of potential discounts and socially equitable subsidies. The findings contribute to the on-going debate of whether goods with public health implications should be priced or not.

6.1 Policy and programme relevance: evidence uptake and use

Public, non-profit and private organizations working on sanitation access are essential actors (independently as well as jointly) in creating and implementing solutions that pursue SDG #6: “Ensure availability and sustainable management of water and sanitation for all”. Therein, we assume that appropriate guidance on what constrains access to and adoption of sanitation facilities is important to address which different intervention approaches are effective and can increase short- and long-term wellbeing of beneficiaries.

Reducing disease burden in low and middle-income countries requires the identification of barriers to access and adoption of products, technologies, and solutions. Our study is particularly important for the public management of slums in major cities. According to UN-HABITAT, a prosperous city is one that provides productivity, infrastructure and development, quality of life, equity and social inclusion and environmental sustainability.

This study provides an insight on the understanding of the determinants of WTP for improved sanitation and on the understanding how life within slums is particularly important for promoting social inclusion and environmental sustainability.

To begin with, the interventions studied do not translate into an increase in the willingness to pay (WTP) for community toilet use during the whole duration of the study, suggesting that neither the supply side intervention, nor the information campaign have an impact on the WTP. If any, while the supply side intervention leads to small (insignificant) increases in WTP, we do observe short-run negative effects on WTP driven by supplementing the supply side interventions with the information campaign. The information campaign could have affected negatively the extent to which households internalize the negative externalities from inadequate sanitation behaviour. This effect is associated with a significant increase in the share of users that pay the fee to use the CT. This effect is driven by both male and female users, the effect for males is more precisely estimated. On the one hand, the expectation and actual reception of a financial reward conditional on keeping the CT clean and with available soap may have incentivized caretakers to collect resources to achieve these goals. Hence, they become less lenient with users that try to enter for free. On the other hand, the information campaign may have reinforced the need to pay fees for households that were already regular users of the CT.

Secondly, we do not observe significant effects on self-reported CT use. While a larger share of households reports exposure to the information campaign, gains in terms of knowledge are very limited. The supply-side plus information treatment arm increases knowledge of private health risks from open defecation, but none of the other dimensions is increased significantly. Nevertheless, both treatment arms increased hand-washing with soap, probably linked to soap availability in the CT driven by the supply-side intervention and an increase in hygiene awareness from the information campaign. In addition, the supply-side plus information treatment arm increased seems to have affected expectations, for instance, the difference in expected child illness between a community using a dirty CT and a community using a clean CT. Looking at CT-level outcomes, we find nil effects of the intervention on caretakers' behaviour, such the percentage of time that caretakers report allocating to cleaning or supervising the cleaner and collecting fees, whether the caretaker reports the CT is cleaned at least twice per day, and a measure of whether the caretaker reported adequate cleaning of the CT.

Even though we observe small changes in individual behaviour, we document that the treatments increase the salience of CT cleanliness when reporting issues to policymakers. When given the opportunity to fill-in "voice-to-the-people" style of cards, in which they are given the possibility to report the most important concerns from a set of different topics to local public officers, more treated households report the dirtiness of the CT as their primary concern. This result highlights the role of policymakers in guaranteeing cleanliness of CTs by solving the potential coordination problem in slums.

6.2 Challenges and lessons

A first set of challenges related to the study eligibility of households. Slum dwellers are highly mobile. Tenure security is low due to lack of property rights, mostly in informal settlements. Often slum dwellers move within the vicinity of the community, so the mechanisms we put in place allowed us to track such households. However, between surveys, many households moved away to other communities because of major public construction works done in and near their slum. We were able to replace (at random) some of these households. On the other extreme, some households stopped meeting all eligibility criteria since they constructed private toilets and reported that all household members use only this newly built latrine. While this implies that they are less likely to move in the (near) future, our interventions became less impactful for their sanitation behaviour.

A second challenge was to bring all stakeholders. The management of CTs in the two cities is highly heterogeneous. Even if a large part of CTs are owned by one NGO, each CT requires the engagement of the caretakers, the managers and the person in charge from the owning organization that is managing the area. In particular, we faced a major challenge in obtaining the final approval from all higher up management of CTs from which we would offer bundles of tickets to the households structured in the WTP game. Most stakeholders feared households getting used to not paying for the community toilet and causing an undesired behavioural change. To deal with this, our partners in India held several meetings with the stakeholders and the research team travelled to India to hold in-person meetings to present the methodology used. We carefully explained that the WTP game used as instrument was not giving away tickets for free, but people were choosing to either keep the money provided or giving it away (similar to paying) in order to get the tickets. We also invited the stakeholders to witness a WTP game in a sampled household. These activities led to the local government and NGO partner understanding the mechanism behind our study to elicit WTP for community toilets and approved the design to be implemented within their CTs.

7. Conclusions and recommendations

Understanding the suitability of shared sanitation facilities is central in a world in which the pace of urbanization in the developing world is generating an increased number of informal settlements. In this setting guaranteeing a viable standard of sanitation infrastructure is not straightforward. Interventions from external agents aimed at improving the quality of CTs, both from the supply and the demand side, can reduce rather than increase the individual evaluation for the CT use. This suggests that government- or NGO-managed interventions trying to address issues related to sanitation in urban informal settlements should consider the crowding-out effect of private contributions in response to an increased public investment. Policies targeting improvements in quality of CTs are not financially viable, and therefore cannot self-fund an improved standard in the long-run through an augmented WTP among users. Hence, improvements require external subsidies. Given the increase in handwashing with soap and our knowledge of the importance of this behaviour, these might be justified in light of externalities associated with improved hygiene.

In response to the interventions, households demand for further public intervention targeted at sanitation. This suggest that privately-oriented solutions to sanitation issues in slums suffer from a coordination problem. Policies should therefore be accompanied mechanisms favouring coordination among slum-dwellers or punishing free-riders. Information campaigns targeted at the sensitization on the role of individual behaviour within the communities have shown to have small impact on individual behaviour. An information campaign in the current setting did not lead to important changes in behaviour, while continued reminders (through voice messages) neither lead to significant changes, despite a significant number of people listening to (parts) of the messages.

Incentive mechanisms aimed at raising effort of caretakers have limited effects, probably due to the limited control of the caretakers on the behaviour of slum dwellers. Incentivizing cleanliness does not translate into improved quality, but it only creates incentives for raising the share of slum dwellers that would pay the CT fee. There is a clear demand by CT caretakers for additional support with repairs/hygiene improvements/training. The average caretaker lacked the knowledge of how to properly clean and the resources to keep the CT in a well-functioning state. However, the support provided did not lead to significant longer-term improvements in the CTs. This points to more structural problems that have to be addressed, such as continuous CT maintenance. Providing financial incentives to the caretakers is not a potential solution to these issues as our research reveals. Similarly, while the state and quality of infrastructure is frequently mentioned as a determinant for use, pouring one-off resources at the problem does not lead to a significant change in usage, which links back to the need for longer-term solutions.

Appendix A: Field notes from formative work

In the design phase of the project, we made use of qualitative methods to understand constraints to usage of community toilets in slums. In particular, we used focus group discussions (FGDs) of different user groups and open-ended interviews with CT managers, caretakers and other community officials.

The FGDs aimed at collecting opinions from different groups of potential users. The selection of respondents for group discussions was based on the following three criteria: women and men aged 18-49 years, irrespective of current marital status; permanent resident of the selected community; living in proximity of CTs. The conceptual framework and the questions were framed based on the literature review of existing research done specially in India. The following research domains have been analysed:

- Which are the priorities and demand for sanitation facilities?
- What are the perceptions and experiences of slum-dwellers regarding safety and accessibility to water, sanitation and hygiene in public/community toilets?
- What are the gender barriers faced by women and girls in accessing public/community toilet facilities, including gender-based violence?
- What are the various coping strategies adapted to overcome challenges faced due to unsafe and inadequate sanitation facilities?
- What can be learnt from the existing models for acceptability and adaptability?
- What are the possible means to increase willingness to use (and possibly pay for usage of) CTs?
- Could voucher packages for CT usage (such as monthly individual passes, or weekly family passes, etc.) meet the demand of the poorest slum dwellers?

FDGs triggered further questions that led us to go back to the field and talk to households, caretakers and managers/government officials. We hence conducted open-ended interviews in randomly selected slums and engaged with caretakers, managers of CTs in those slums and government officials in charge of CTs. The main purpose of discussing with households was to understand their knowledge and attitudes about open defecation and gain further information in shaping the BCC campaign. Whereas the discussion with caretakers was focused around improving maintenance and collecting fees/monthly passes, further improving the design of the supply-side intervention.

During the open-ended interviews with slum dwellers, we structured the interview on the following questions:

- What are the private consequences of practicing open defecation?
- What are the consequences on your family from neighbours practicing open defecation?
- What does your community believe about open defecation and its consequences?

During the open-ended interviews with the care-takers, we structured the interview on the following questions:

- What are your concerns about the service provided by the CT?
- What do you think would increase the number of users?
- How could you increase usage and profits?

- What type of cleaning do you do and what type of cleaning is needed?
- Do you need repairs? Do you need to paint walls?
- Do you have a family pass system? For how many families?
- Would you accept selling bundles of pay-as-you-go tickets to users?

During the open-ended interviews with managers and the government officials, we structured the interview on the following questions:

- What do you think would increase the number of users?
- What are your concerns about the service provided by the CT?
- How could you increase usage and profits?
- Would you accept selling bundles of pay-as-you-go tickets to users?

The FGDs were implemented in the city of Lucknow. The selection of slums was based on the CT census done by Morsel, whereby GPS readings were taken at the CT facility and the slum (organized and unorganized) in closest proximity to the CT. These were used to calculate the number of community toilets available and distances between the houses and the CT facility. FGDs were organized according to ownership status of the CTs, which are generally divided in two categories: run by the municipality (Lucknow Municipal Corporation, Lucknow Development Agency, Uttar Pradesh New and Renewable Energy) or run by NGOs. One CT from the municipal management model and two managed by an NGO were randomly selected from a list of pay-to-use CTs. The following CTs were selected:

- Vandemarket sawa market (owned by an NGO);
- Sahadatat ganj (owned by MP Fund / NGO);
- Udayganj (owned by Lucknow Municipal Corporation).

FGDs were implemented on April 19-22 2017. Table 30 presents information on each focus group in terms on group size, gender and age composition, education level of participants and location:

Table 30: Focus Group composition

Participants	Gender	Age group	Marital status	Education	Location
8	Female	20-42	3 unmarried; 5 married	6 primary; 2 high school	Joshi Colony, Vande Market Saba Market
7	Female	18-35	4 unmarried; 3 married	Primary to high school	Rampur Nahar Basti, Udayganj
3	Female	32-35	3 married	Illiterate	Community Centre, Udayganj
7	4-Female 3- Male	30-42	All married	Primary school	Nautalla, Shahadatganj
4	Female	38-46	All married	Primary school	Makeshwarpur, Saba Market
3	Female	20-25	All married	High school	Udayganj, Community Centre
21	Men	21-35	12 unmarried; 9 married	Mostly High school; 5 primary school	Rampur Nahar Basti, Udayganj

Open-ended interviews with heads of households were implemented in both Lucknow and Kanpur, Uttar Pradesh, in July 2017. We randomly selected three CTs (two in Lucknow and one in Kanpur) linked to a slum where residents are the main users. The following CTs were selected:

- Lavkush Nagar kabristan basti in (Lucknow, managed by an NGO);
- Purniya 2 crossing ke pass (Lucknow, managed by an NGO);
- Faithful Ganj, Samudayik Shauchalay (Kanpur, managed by Uttar Pradesh New and Renewable Agency).

From these toilets, we randomly selected nine households located in the nearby slum that do not own a private toilet. More than half of the sample corresponds to female respondents.

In addition, we conducted open-ended interviews to caretakers from randomly selected CTs in Lucknow and Kanpur. From a list of pay-to-use CTs, we randomly selected 5 CTs in Lucknow and 3 CTs in Kanpur. The following CTs were selected:

- Purniya 2 crossing ke pass (Lucknow, managed by an NGO);
- Lavkush Nagar kabristan basti (Lucknow, managed by an NGO);
- Lavkush Nagar basti (Lucknow, managed by an NGO);
- Bilegarad (Lucknow, privately managed);
- Katra paltan tadikhana (Lucknow, managed by District Urban Development Agency);
- Faithful Ganj, Samudayik Shauchalay (Kanpur, managed by an NGO);
- Faithful Ganj (Kanpur, managed by an NGO);
- Shiv Katra, Sarvajanik Shauchalay (Kanpur, managed by Kanpur Slum Dwellers Federation SUDA).

Finally, we conducted open-ended interviews with the managers of the most active NGO for Kanpur and Lucknow area. These interviews were highly relevant due to the large percentage of CTs managed by this specific NGO in our study area.

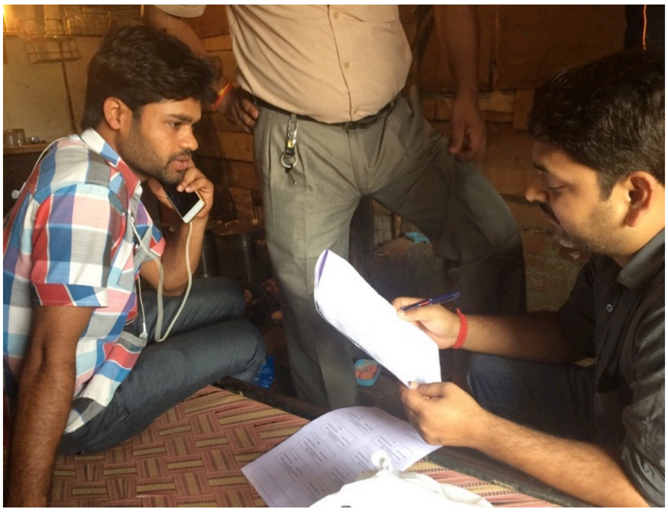
The results from FDGs aimed to understand sanitation behaviour, barriers faced and coping mechanisms adopted by slum dwellers without access to individual household latrines (IHHLs), particularly disadvantaged females. We find that, although CTs are largely used by slum dwellers, they have not ended open defecation because of several barriers that individuals face. Barriers identified mostly mirror those found in our literature review. Respondents highlighted distance, time spent queuing, location (i.e. close to market areas receiving strange people outside of community), inadequate maintenance, fees, female's safety and structural aspects (i.e. unfit for young children and elderly) as main barriers. Even within relatively shorts distances from CTs, many households with no IHHLs were using makeshift toilets connected to open ditches instead of CTs.

All respondents agreed on the necessity to improve CTs to suit their needs. Men suggested increasing the number of CTs and facilitating a community ownership for its maintenance. Women focused largely on the infrastructural and maintenance aspect of the CTs. All the participants belonged to low-income groups; hence, pay per use was not considered feasible and fair, as they stressed having the natural option of going to the open. They considered fairer to pay a fixed amount up-front in order to contribute towards cleaning and maintenance of the CT.

All respondents agreed on the necessity to improve CTs to suit their needs. Men suggested increasing the number of CTs and facilitating a community ownership for its maintenance. Women focused largely on the infrastructural and maintenance aspect of the CTs. All the participants belonged to low-income groups; hence, pay per use was not considered feasible and fair, as they stressed having the natural option of going to the open. They considered fairer to pay a fixed amount up-front in order to contribute towards cleaning and maintenance of the CT.

The aim of additional open-ended interviews with (potential) users was to dig deeper into the status-quo knowledge of open defecation risks, challenges that CTs face and opportunities to increase usage and willingness to pay for CTs. Figure 15 shows a moment from an interview in Lucknow.

Figure 15: Open-ended interview in Lucknow



Note. The figure shows a moment during the open-ended interview with a slum dweller in Lucknow.

The interviews highlighted (and hence support conclusions from the FGDs) that slum dwellers have knowledge about the public health risks of open defecation. However, there is still scope to improve the understanding of the association between causes and consequences of open defecation and the benefits of using CTs. Also, there is only very limited understanding of the need and importance of paying fees to ensure adequate maintenance and cleanliness. Even if slum dwellers understand the risks of inadequate sanitation behaviour, the benefits offered by using bad quality CTs are not as high as to outweigh the costs of using them (i.e. monetary and transaction). These results support the introduction of information dissemination efforts, accompanied by improvements in the quality of CTs, to guarantee that slum dwellers can make informed decisions related to their sanitation practices and increase their willingness to pay given the quality of the service provided.

Interviews with CT managers and care-takers revealed that this low willingness to pay results in significant challenges to maintain CTs, in most cases leading to low hygienic and safety standards of the facilities. This in turn leads to lower usage and lower willingness to pay, meaning that there is a vicious cycle. Without adequate resources to break this, CT managers and caretakers have little means and incentives to exert effort in order to improve conditions.

Given these findings, it was perceived that providing an initial maintenance support and incentivizing care-takers to maximize pay-to-use usage and to further improve the quality of service can lead to improvements in the supply side that could further increase usage and willingness to pay – breaking the vicious cycle and providing revenues that can be re-invested.

Maintenance needs faced by CTs vary greatly, ranging from deep cleaning (i.e. emptying septic tanks and unclogging toilets) to infrastructure repairs like water connections, doors and locks and painting walls. This suggested that improvements in supply side should not focus on one specific intervention, such as improving the frequency of cleaning (as suggested in the original research proposal), but need to be addressing the individual needs of each CT. An increase in usage and thus revenues may be achieved through informing households about the need to pay fees in exchange of a service and the re-investment of revenue in different maintenance activities suitable to each CT's needs.

Appendix B: Pre-analysis Plan

See attachment "*Pre-Analysis Plan: Community toilet use in slums - willingness to pay and the role of informational and supply side constraints*". The pre-analysis plan has also been registered with the AEA RCT Registry (AEARCTR-0003087).

References

Ahuja, Amrita, Michael Kremer, and Alix Peterson Zwane. 2010. “Providing Safe Water: Evidence from Randomized Evaluations.” https://doi.org/10.1146/annurev.resource.012809.103919.

Armand, Alex, et al. "Do public health interventions crowd out private health investments? Malaria control policies in Eritrea." Labour Economics 45 (2017): 107-115.

Ashraf, Nava, Oriana Bandiera, and B. Kelsey Jack. 2014. “No Margin, No Mission? A Field Experiment on Incentives for Public Service Delivery No Margin, No Mission? A Field Experiment on Incentives for Public Service Delivery *.” *Journal of Public Economics* 120: 1–17. https://doi.org/10.1016/j.jpubeco.2014.06.014.

Ashraf, Nava, Oriana Bandiera, and Scott Lee. 2014. “Awards Unbundled: Evidence from a Natural Field Experiment.” *Journal of Economic Behaviour & Organization* 100: 44–63. https://doi.org/10.1016/j.jebo.2014.01.001.

Ashraf, Nava, James Berry, and Jesse M. Shapiro. 2010. “Can Higher Prices Stimulate Product Use? Evidence from a Field Experiment in Zambia.” *American Economic Review* 100: 2383–2413. https://doi.org/10.1257/aer.100.5.2383.

Attanasio, Orazio, Luca Pellerano, and Sandra Polanía Reyes. 2009. “Building Trust? Conditional Cash Transfer Programmes and Social Capital.” *Fiscal Studies* 30 (2): 139–77. https://doi.org/10.1111/j.1475-5890.2009.00092.x.

Basinga, Paulin, Paul J. Gertler, Agnes Binagwaho, Agnes Lb Soucat, Jennifer Sturdy, and Christel Mj Vermeersch. 2011. “Effect on Maternal and Child Health Services in Rwanda of Payment to Primary Health-Care Providers for Performance: An Impact Evaluation.” *The Lancet* 377 (9775): 1421–28. https://doi.org/10.1016/S0140-6736(11)60177-3.

Benabou, Roland, and Jean Tirole. 2002. “Self-Confidence and Personal Motivation.” *Quarterly Journal of Economics*, no. 1: 871–914. https://doi.org/10.1007/s13398-014-0173-7.2.

———. 2003. “Intrinsic and Extrinsic Motivation ´.” *Review of Economic Studies* 70: 489–520. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.89.1235&rep=rep1&type=pdf.

Benn, Hilary. 2006. “Meeting Our Promises in Poor Countries.”

Bennett, D. (2012). Does clean water make you dirty? water supply and sanitation in the Philippines. Journal of Human Resources 47(1), 146–173.

BenYishay, Ariel, Andrew Fraker, Raymond Guiteras, Giordano Palloni, Neil Buddy Shah, Stuart Shirrell, and Paul Wang. 2016. “Microcredit and Willingness to Pay for Environmental Quality: Evidence from a Randomized-Controlled Trial of Finance for Sanitation in Rural Cambodia.” PWP-MPRC-2016-003. Maryland.

Berry, James, Greg Fischer, and Raymond Guiteras. 2012. “Eliciting and Utilizing Willingness to Pay: Evidence from Field Trials in Northern Ghana.” London. http://eprints.lse.ac.uk/47913/.

Besley, Timothy, and Maitreesh Ghatak. 2005. “Competition and Incentives with Motivated Agents.” *American Economic Review* 95 (3): 616–36. http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=4D947633FB3D76F4A2B36B63BE23E127?doi=10.1.1.487.9260&rep=rep1&type=pdf.

Biran, A., M. W. Jenkins, P. Dabrase, and I. Bhagwat. 2011. “Patterns and Determinants of Communal Latrine Usage in Urban Poverty Pockets in Bhopal, India.” *Tropical Medicine & International Health: TM & IH* 16 (7): 854–62. https://doi.org/10.1111/j.1365-3156.2011.02764.x.

Cairncross, S., K. Shordt, S. Zacharia, and B. K. Govindan. 2005. “What Causes Sustainable Changes in Hygienic Behaviour? A Cross-Sectional Study from Kerala, India.” *Social Science & Medicine* 61 (10): 2212–20.

Cameron, Lisa, Manisha Shah, and Susan Olivia. 2013. “Impact Evaluation of a Large-Scale Rural Sanitation Project in Indonesia.” 6360. Policy Research Working Paper. Washington D.C. https://www.wsp.org/sites/wsp.org/files/publications/WSP-Indonesia-Impact-Evaluation-Large-Scale-Rural-Sanitation.pdf.

Caruso, Bethany A., Robert Dreibelbis, Emily Awino Ogutu, and Richard Rheingans. 2014. “If You Build It Will They Come? Factors Influencing Rural Primary Pupils’ Urination and Defecation Practices at School in Western Kenya.” *Journal of Water, Sanitation and Hygiene for Development* 4 (4): 642. https://doi.org/10.2166/washdev.2014.028.

Casey, Katherine, Rachel Glennerster, and Edward Miguel. 2012. “Reshaping Institutions: Evidence and Impacts Using a Pre-Analysis Plan.” *The Quarterly Journal of Economics*, 1755–1812. https://doi.org/10.1093/qje/qje027.

Cutler, D. M. and J. Gruber (1996). Does public insurance crowd out private insurance? The Quarterly Journal of Economics 111(2), 391–430.

Das, J., S. Dercon, J. Habyarimana, P. Krishnan, K. Muralidharan, and V. Sundararaman (2011). School Inputs, Household Substitution, and Test Scores. National Bureau of Economic Research (NBER) Working Paper 16830.

Dupas, Pascaline, Vivian Hoffmann, Michael Kremer, and Alix Peterson Zwane. 2016. “Targeting Health Subsidies through a Non-Price Mechanism: A Randomized Controlled Trial in Kenya.” *Science* 353 (6302): 889–895. https://doi.org/10.1126/science.aaf6288.

Evans, W D, S K Pattanayak, S Young, J Buszin, S Rai, and Jasmine Wallace Bihm. 2014. “Social Marketing of Water and Sanitation Products: A Systematic Review of Peer-Reviewed Literature.” *Social Science & Medicine* 110: 18–25. https://doi.org/10.1016/j.socscimed.2014.03.011.

Garn, Joshua, Bethany Caruso, Carolyn Drews-Botsch, Michael Kramer, Babette Brumback, Richard Rheingans, and Matthew Freeman. 2014. “Factors Associated With Pupil Toilet Use in Kenyan Primary Schools.” *International Journal of Environmental Research and Public Health* 11 (9): 9694–9711. https://doi.org/10.3390/ijerph110909694.

Garn, Joshua V., Gloria D. Sclar, Matthew C. Freeman, Gauthami Penakalapati, Kelly T. Alexander, Patrick Brooks, Eva A. Rehfuess, Sophie Boisson, Kate O. Medlicott, and Thomas F. Clasen. 2016. “The Impact of Sanitation Interventions on Latrine Coverage and Latrine Use: A Systematic Review and Meta-Analysis.” *International Journal of Hygiene and Environmental Health*. https://doi.org/10.1016/j.ijheh.2016.10.001.

Gertler, Paul, Manisha Shah, Maria Laura Alzua, Lisa Cameron, Sebastian Martinez, and Sumeet Patil. 2015. “HOW DOES HEALTH PROMOTION WORK? EVIDENCE FROM THE DIRTY BUSINESS OF ELIMINATING OPEN DEFECATION.” 20997. NBER Working Paper Series. Cambridge, MA. http://www.nber.org/papers/w20997.

Guiteras, Raymond P., James Levinsohn, and Ahmed Mushfiq Mobarak. 2015. “Encouraging Sanitation Investment in the Developing World: A Cluster-Randomized Trial.” *Science* 348 (6237): 903–6. https://doi.org/10.1126/science.aaa5727.

Hastings, Gerard. 2007. *Social Marketing : Why Should the Devil Have All the Best Tunes?* Elsevier/Butterworth-Heinemann. https://books.google.co.uk/books/about/Social_Marketing.html?id=MbvqeuzyJuEC&redir_esc=y.

Hathi, Payal, Sabrina Haque, Lovey Pant, Diane Coffey, Dean Spears, and Coffey@utexas Edu. 2017. “Place and Child Health: The Interaction of Population Density and Sanitation in Developing Countries.” *Demography* 54: 337–60. https://doi.org/10.1007/s13524-016-0538-y.

Indian Ministry of Home Affairs. 2011. “Census - Slum Households by Availability of Type of Latrine Facility.” Delhi. 2011. http://www.censusindia.gov.in/2011census/Hlo-series/HH08.html.

Jalan, Jyotsna, and E. Somanathan. 2008. “The Importance of Being Informed: Experimental Evidence on Demand for Environmental Quality.” *Journal of Development Economics* 87 (1): 14–28. https://doi.org/10.1016/j.jdeveco.2007.10.002.

JMP WHO-UNICEF. 2017. “Sanitation | JMP.” Monitoring. 2017. https://washdata.org/monitoring/sanitation.

Kremer, Michael, Jessica Leino, Edward Miguel, and Alix Peterson Zwane. 2009. “Spring Cleaning: Rural Water Impacts, Valuation and Property Rights Institutions.” 15280. NBER Working Paper Series. Cambridge, MA. http://www.nber.org/papers/w15280.

Luoto, Jill E., David I. Levine, and Jeff Albert. 2011. “Information and Persuasion.” 2011. http://www.rand.org/pubs/working_papers/WR885.html.

Marx, Benjamin, Thomas Stoker, and Tavneet Suri. 2013. “The Economics of Slums in the Developing World.” *Journal of Economic Perspectives* 27 (4): 187–210. https://doi.org/10.1257/jep.27.4.187.

Mazeau, A P, N Wellington, S Drabble, F Asante, and D Awantungo. 2014. “Bringing Toilets Back to Kumasi’s Compound Houses: Landlord and Tenant Behaviours and Motivators.” In *Sustainable Water and Sanitation for All in a Fast Changing World*. Hanoi: 37th WEDC International Conference.

McNeil Jr., Donal J. 2005. “A Program to Fight Malaria in Africa Draw Questions.” *New York Times*, June 11, 2005.

Minh, Hoang Van, Hung Nguyen-Viet, Nguyen Hoang Thanh, and Jui-Chen Yang. 2013. “Assessing Willingness to Pay for Improved Sanitation in Rural Vietnam.” *Environmental Health and Preventive Medicine* 18 (4): 275–84. https://doi.org/10.1007/s12199-012-0317-3.

Muralidharan, Karthik, and Venkatesh Sundararaman. 2011. “Teacher Performance Pay: Experimental Evidence from India.” *Journal of Political Economy* 119 (1): 39–77. https://doi.org/10.1086/659655.

Neal, Derek. 2011. “The Design of Performance Pay in Education.” In *Handbook of the Economics of Education*, 4th ed., 495–548. Chicago: Elsevier B.V. https://doi.org/B978-0-444-53444-6.00006-7.

Ness, S.J. 2015. “Evaluation of School VIP Latrines and User Preferences and Motivations for Adopting Communal Sanitation Technologies in Zwedru, Liberia.” South Florida.

Null, Clair, Jorge Garcia Hombrados, Michael Kremer, and Edward Miguel. 2012. “WILLINGNESS TO PAY FOR CLEANER WATER IN LESS DEVELOPED COUNTRIES: SYSTEMATIC REVIEW OF EXPERIMENTAL EVIDENCE.” Washington D.C.

Oster, Sharon M. 1995. *Strategic Management for Non-Profit Organizations: Theory and Cases*. Oxford: Oxford University Press.

Peltzman, S. (1973). The Effect of Government Subsidies-in-Kind on Private Expenditures: The Case of Higher Education. Journal of Political Economy 81(1), 1–27.

Pickering, Amy J, Habiba Djebbari, Carolina Lopez, Massa Coulibaly, and Maria Laura Alzua. 2015. “Effect of a Community-Led Sanitation Intervention on Child Diarrhoea and Child Growth in Rural Mali: A Cluster-Randomised Controlled Trial.” *The Lancet Global Health* 3 (11): e701–11. https://doi.org/10.1016/S2214-109X(15)00144-8.

Roy, A.D. 1951. “Some Thoughts on the Distribution of Earnings.” *Oxford Economic Papers* 3 (2): 135–46.

Spears, Dean. 2014. “Decision Costs and Price Sensitivity: Field Experimental Evidence from India.” *Journal of Economic Behaviour & Organization* 97 (January): 169–84. https://doi.org/10.1016/j.jebo.2013.06.012.

Wagner, Edmund Glenn, and Joseph Nicker Lanoix. 1958. “Excreta Disposal for Rural Areas and Small Communities.” 39. Monograph Series. Geneva: World Health Organization.

Wilson, J.M., and G. N. Chandler. 1993. “Sustained Improvements in Hygiene Behaviour amongst Village Women in Lombok, Indonesia.” *Transactions of the Royal Society of Tropical Medicine and Hygiene* 87 (6): 615–16.

Wood, W., L. Tam, and M. G. Witt. 2005. “Changing Circumstances, Disrupting Habits.” *Journal of Personality and Social Psychology* 88: 918–33.

Yeatman, Sara, and Jenny Trinitapoli. 2011. “Best-Friend Reports: A Tool for Measuring the Prevalence of Sensitive Behaviours.” *American Journal of Public Health* 101 (9): 1666–67. https://doi.org/10.2105/AJPH.2011.300194.

Other publications in the 3ie Impact Evaluation Report Series

The following reports are available from http://3ieimpact.org/evidence-hub/publications/impact-evaluations

Impacts, maintenance and sustainability of irrigation in Rwanda, 3ie Impact Evaluation Report 112. Byiringo, E, Jones M, Kondylis F, Loeser J, Magruder, J and Ndahimana, C, 2020.

Continuous Emissions Monitoring Systems (CEMS) in India, 3ie Impact Evaluation Report 111. Greenstone, M, Pande, R, Ryan, N and Sudarshan A, 2020.

Evaluating the impacts of the Dar es Salaam Bus Rapid Transit System, 3ie Impact Evaluation Report 110. Morten, M, Bryan, G, Siddiqi, B, Balboni, C, 2020.

Access to safe drinking water: experimental evidence from new water sources in Bangladesh, 3ie Impact Evaluation Report 109. Cocciolo, S, Ghisolfi, S, Habib, A, Rashid, SMA and Tompsett, A, 2020.

Impact of alternate wetting and drying on farm incomes and water savings in Bangladesh, 3ie Impact Evaluation Report 108. Chakravorty, U, Dar, MH, Emerick, K, 2020.

The effects of vouchers for essential household items on child health, mental health, resilience and social cohesion among internally displaced persons in the Democratic Republic of Congo, 3ie Impact Evaluation Report 107. Quattrochi, J, Bisimwa, G, Thompson, T, van der Windt, P and Voors, M, 2020.

Measuring impacts of conservation interventions on human well-being and the environment in Northern Cambodia, 3ie Impact Evaluation Report 106. Clements, T, Neang, M, Milner-Gulland, EJ and Travers, H, 2020.

The 5 Star Toilet Campaign: improving toilet use in rural Gujarat, 3ie Impact Evaluation Report 105. Chauhan, K, Schmidt, WP, Aunger, R, Gopalan, B, Saxena, D, Yashobant, S, Patwardhan, V, Bhavsar, P, Mavalankar, D and Curtis, V, 2020.

How education about maternal health risk can change the gender gap in the demand for family planning in Zambia, 3ie Impact Evaluation Report 104. Ashraf, N, Field, E, Voena, A and Ziparo, R, 2019.

In search of the holy grail: can unconditional cash transfers graduate households out of poverty in Zambia?, Impact Evaluation Report 103. Handa, S, Tembo, G, Natali, L, Angeles, G and Spektor, G, 2019.

Increasing HIV self-testing and linkage to care for partners of women in antenatal care in Uganda, Impact Evaluation Report 102. Wanyenze, R, Buregyeya, E, Matovu, J, Kisa, R, Kagaayi, J, Vrana-Diaz, C, Malek, A, Musoke, W, Chemusto, H, Mukama, S and Korte, J, 2019.

Improving the quality of care for children with acute malnutrition in Uganda, 3ie Impact Evaluation Report 101. Marzia, L, Wanzira, H, Lochoro, P and Putoto, G, 2019.

Impacts of increasing community resilience through humanitarian aid in Pakistan, 3ie Impact Evaluation Report 100. Avdeenko, A and Frölich, M, 2019.

Impacts of community monitoring of socio-environmental liabilities in the Ecuadorian and Peruvian Amazon, 3ie Impact Evaluation Report 99. Pellegrini, L, 2019.

Increasing HIV testing demand among Kenyan truck drivers and female sex workers, 3ie Impact Evaluation Report 98. Kelvin, E, George, G, Mwai, E, Kinyanjui, S, Inoti, S, Chetty, T, Strauss, M, Romo, M, Oruko, F, Odhiambo J, Nyaga, E, Mantell, J and Govender, K, 2019.

Impacts of community stakeholder engagement interventions in Ugandan oil extractives, 3ie Impact Evaluation Report 97. Parker, R, Coleman, E, Manyindo, J, Schultz, B and Mukuru, E, 2019.

The impacts of formal registration of businesses in Malawi, 3ie Impact Evaluation Report 96. Campos, F, Goldstein, M and McKenzie, D, 2019.

Unpacking the determinants of entrepreneurship development and economic empowerment for women in Kenya, 3ie Impact Evaluation Report 95. McKenzie, D, Puerto, S and Odhiambo, F, 2019.

Impacts of key provisions in Ghana's Petroleum Revenue Management Act, 3ie Impact Evaluation Report 94. Edjekumhene, I, Voors, M, Lujala, P, Brunnschweiler, C, Owusu, CK and Nyamekye, A, 2019.

Using information to break the political resource curse in natural gas management in Mozambique, 3ie Impact Evaluation Report 93. Armand, A, Costa, AI, Coutts, A, Vicente, P and Vilela, I, 2019.

Harnessing transparency initiatives to improve India's environmental clearance process for the mineral mining sector, 3ie Impact Evaluation Report 92. Pande, R and Sudarshan, A, 2019.

Impacts of removing user fees for maternal health services on universal health coverage in Kenya, 3ie Impact Evaluation Report 91. Abuya, T, Dennis, M, Matanda, D, Obare, F and Bellows, B, 2018.

Impact of voice reminders to reinforce harvest aggregation services training for farmers in Mali, 3ie Impact Evaluation Report 90. Osei, RD, Dzanku, FM, Osei-Akoto, I, Asante, F, Hodey, LS, Adu, PN, Adu-Ababio, K and Coulibaly, M, 2018.

Impacts of Breakthrough's school-based gender attitude change programme in Haryana, India, 3ie Impact Evaluation Report 89. Jayachandran, S, Jain, T and Dhar, D, 2018.

Hotspot interventions at scale: the effects of policing and city services on crime in Bogotá, Colombia, 3ie Impact Evaluation Report 88. Blattman, C, Green, D, Ortega, D and Tobón, S, 2018.

Impact evaluation of the Philippine Special Program for Employment of Students, 3ie Impact Evaluation Report 87. Beam, E, Linden, L, Quimbo, S and Richmond, H, 2018.

Community-based distribution of oral HIV self-testing kits: experimental evidence from Zambia, 3ie Impact Evaluation Report 86. Hensen, B, Ayles, H, Mulubwa, C, Floyd, S, Schaap, A, Chiti, B, Phiri, M, Mwenge, L, Simwinga, M, Fidler S, Hayes, R, Bond, V and Mwinga, A, 2018.

Evaluating the economic impacts of rural banking: experimental evidence from southern India, 3ie Impact Evaluation Report 85. Field, E and Pande, R, 2018.

Direct provision versus facility collection of HIV tests: impacts of self-testing among female sex workers in Uganda. 3ie Impact Evaluation Report 84. Ortblad, K, Musoke, DK, Ngabirano, T, Oldenburg, C and Bärnighausen, T, 2018.

Increasing female sex worker HIV testing: effects of peer educators and HIV self-tests in Zambia, 3ie Impact Evaluation Report 83. Chanda, MM, Ortblad, KF, Mwale, M, Chongo, S, Kanchele, C, Kamungoma, N, Fullem, A, Bärnighausen, T and Oldenburg, CE, 2018.

Community delivery of antiretroviral drugs: a non-inferiority matched-pair pragmatic cluster-randomized trial in Dar es Salaam, Tanzania, 3ie Impact Evaluation Report 82. Francis, JM, Geldsetzer, P, Asmus, G, Ulenga, N, Ambikapathi, R, Sando, D, Fawzi, W and Bärnighausen, T, 2018.

Nourishing the future: targeting infants and their caregivers to reduce undernutrition in rural China, 3ie Impact Evaluation Report 81. Cai, J, Luo, R, Li, H, Lien, J, Medina, A, Zhou, H and Zhang, L, 2018.

Impacts of the World Food Programme's interventions to treat malnutrition in Niger. 3ie Impact Evaluation Report 80. Brück, T, Ferguson, NTN, Ouédraogo, J and Ziegelhöfer, Z, 2018.

Impact evaluation of the World Food Programme's moderate acute malnutrition treatment and prevention programmes in Sudan. 3ie Impact Evaluation Report 79. Guevarra, E, Mandalazi, E, Balegamire, S, Albrektsen, K, Sadler, K, Abdelsalam, K, Urrea, G and Alawad, S, 2018.

Impact evaluation of WFP's programs targeting moderate acute malnutrition in humanitarian situations in Chad. 3ie Impact Evaluation Report 78. Saboya, M, Rudiger, J, Frize, J, Ruegenberg, D, Rodríguez Seco, A and McMillon, C, 2018.

Improving midday meal delivery and encouraging micronutrient fortification among children in India, 3ie Impact Evaluation Report 77. Shastry, GK, Berry, J, Mukherjee, P, Mehta, S and Ruebeck, H, 2018.

Evaluation of infant development centres: an early years intervention in Colombia, 3ie Impact Evaluation Report 76. Andrew, A, Attanasio, O, Bernal, R, Cordona, L, Krutikova, S, Heredia, DM, Medina, C, Peña, X, Rubio-Codina, M and Vera-Hernandez, M, 2018.

Can the wounds of war be healed? Experimental evidence on reconciliation in Sierra Leone. 3ie Impact Evaluation Report 75. Cilliers, J, Dube, O and Siddiqi, B, 2018.

Impact evaluation of the Menabe and Melaky development programme in Madagascar, 3ie Impact Evaluation Report 74. Ring, H, Morey, M, Kavanagh, E, Kamto, K, McCarthy, N, Brubaker, J and Rakotondrafara, C, 2018.

Impact evaluation of the Smallholder Dairy Commercialization Programme in Kenya, 3ie Impact Evaluation Report 73. Bonilla, J, McCarthy, N, Mugatha, S, Rai, N, Coombes, A and Brubaker, J, 2018.

Impact and adoption of risk-reducing drought-tolerant rice in India, 3ie Impact Evaluation Report 72. Yamano, T, Dar, MH, Panda, A, Gupta, I, Malabayabas, ML and Kelly, E, 2018.

Poverty and empowerment impacts of the Bihar Rural Livelihoods Project in India, 3ie Impact Evaluation Report 71. Hoffmann, V, Rao, V, Datta, U, Sanyal, P, Surendra, V and Majumdar, S 2018.

How should Tanzania use its natural gas? Citizens' views from a nationwide Deliberative Poll, 3ie Impact Evaluation Report 70. Birdsall, N, Fishkin, J, Haqqi, F, Kinyondo, A, Moyo, M, Richmond, J and Sandefur, J, 2018.

Impact evaluation of the conditional cash transfer program for secondary school attendance in Macedonia, 3ie Impact Evaluation Report 69. Armand, A and Carneiro, P, 2018.

Age at marriage, women's education, and mother and child outcomes in Bangladesh, 3ie Impact Evaluation Report 68. Field, E, Glennerster, R, Nazneen, S, Pimkina, S, Sen, I and Buchmann, N, 2018.

Evaluating agricultural information dissemination in western Kenya, 3ie Impact Evaluation Report 67. Fabregas, R, Kremer, M, Robinson, J and Schilbach, F, 2017.

General equilibrium impact assessment of the Productive Safety Net Program in Ethiopia, 3ie Impact Evaluation Report 66. Filipski, M, Taylor, JE, Abegaz, GA, Ferede, T, Taffesse, AS and Diao, X, 2017.

Impact of the Uddeepan programme on child health and nutrition in India, 3ie Impact Evaluation Report 65. Kochar, A, Sharma, A and Sharma, A, 2017.

Evaluating oral HIV self-testing to increase HIV testing uptake among truck drivers in Kenya, 3ie Impact Evaluation Report 64. Kelvin, EA, Mwai, E, Romo, ML, George, G, Govender, K, Mantell, JE, Strauss, M, Nyaga, EN and Odhiambo, JO, 2017.

Integration of EPI and paediatric HIV services for improved ART initiation in Zimbabwe, 3ie Impact Evaluation Report 63. Prescott, M, Boeke, C, Gotora, T, Mafaune, HW, Motsi, W, Graves, J, Mangwiro, A and McCarthy, E, 2017.

Increasing male partner HIV testing using self-test kits in Kenya, 3ie Impact Evaluation Report 62. Gichangi, A, Korte, JE, Wambua, J, Vrana, C and Stevens, D, 2017.

Evaluating the impact of community health worker integration into prevention of mother-to-child transmission of HIV services in Tanzania, 3ie Impact Evaluation Report 61. Nance, N, McCoy, S, Ngilangwa, D, Masanja, J, Njau, P and Noronha, R, 2017.

Using HIV self-testing to promote male partner and couples testing in Kenya, 3ie Impact Evaluation Report 60. Thirumurthy, H, Omanga, E, Obonyo, B, Masters, S and Agot, K, 2017.

Increasing male partner HIV self-testing at antenatal care clinics in Kenya, 3ie Impact Evaluation Report 59. Gichangi, A, Korte, JE, Wambua, J, Vrana, C and Stevens, D, 2017.

Impact of free availability of public childcare on labour supply and child development in Brazil, 3ie Impact Evaluation Report 58. Attanasio, O, Paes de Barros, R, Carneiro, P, Evans, D, Lima, L, Olinto, P and Schady, N, 2017.

Estimating the effects of a low-cost early stimulation and parenting education programme in Mexico, 3ie Impact Evaluation Report 57. Cardenas, S, Evans, D and Holland, P, 2017.

The Better Obstetrics in Rural Nigeria study: an impact evaluation of the Nigerian Midwives Service Scheme, 3ie Impact Evaluation Report 56. Okeke, E, Glick, P, Abubakar, IS, Chari, AV, Pitchforth, E, Exley, J, Bashir, U, Setodji, C, Gu, K and Onwujekwe, O, 2017.

*The Productive Safety Net Programme in Ethiopia: impacts on children's schooling, labour and nutritional statu*s, 3ie Impact Evaluation Report 55. Berhane, G, Hoddinott, J, Kumar, N and Margolies, A, 2016.

The impact of youth skills training on the financial behaviour, employability and educational choice in Morocco, 3ie Impact Evaluation Report 54. Bausch, J, Dyer, P, Gardiner, D, Kluve, J and Mizrokhi, E, 2016.

Using advertisements to create demand for voluntary medical male circumcision in South Africa, 3ie Impact Evaluation Report 53. Frade, S, Friedman, W, Rech, D and Wilson, N, 2016.

The use of peer referral incentives to increase demand for voluntary medical male circumcision in Zambia, 3ie Impact Evaluation Report 52. Zanolini, A, Bolton, C, Lyabola, LL, Phiri, G, Samona, A, Kaonga, A and Harsha Thirumurthy, H, 2016.

Using smartphone raffles to increase demand for voluntary medical male circumcision in Tanzania, 3ie Impact Evaluation Report 51. Mahler, H and Bazant, E, 2016.

Voluntary medical male circumcision uptake through soccer in Zimbabwe, 3ie Impact Evaluation Report 50. DeCelles, J, Kaufman, Z, Bhauti, K, Hershow, R, Weiss, H, Chaibva, C, Moyo, N, Braunschweig, E, Mantula, F, Hatzold, K and Ross, D, 2016.

Measuring the impact of SMS-based interventions on uptake of voluntary medical male circumcision in Zambia, 3ie Impact Evaluation Report 49. Leiby, K, Connor, A, Tsague, L, Sapele, C, Koanga, A, Kakaire, J and Wang, P, 2016.

Assessing the impact of delivering messages through intimate partners to create demand for voluntary medical male circumcision in Uganda, 3ie Impact Evaluation Report 48. Semeere, AS, Bbaale, DS, Castelnuovo, B, Kiragga, A, Kigozi, J, Muganzi, A, Kambugu, A and Coutinho, AG, 2016.

Optimising the use of economic interventions to increase demand for voluntary medical male circumcision in Kenya, 3ie Impact Evaluation Report 47. Thirumurthy, H, Omanga, E, Rao, SO, Murray, K, Masters, S and Agot, K, 2016.

The impact of earned and windfall cash transfers on livelihoods and conservation in Sierra Leone, 3ie Impact Evaluation Report 46. Bulte, E, Conteh, B, Kontoleon, A, List, J, Mokuwa, E, Richards, P, Turley, T and Voors, M, 2016.

Property tax experiment in Pakistan: Incentivising tax collection and improving performance, 3ie Impact Evaluation Report 45. Khan, A, Khwaja, A and Olken, B, 2016.

Impact of mobile message reminders on tuberculosis treatment outcomes in Pakistan, 3ie Impact Evaluation Report 44. Mohammed, S, Glennerster, R and Khan, A, 2016.

Making networks work for policy: Evidence from agricultural technology adoption in Malawi, 3ie Impact Evaluation Report 43. Beaman, L, BenYishay, A, Fatch, P, Magruder, J and Mobarak, AM, 2016.

Estimating the impact and cost-effectiveness of expanding access to secondary education in Ghana, 3ie Impact Evaluation Report 42. Dupas, P, Duflo, E and Kremer, M, 2016.

*Evaluating the effectiveness of computers as tutors in China, 3*ie Impact Evaluation Report 41. Mo, D, Bai, Y, Boswell, M and Rozelle, S, 2016.

Micro entrepreneurship support programme in Chile, 3ie Impact Evaluation Report 40. Martínez, CA, Puentes, EE and Ruiz-Tagle, JV, 2016.

Thirty-five years later: evaluating the impacts of a child health and family planning programme in Bangladesh, 3ie Impact Evaluation Report 39. Barham, T, Kuhn, R, Menken, J and Razzaque, A, 2016.

Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection and malnutrition in India, 3ie Impact Evaluation Report 38. Clasen, T, Boisson, S, Routray, P, Torondel, B, Bell, M, Cumming, O, Ensink, J, Freeman, M and Jenkins, M, 2016.

Evaluating the impact of vocational education vouchers on out-of-school youth in Kenya, 3ie Impact Evaluation Report 37. Hicks, JH, Kremer, M, Mbiti, I and Miguel, E, 2016.

Removing barriers to higher education in Chile: evaluation of peer effects and scholarships for test preparation, 3ie Impact Evaluation Report 36. Banerjee, A, Duflo E and Gallego, F, 2016.

Sustainability of impact: dimensions of decline and persistence in adopting a biofortified crop in Uganda, 3ie Impact Evaluation Report 35. McNiven, S, Gilligan, DO and Hotz, C 2016.

A triple win? The impact of Tanzania's Joint Forest Management programme on livelihoods, governance and forests, 3ie Impact Evaluation Report 34. Persha, L and Meshack, C, 2016.

The effect of conditional transfers on intimate partner violence: evidence from Northern Ecuador, 3ie Impact Evaluation Report 33. Hidrobo, M, Peterman, A and Heise, L, 2016.

The effect of transfers and preschool on children's cognitive development in Uganda, 3ie Impact Evaluation Report 32. Gillian, DO and Roy, S, 2016.

Can egovernance reduce capture of public programmes? Experimental evidence from India's employment guarantee, 3ie Impact Evaluation Report 31. Banerjee, A, Duflo, E, Imbert, C, Mathew, S and Pande, R, 2015.

Improving maternal and child health in India: evaluating demand and supply strategies, 3ie Impact Evaluation Report 30. Mohanan, M, Miller, G, Forgia, GL, Shekhar, S and Singh, K, 2016.

Smallholder access to weather securities in India: demand and impact on production decisions, 3ie Impact Evaluation Report 28. Ceballos, F, Manuel, I, Robles, M and Butler, A, 2015.

What happens once the intervention ends? The medium-term impacts of a cash transfer programme in Malawi, 3ie Impact Evaluation Report 27. Baird, S, Chirwa, E, McIntosh, C and Özler, B, 2015.

Validation of hearing screening procedures in Ecuadorian schools, 3ie Impact Evaluation Report 26. Muñoz, K, White, K, Callow-Heusser, C and Ortiz, E, 2015.

Assessing the impact of farmer field schools on fertilizer use in China, 3ie Impact Evaluation Report 25. Burger, N, Fu, M, Gu, K, Jia, X, Kumar, KB and Mingliang, G, 2015.

The SASA! study: a cluster randomised trial to assess the impact of a violence and HIV prevention programme in Kampala, Uganda, 3ie Impact Evaluation Report 24. Watts, C, Devries, K, Kiss, L, Abramsky, T, Kyegombe, N and Michau, L, 2014.

Enhancing food production and food security through improved inputs: an evaluation of Tanzania's National Agricultural Input Voucher Scheme with a focus on gender impacts, 3ie Impact Evaluation Report 23. Gine, X, Patel, S, Cuellar-Martinez, C, McCoy, S and Lauren, R, 2015.

A wide angle view of learning: evaluation of the CCE and LEP programmes in Haryana, 3ie Impact Evaluation Report 22. Duflo, E, Berry, J, Mukerji, S and Shotland, M, 2015.

Shelter from the storm: upgrading housing infrastructure in Latin American slums, 3ie Impact Evaluation Report 21. Galiani, S, Gertler, P, Cooper, R, Martinez, S, Ross, A and Undurraga, R, 2015.

Environmental and socioeconomic impacts of Mexico's payments for ecosystem services programme, 3ie Impact Evaluation Report 20. Alix-Garcia, J, Aronson, G, Radeloff, V, Ramirez-Reyes, C, Shapiro, E, Sims, K and Yañez-Pagans, P, 2015.

A randomised evaluation of the effects of an agricultural insurance programme on rural households' behaviour: evidence from China, 3ie Impact Evaluation Report 19*.* Cai, J, de Janvry, A and Sadoulet, E, 2014.

Impact of malaria control and enhanced literacy instruction on educational outcomes among school children in Kenya: a multi-sectoral, prospective, randomised evaluation, 3ie Impact Evaluation Report 18*.* Brooker, S and Halliday, K, 2015.

Assessing long-term impacts of conditional cash transfers on children and young adults in rural Nicaragua, 3ie Impact Evaluation Report 17*.* Barham, T, Macours, K, Maluccio, JA, Regalia, F, Aguilera, V and Moncada, ME, 2014.

The impact of mother literacy and participation programmes on child learning: evidence from a randomised evaluation in India, 3ie Impact Evaluation Report 16*.* Banerji, R, Berry, J and Shortland, M, 2014.

A youth wage subsidy experiment for South Africa, 3ie Impact Evaluation Report 15. Levinsohn, J, Rankin, N, Roberts, G and Schöer, V, 2014.

Providing collateral and improving product market access for smallholder farmers: a randomised evaluation of inventory credit in Sierra Leone, 3ie Impact Evaluation Report 14*.* Casaburi, L, Glennerster, R, Suri, T and Kamara, S, 2014.

Scaling up male circumcision service provision: results from a randomised evaluation in Malawi, 3ie Impact Evaluation Report 13. Thornton, R, Chinkhumba, J, Godlonton, S and Pierotti, R, 2014.

Targeting the poor: evidence from a field experiment in Indonesia, 3ie Impact Evaluation Report 12. Atlas, V, Banerjee, A, Hanna, R, Olken, B, Wai-poi, M and Purnamasari, R, 2014.

An impact evaluation of information disclosure on elected representatives' performance: evidence from rural and urban India, 3ie Impact Evaluation Report 11*.* Banerjee, A, Duflo, E, Imbert, C, Pande, R, Walton, M and Mahapatra, B, 2014.

Truth-telling by third-party audits and the response of polluting firms: Experimental evidence from India, 3ie Impact Evaluation Report 10. Duflo, E, Greenstone, M, Pande, R and Ryan, N, 2013.

No margin, no mission? Evaluating the role of incentives in the distribution of public goods in Zambia, 3ie Impact Evaluation Report 9*.* Ashraf, N, Bandiera, O and Jack, K, 2013.

Paying for performance in China's battle against anaemia, 3ie Impact Evaluation Report 8*.* Zhang, L, Rozelle, S and Shi, Y, 2013.

Social and economic impacts of Tuungane: final report on the effects of a community-driven reconstruction programme in the Democratic Republic of Congo, 3ie Impact Evaluation Report 7. Humphreys, M, Sanchez de la Sierra, R and van der Windt, P, 2013.

The impact of daycare on maternal labour supply and child development in Mexico, 3ie Impact Evaluation Report 6. Angeles, G, Gadsden, P, Galiani, S, Gertler, P, Herrera, A, Kariger, P and Seira, E, 2014.

Impact evaluation of the non-contributory social pension programme 70 y más in Mexico, 3ie Impact Evaluation Report 5. Rodríguez, A, Espinoza, B, Tamayo, K, Pereda, P, Góngora, V, Tagliaferro, G and Solís, M, 2014.

Does marginal cost pricing of electricity affect groundwater pumping behaviour of farmers? Evidence from India, 3ie Impact Evaluation Report 4. Meenakshi, JV, Banerji, A, Mukherji, A and Gupta, A, 2013.

The GoBifo project evaluation report: Assessing the impacts of community-driven development in Sierra Leone, 3ie Impact Evaluation Report 3. Casey, K, Glennerster, R and Miguel, E, 2013.

A rapid assessment randomised-controlled trial of improved cookstoves in rural Ghana, 3ie Impact Evaluation Report 2. Burwen, J and Levine, DI, 2012.

The promise of preschool in Africa: A randomised impact evaluation of early childhood development in rural Mozambique, 3ie Impact Evaluation Report 1. Martinez, S, Naudeau, S and Pereira, V, 2012.

In India, people living in urban slums do not have equitable access to water and sanitation, which can have a direct impact on their health. Since there are numerous constraints to increasing access to private toilets, providing public and community toilets with a nominal user fee can be a viable solution to tackle urban sanitation challenges. Authors of this study examine if improvements in supply-side issues related to community toilets and information campaigns can impact the users' willingness to pay, as well as the overall quality and usage of these toilets in the cities of Lucknow and Kanpur.

Impact Evaluation Series

International Initiative for Impact Evaluation
202-203, Rectangle One
D-4, Saket District Centre
New Delhi – 110017
India

3ie@3ieimpact.org
Tel: +91 11 4989 4444

www.3ieimpact.org