Valuing Open Defecation Free Surroundings: Experimental Evidence from a Norm-Based Intervention in India

Authors:

(i) Upasak Das (corresponding author)

Presidential Fellow in Economics of Poverty Reduction, Global Development Institute, University of Manchester

Research Affiliate, Center for Social Norms and Behavioral Dynamics, University of Pennsylvania, Philadelphia, PA, United States

(Email: upasak.das@manchester.ac.uk)

(ii) Sania Ashraf

Center for Social Norms and Behavioral Dynamics, University of Pennsylvania, Philadelphia, PA, United States

(Email: asania@upenn.edu)

(iii) Kavita Chauhan

Center for Social Norms and Behavioral Dynamics, University of Pennsylvania, Philadelphia, PA, United States

(Email: kavita.chauhan@gmail.com)

(iv) Alex Sphenev

Center for Social Norms and Behavioral Dynamics, University of Pennsylvania, Philadelphia, PA, United States

(Email: shpenev@sas.upenn.edu)

(v) Cristina Bicchieri

Center for Social Norms and Behavioral Dynamics, University of Pennsylvania, Philadelphia, PA, United States

(Email: cb36@sas.upenn.edu)

Abstract

Open defecation, which is found to be associated with poorer health outcomes apart from lower cognitive ability and productivity has widespread in India. This paper assesses the impact of a randomized norm-centric intervention implemented in peri-urban areas of Tamil Nadu in India on raising the value attached to residence in areas with a lower prevalence of open defecation measured through Willingness to Pay (WTP). The treatment hinges on the hypothesis that one's perception of the prevalence of certain behavior within the community and its approval may influence her own behavior and beliefs. Accordingly, through wall paintings, household visits, and community meetings, the intervention targeted changing the norms about toilet usage for defecation. The findings indicate a significant increase in the WTP for relocating in areas with zero and a half prevalence of open defecation from those with high prevalence. The estimates are robust to potential bias due to local political factors and COVID-led attrition. We also find changes in empirical expectations (what one believes about the prevalence of toilet usage in the community) is the key mediating channel because of which we also find a significant increase in toilet ownership and usage. Normative expectations (what one believes about approval of toilet usage) also yield some mediating effects, albeit to a lesser extent. Importantly, because of the nature of the intervention, we account for the possibility of contamination of control areas. The local average treatment effect estimations through two-stage least square regressions also indicate a similar impact. The findings underscore the need for norm-centric intervention to propel change in beliefs and achieve long-term and sustainable sanitation behavior.

Acknowledgement

This project was funded by the Bill and Melinda Gates Foundation (Grant No. INV-009118 / OPP1157257). The funder was not involved in study design, data collection, data analysis, and data interpretation. Swasti Health Catalyst implemented the intervention on the ground. Kantar Republic and Academy of Management Studies managed the data collection but had no role in the study design, decision to publish, or preparation of the manuscript. The study protocol has been reviewed and approved by the ethics board at the University of Pennsylvania, USA, and the Catalyst Foundation, India. Trial registration: NCT04269824. We thank Jinyi Kuang, Erik Thulin, Peter McNally, Maryann Delea, Rachel Sanders, and Hans Peter Kohler for their suggestions and help while setting up the project along with all the respondents and the investigators of the survey work. We also thank Anubha Tyagi and Paulius Yamin for their logistical assistance.

1. Introduction

Inadequate access to improved sanitation leads to a higher prevalence of soil-transmitted helminths, enteric diseases, infant mortality, reduced cognitive growth among children, their height and well-being with poorer educational outcomes (Saleem et al., 2019; Strunz et al., 2014; Hammer and Spears, 2016; Spears and Lamba, 2016; Geruso and Spears, 2018; Spears, 2020; Cameron et al. 2021). Therefore, the Sustainable Development Goals (SDGs) target 6.2 calls for universal access to sanitation and putting an end to open defecation (OD). Nevertheless, the problem with regard to access to safe sanitation has been substantial in India. About 60% of the world's population, who defecate in the open, have been estimated to come from India (WHO/UNICEF, 2014). Accordingly, successive governments have implemented nationwide sanitation promotion programs, the most recent of which is the Swachh Bharat Abhiyan (SBA), which promotes behavioral change interventions to encourage toilet construction and its usage. Despite some documented evidence surrounding the impact of the program on toilet access, toilet usage remains low in India (Barnard et al. 2013; Coffey et al. 2017) The recently conducted National Family Health Survey in 2019-21 found about 17 percent of the households in India have no access to toilet and in rural areas it goes up to 25 percent.

One of the striking features of OD practice in India as argued by Coffey et al. (2014) is its revealed preference. Bicchieri et al. (2018) found that OD among the ones who have access to toilets in parts of Bihar and Tamil Nadu is about 24 percent. Coffey et al. (2014) argue that many engage in OD because they find it “pleasurable, comfortable, or convenient.” It is also reported that OD allows them to lead a healthy life. Further, among many, it is not recognized as a threat to health. Therefore, staying in areas that are open defecation free (ODF) may not be valued by many.

This paper utilizes a randomly implemented norms-based intervention in peri-urban areas of Tamil Nadu, India to examine its effect in altering the value attached to residing in areas with a lower prevalence of OD. More specifically, we measure the value attached by capturing the Willingness to Pay (WTP) for residents in varying levels of OD and assess the impact of the intervention on it. It is important to note that irrespective of whether a household uses the toilet themselves, an ODF environment might be effective in reducing the spread of harmful bacteria and germs. Hence, everyone benefits from an excreta-free environment, and others residing in

the area cannot be excluded from enjoying them, thereby it assumes the character of a public good. Accordingly, improved sanitation practices in a neighborhood might be potentially valued. In addition, evidence suggests that there exist a significant slippage in toilet usage and switch to open for defecation purpose (Coffey et al. 2017; Abebe and Tucho, 2020).

The intervention revolves around the fact that one's belief about others within their community can yield a significant impact on one's own behavior and beliefs. In other words, one's belief about the prevalence of certain behavior within the community and its approval among its members may motivate her to behave pro-socially (Steg and Vlek, 2009). Social Norms Theory (SNT) argues that interdependent behaviors can be influenced through social expectations that might take the form of Empirical Expectations (EE) (expectations regarding what others around them do) or Normative Expectations (NE) (expectations regarding what others around them think one should do) (Bicchieri 2005; Bicchieri et al, 2016). In this regard, studies have found evidence of how such social motivations and beliefs can induce social transformation in terms of higher female labor participation and breastfeeding practices among others (Jayachandran, 2015, Bursztyn et al., 2020).

Early diagnostic research using non-experimental data on OD across rural and urban parts of Bihar and Tamil Nadu indicates that toilet usage for defecation is influenced by EE and to a lesser extent by NE in terms of what others in the community use for defecation purposes (Bicchieri et al. 2018). In other words, one is more likely to use a toilet for defecation if he/she believes that others in the community also use a toilet and approve of this practice. Based on this research, we developed a norm-centric intervention that hinges on messages and activities that aim to increase the EE and NE of the community members. Specifically, the components of the intervention include the following: personalized counseling sessions using flipbooks, and stickers to signal improved sanitary practices; peer learning groups using multimedia and discussion guides; community level using audio announcements and community mobilization events. The target was to shift social beliefs about other people's latrine usage, which can motivate people to use toilets for defecation.

In this paper, we focus on the impact on how individuals value living in a community with improved sanitation practices by studying their willingness to pay. The motivation stems from the idea on "common is moral" which argues the behavior of others can drive what people

think about what is right and what is wrong (Kelley, 1971). People often use their perception of the prevalence or the frequency of a behavior around them as a signal to judge morally acceptable it is (Eriksson et al., 2021). Therefore, it is likely that a behaviour would be rated moral or right if they think it is common around them. Existing evidence on the morality of tax evasion found if tax evasion is perceived to be prevalent within their community, it is less likely to be judged harshly (Welch et al., 2005). Brauer and Chaurand (2010) documented similar findings for 46 "uncivil" behaviors

In addition, injunctive norms or NE can also drive the "common is moral" association (White et al. 2009). Therefore, people may be motivated to favor a behavior if they are around individuals who approve of the behavior or conform to the social injunctive norms. Putting the above two arguments together, if the EE and/or of any behavior appears to be higher, the corresponding behavior might possibly be considered morally right. We use these findings to examine if our intervention that relies on changing social expectations surrounding toilet usage and safe defecation practice can influence how people value ODF environments.

By comparing the respondents residing in areas that are exposed to the treatment with those in the control areas, we find that the intervention has been successful in significantly increasing the value they attach to improved sanitation practices. The willingness to pay for residing in an ODF area is found to go up by Rs. 336 a month ($51 annually) and that in an area with a 50 percent prevalence of OD by about Rs. 237 ($36 annually). These findings are found to be robust to adjusting for the potential bias from non-random attrition because of COVID and other socio-political events. On the mechanisms that drive this, we find that our intervention led to a significant increase in people's perception of the prevalence of toilet usage and ownership (EE). Further, we observe more people started approving toilet usage because of the intervention, thereby increasing the NE. The mediation effects suggest that the increase in EE contributed to 12-14 percent of the impact while the contribution of NE is lesser at less than 5 percent. Importantly, we find that the intervention led to an overall increase in toilet usage and ownership, which lends credence to our findings on WTP. Because of the nature of our intervention, we consider substantial contamination of control wards which might bias our estimates. Therefore, we also make use of the Local Average Treatment Effect (LATE) estimates to run two-stage least square regressions using the actual randomly assigned treatment as the

Instrument Variable (IV) and the reported treatment exposure as the main variable of interest. The impact on WTP remains similar even after making this adjustment underscoring the significant and robust gains from the intervention.

The paper offers a number of contributions. First, in terms of interventions to promote health behavior, there are two broad categories. The first relies on subsidies that lower the cost of healthy behavior. If the unwanted behavior spreads infection and diseases among others, private investment in health products may not optimal (Gertler, 2015). In addition, in poorer settings, subsidies might be necessary to ensure widespread adoption of the behavior. The second set comprises behavioral change interventions that depend on norm messaging and nudging that can lower the marginal cost of pro-social behavior (Ashraf et al. 2006; Thaler and Sunstein. 2009; Giné et al. 2010). This intervention used in this paper adds to the second possible set of experiments that are found to be effective in promoting healthy beliefs. In this context, as a second contribution, the paper adds to the growing literature on social norms. Existing literature in Economics has focussed on the persistence of cultural traits and norms (Giuliano 2007, Alesina et al. 2013). We study how provisions of information or norm messaging on empirical expectations can potentially change the existing descriptive and injunctive norms and then beliefs. While studies have utilized lab-based experiments or randomized vignettes to elicit these effects (Bicchieri and Xiao, 2008; Bicchieri et al. 2022a; Bicchieri et al. 2022b), field-based experimental studies that revolve around norm-centric interventions are limited. Third, existing studies have looked into how behavior change interventions affect behavior. For example, norm messaging have successfully utilized to reduce the prevalence of drug use, sexual assaults, and drinking problems (Perkins, 2003; Kramer and Levy, 2008; Hillenbrand-Gunn et al. 2010). We assess the impact of such intervention on the overall change in attitude and beliefs. This is especially pertinent in the context of sanitation behavior because how people value improved sanitation practices within their community is key for maintaining ODF surroundings. As already mentioned, there is a substantial prevalence of slippage to OD even after individuals start using toilets and if individuals start to value residence in OD areas, long-term sustenance can be effectively achieved. Third, existing studies have used norm-based messaging to promote latrine usage in India, predominantly guided by Behavior Centered Design (BCD); Community-Led Total Sanitation (CLTS), and the Risks, Attitudes, Norms, Abilities, and Self-regulation (RANAS) approach. These designs consider social norms as one of the drivers of behavior

change but do not explicitly leverage specific social expectations that can effectively influence norms (Aunger and Curtis, 2016; Kar and Chambers, 2019; Friedrich et al. 2020). In this way, our intervention is among the first that primarily relies on altering social expectations to achieve improved sanitation practices. Finally, the study assumes importance in the context of ensuring healthy environments and better health outcomes, especially in the post-COVID scenario. In terms of policy implications, the study can potentially help in effectively devising norm-centric interventions to curb OD, a problem India and other countries in the Global South have been endeavoring to curb for decades.

The paper is structured as follows. Section 2 discusses the motivation behind the intervention and its components. Section 3 explains the randomization process and data used in the analysis. Section 4 describes the empirical strategy and section 5 presents the estimations from the main regressions, robustness checks, mediation analysis, and other further analyses. Section 6 concludes with a discussion.

Intervention- motivation and design

Our intervention hinges on the following question: does one's belief about others within their community affect their own beliefs and actions? Studies have indicated that concerns about how individuals are perceived by others influence important decisions and behavior (Della Vigna et al. 2012; Bursztyn and Jensen, 2015; Bursztyn et al. 2017; DellaVigna et al., 2017). In the context of collective behavior, this question assumes pertinence, especially for behaviors that are intensely routed in community discourse. Studies have indicated that interventions that depend on changing one's own factual beliefs have often produced a negligible impact on altering one's own beliefs and behavior (Petit and Zalk, 2019). In contrast, interventions that target collective behavior change can often fast-track social transformation at the individual and community level. Therefore, such interventions can potentially motivate people to behave pro-socially.

Our intervention is grounded on the SNT, which argues that behavior is often driven by social norms or beliefs and/or practices of others in their community. As argued by Bicchieri (2006) and Bicchieri (2016), engagement in independent behaviors is not motivated by what others do or approve/disapprove of. Alternatively, interdependent behaviors are dependent on social expectations that take the form of (i) EE, which is what one believes about the prevalence

of that behavior within the community) or (ii) NE, which is what one believes about the approval of that behavior within the community). Experimental and non-experimental studies have documented that social expectations can explain the prevalence of exclusive breastfeeding, corruption, child marriage, and low female labor force participation among others (Bicchieri et al., 2014, Jayachandran, 2015, Bursztyn et al., 2020; Bicchieri et al. 2022a). To this end, using non-experimental data and experimental vignettes, Bicchieri et al. (2018) find that toilet usage is highly predicted by the EE and to a lesser extent by NE.

We frame our intervention that revolves around changing the social expectations surrounding toilet usage and defecation practice and call this demand-side behavior change intervention package as *Nam Nalavazhvu*, (NN) which in the Tamil language means well-being. The intervention focussed on broadcasting improved toilet usage of other community members through activities at all levels, thereby shifting the EE. Information on household toilet usage behavior in the study wards was collected and then the related information was disseminated during community events at the ward level and household visits. At the ward level, we depended on audio broadcasting using automobiles and wall paintings that promoted toilet ownership and usage. We also delivered these messages through social media networks through community influencers. At the household level, visual stickers were used to signal households following improved sanitary practices. In addition, counseling sessions and personalized advice were also provided.[1] The people involved in the implementation of the activities are the ward outreach workers, who are residents of the ward and have at least 12 years of formal education. Field supervisors have also been involved to facilitate ward-level activities. In addition, influential members of the community are utilized to disseminate promotional and motivational messages surrounding improved sanitary practices. It is important to note that we conducted the trial of improved practices for three months to refine and revise the intervention activities in the same study districts but in a different Town Panchayat far off from the study areas.[2] We hypothesize that these activities will subsequently change the EE as well as the NE which would drive up toilet ownership and usage. In addition, the changes in the social expectation would improve how

[1] For detailed information on the intervention, refer to Ashraf et al. (2021).

[2] The peri-urban areas within a district are divided into 3-4 Town Panchayats in Tamil Nadu. These Town Panchayats are broader administrative units in the peri-urban areas.

people value improved surroundings with improved sanitation behavior, which is what is examined in this paper.

The *Nam Nalavazhvu* intervention design is influenced by 2 years of mixed-method formative research that comprises two rounds of surveys aimed to understand the social determinants of toilet usage and diagnose the collective behavior around it using the SNT. In particular, we conducted a social network survey among 3370 individuals across rural, peri-urban and urban slums of Bihar and Tamil Nadu to understand the features of social networks and their relationship with sanitation behaviors (Bicchieri et al. 2018). Next, we conducted 18 focus group discussions to study the social and gender norms surrounding sanitation behavior (Ashraf e al. 2022). In addition, we conducted a longitudinal survey with additional respondents (of 5052 individuals) to understand social beliefs and expectations concerning toilet usage for defecation purposes.

Randomization process and data

Randomization

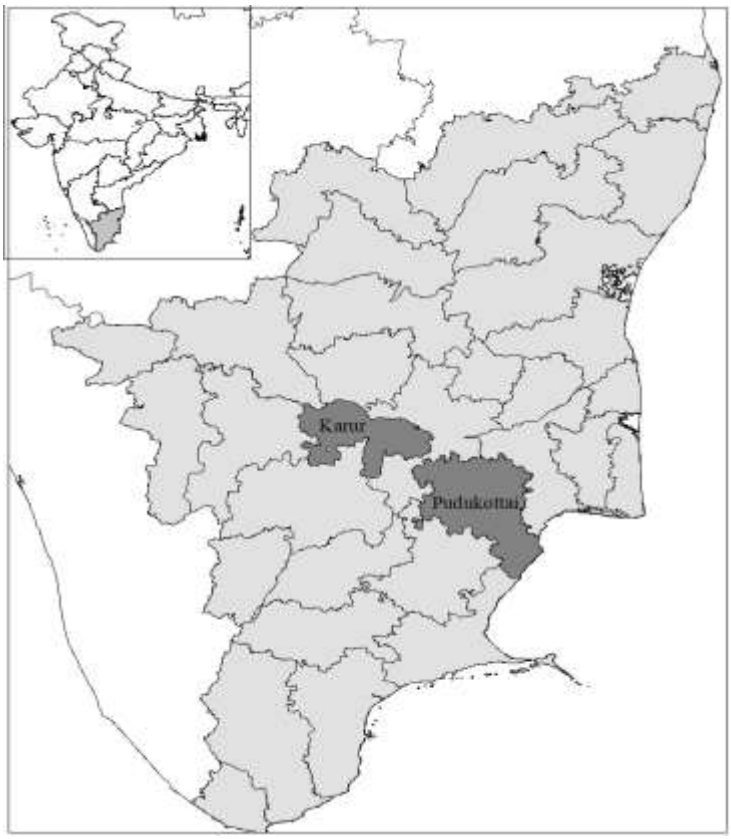
As discussed, the *Nam Nalavazhvu* intervention was randomly implemented across peri-urban areas in the state of Tamil Nadu in India. More specifically, we randomly chose five town panchayats each from the Pudukottai and Karur districts of Tamil Nadu.[3] The locations of Tamil Nadu along with that of these two districts are shown in Figure 1. Within these town panchayats, are clusters called wards, which forms the unit of randomization. For identification of the potential wards for our study purpose, we took help from the local executive officers to obtain the ward maps of each town panchayat. Wards that are identified as commercialized wards with very few residential households are excluded from our study. Further, other urbanized wards with complete coverage of toilet access according to the official records are excluded.[4] In addition, wards which are situated in the border of two or more adjoining wards are excluded

[3] The three town panchayats, which were used for piloting the intervention, have been excluded from the sampling frame. Further, we had to drop another town panchayat from which we could not receive permission to study because of political concerns.

[4] These wards are typically those where government officials or school/ college teachers stay. Some of these wards also have factories in which factory workers stay in allocated housing complex. Every apartment in such complex has been provided with improved toilets and hence OD is negligible in these wards.

from our analysis to reduce contamination of treatment and control wards. Of note is the fact that a minimum of 1 km distance has been ensured between each of the non-excluded ward.

Figure 1: Location of Tamil Nadu and Pudukottai and Karur districts within Tamil Nadu



In the ten randomly selected town panchayats, out of the existing 153 wards, 79 were found to be eligible wards based on the above-mentioned criteria. From these 79, a sample of 76 wards is randomly chosen for our study. These wards are more likely to represent peri-urban wards, used for residential purposes. Out of these 76 wards, in 38 randomly chosen wards, the *Nam Nalavazhvu* intervention has been implemented, thereby being the treatment wards with the remaining being the control wards. All the households residing in the treatment wards are eligible to participate in the activities listed above.

Data

We use primary data from two waves of the survey collected from 2,571 households. The first wave was conducted from January to March 2022 (baseline) and the second one from July to

September 2021 (endline). Note that the period in between was utilized for the intervention. From the sampled households, individuals of 18 years or above, who are deemed cognitively able to participate in the survey and also have been residing in the household since the last six months are listed. One individual from this list is then randomly chosen from each household. Around 34 households are chosen for the survey from each of the 76 wards that are considered for this analysis.[5]

In the two waves of the survey, we tried to ensure that the survey was conducted among the same households and respondents. However, majorly owing to the pandemic, there was a significant attrition that our survey suffered from. Because of reasons that varied from migration, illness, death, and non-entry of the surveyors, out of 2,571, we were able to re-survey only 1,874 households, amounting to attrition of about 27 percent. In the endline survey, for every household that we were not able to survey, we replaced it with another household in the vicinity of that one. In our main analysis, we use data from all the households surveyed in the endline. As a robustness check, we also present the findings from the non-attrition sample.

The survey questionnaire collected a wide range of information that included household socio-demographic and economic characteristics. We also gathered detailed data on household toilet ownership, respondent's defecation practice, and their social beliefs about toilet usage. Specific questions on willingness to pay for residing in ODF areas were asked. The endline survey also collected information on the extent of exposure of the respondents to the intervention across the treatment and control wards.

Our main outcome variable is the willingness to pay for residing in areas that have a lesser prevalence of OD. In particular, we asked the following:

(i) "Suppose there are two areas with similar houses you can rent. One, where there is zero open defecation, and another, where many defecate in the open. If you want to live in the area, where no one defecates in the open, how much extra are you willing to pay for rent per month?"

[5] The pre-analysis plan can be retrieved from https://clinicaltrials.gov/ProvidedDocs/24/NCT04269824/SAP_000.pdf (accessed on September 13, 2022)

(ii) "How much extra money per month will you pay to live in an area, where half of the people defecate in the open compared to an area where almost everyone defecates in the open?"

We use these questions to form our outcome variables. Please note here that we are gauging the valuation for two scenarios, the first being relocating from an area with high OD to an area with no OD and the second being to an area with 50 percent OD. This allows us to assess the impact of the intervention on valuing an area with no OD and another where half of the residents defecate in the open and the remaining half do not.

To ensure that the randomization was successful, we consider a number of household and individual level characteristics and compare them across the treatment and control arms taking their baseline values. Table 1 presents the differences in the means test between the two arms for these variables. The findings indicate that most of the characteristics are similar across the two groups. We also compare toilet ownership and toilet usage for defecation last time, three and seven days, all prior to the survey. In addition, we also assess the difference in EE and NE along with other measures of social expectations. For majority of the variables overall, we do not any significant difference between the treatment and control groups of wards. This lends credence to the randomization process that we adopted. Please note that we did not collect information on WTP in the baseline and hence were unable to compare the values across the two groups. The fact that the individual and household characteristics along with toilet access, defecation place, and social expectation do not show any significant difference allows us to assume that the difference in WTP for residents in low OD areas would have been negligible before the intervention. Nevertheless, in our robustness section, we employ several empirical strategies to adjust for any potential bias that could occur because of the possible difference in valuation in residences in ODF surroundings.

Table 1: Balance between treatment and control arms

	Endline			Baseline: non-attrition		
	Control	Treatment	Difference	Control	Treatment	Difference
			Covariates			
Age	45.803	45.132	0.671	45.135	44.127	1.008
Years of education	7.423	8.084	-0.661**	7.653	8.161	-0.508
Gender: female	0.55	0.515	0.035*	0.551	0.518	0.033
Currently married	0.765	0.797	-0.032*	0.738	0.764	-0.026
Upper caste	0.269	0.205	0.064	0.19	0.208	-0.018
Household size	3.684	3.693	-0.009	3.598	3.588	0.01

Owns color TV	0.956	0.964	-0.008	1.068	1.07	-0.002
Owns computer	0.073	0.077	-0.003	0.046	0.099	-0.053*
Has internet connection	0.494	0.498	-0.004	1.545	1.515	0.03
Gas as cooking fuel	0.867	0.895	-0.028	0.757	0.813	-0.056
Has separate kitchen	0.768	0.804	-0.036	0.747	0.759	-0.013
Occupation: laborer	0.227	0.264	-0.037	0.207	0.201	0.006
Social Expectations						
EE: How many out of 10 people use a toilet every time to defecate?				7.197	7.528	-0.331
EE: Many people in my area are constructing a toilet				0.856	0.815	0.041
NE: How many out of 10 people think one should use a toilet to defecate?				8.751	8.883	-0.132
Toilet ownership and usage						
All members have access to toilet				0.778	0.822	-0.044
Last usage: toilet				0.738	0.77	-0.032
Last two days primary defecation place: toilet				0.735	0.764	-0.029
Exclusive toilets usage in last two days				0.664	0.704	-0.040
Exclusive toilets usage in last seven days				0.655	0.7	-0.045
Observations	1280	1291		893	981	

*** p<0.01, ** p<0.05, * p<0.1.

Estimation Strategy

We utilize the experimental design of our intervention that accounts for the potential selection bias and hence we are able to generate unbiased causal estimates. To assess the impact on WTP, we estimate the following regression equation:

$$Y_{iwp} = \alpha + \beta.Treatment_{wp} + \delta X_{iwp} + \pi_p + \varepsilon_{iwd} \quad (1)$$

Here, Y_{iwd} is the WTP for residence in surroundings with low rates of OD as reported by respondent, i from ward, w located in town panchayat, p. The treatment status of ward, w, is indicated by $Treatment_{wp}$. It takes the value of 1 if it has been intervened and 0 otherwise. X_{iwp} is the vector of individual and household level characteristics and the town panchayat fixed effects is indicated by π_p. The error term is shown through ε_{iwd}. In our case, β forms the main

coefficient of interest. Because the intervention has been assigned at the ward level, we cluster the standard error at the level of wards.

Please note that we use the whole of the endline data to estimate the above equation using not only those which have been re-surveyed but also the newly surveyed ones. We also estimate the above regression only for the respondents surveyed in both waves. Further, for indicators where we have data from both the waves which include toilet ownership, usage, and social expectations among others, we use the Analysis of Covariance (ANCOVA) to control for the baseline values for estimating the treatment effects (Hidrobo et al. 2016; Haushofer et al. 2020, Das et al. 2021). Here, we estimate the following regression equation:

$$Y_{iwpt=1} = \alpha + \beta.Treatment_{wp} + \gamma Y_{iwpt=0} + \delta X_{iwpt=0} + \pi_p + \varepsilon_{iwdt=1} \qquad (2)$$

Here, $t = 1$ indicates values from the endline wave and $t = 0$ is the baseline. The notations of other variables and subscripts remain the same as that in equation 1. Therefore, in essence, we regress the endline outcome variable and estimate the treatment effects after accounting for the baseline values of the outcome variable. In addition, covariates from the baseline wave have been controlled for in this equation.

Results

(i) Main results

Table 2 presents the estimates of the intervention effect on the two outcome variables as discussed from the regression as elucidated in equation 1. Here, we estimate an Ordinary Least Square (OLS) regression and present the Intent to Treat (ITT) estimate of being in the treated ward. We use three specifications: in the first one, we do not control for any covariates, in the second we include the individual and household characteristics and in the third one, we further include the town panchayat fixed effects. The results from the first specifications are presented in columns (1) and (2), columns (3) and (4) present the results from the second one, and those from the third one, which is also our preferred specification, are presented in column (5) and (6).

The findings indicate a statistically significant increase in willingness to pay for residing in ODF communities among respondents from the treated wards. Specifically, we find an increase of WTP by about Rs. 336 a month ($51 annually) on average for moving from an area

with a high prevalence to OD to another with no prevalence, which can be causally linked with our intervention. We also find that individuals from the intervention wards are willing to pay more for moving to areas with 50 percent OD from areas with almost 100 percent defecation. As one would expect, the increase in WTP, in this case, is lower at around Rs. 237 ($36 annually). Nevertheless, the findings indicate a discernible increase in valuing relocation in surroundings with a lesser prevalence of OD. We also use the covariates that were listed in the pre-analysis plan and re-run the regressions. The estimates remain similar.

Table 2: Impact on willingness to pay

	WTP (full OD)	WTP (half OD)	WTP (full OD)	WTP (half OD)	WTP (full OD)	WTP (half OD)
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.334**	0.240**	0.321**	0.233**	0.336***	0.237***
	(0.160)	(0.111)	(0.151)	(0.108)	(0.105)	(0.085)
Controls	No	No	Yes	Yes	Yes	Yes
Town Panchayat FE	No	No	No	No	Yes	Yes
Observations	2571	2571	2571	2562	2562	2562

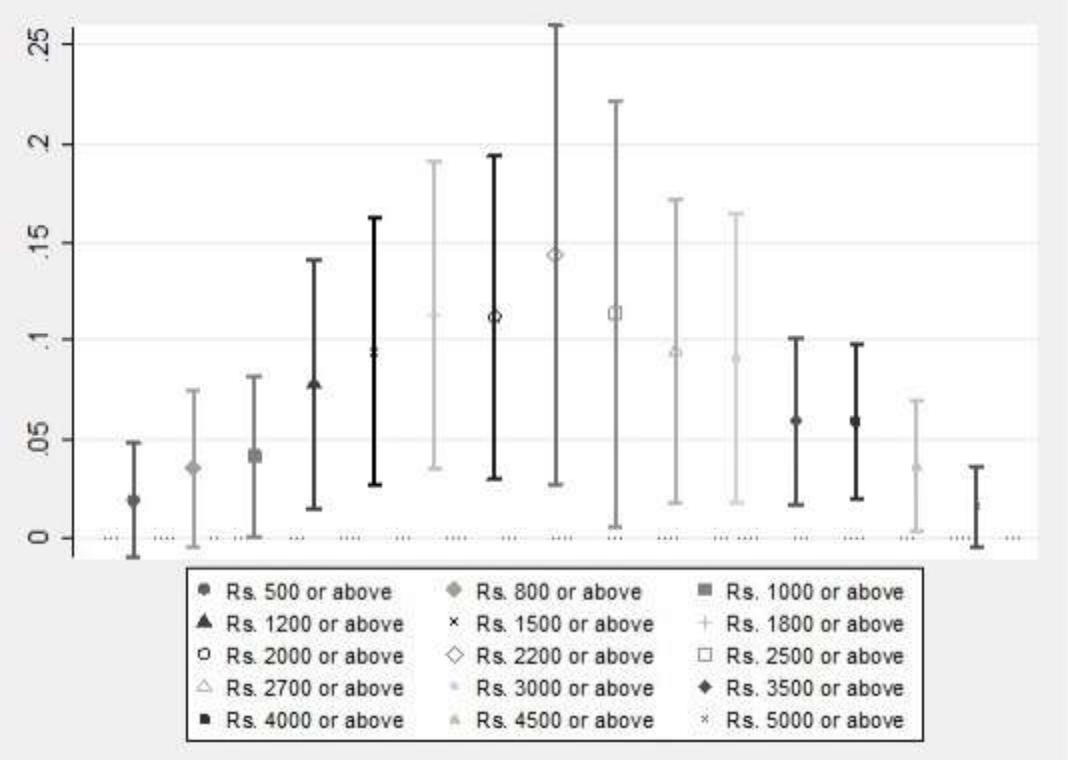
Note: The marginal effects from LPM model are presented. The WTP is divided by 1000 and then used in the regression. The standard errors clustered at the ward level are presented in the parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

It must be noted that we also collected additional indicators to elicit the WTP for residing in ODF surroundings. Here, we use a discrete choice experiment and start off asking the following question: *"Would you be willing to pay Rs. 2000 extra to live in an area with zero open defecation compared to moving to an area where most people defecate in the open for a similar house?"*. If the respondent says "yes" to the above, we further provide the following choices one by one: *Rs. (2200; 2500; 2700; 3000; 3500; 4000; 4500; 5000 and above Rs. 5000)* and stop wherever the respondent says "no". If the response to the above question is "no", we start providing the following choices one by one till the respondent says "yes": *Rs. (1800; 1500; 1200; 1000; 800; 500; 300; 100; below Rs. 100; I would not pay anything).*

We use these sets of questions to examine whether the intervention has been successful in raising the WTP at least by Rs. 2000. Here, the outcome variable assumes the value of "1" if the reported WTP is Rs. 2000 and "0" otherwise; therefore it is binary in nature. In addition, we repeat the same exercise by taking the choice threshold from Rs. 500 to Rs. 5000 instead of Rs. 2000. In other words, this variable takes the value of "1" if the final amount is the threshold amount or above and "0" otherwise. This allows us to assess the intervention effects on increasing the probability of paying a higher amount at least by the revealed threshold as we

increase it. We use the same regression as shown in equation 1 and estimate a Linear Probability Model (LPM). The marginal effects from the regression, along with the 90% confidence interval are shown in figure 1. The findings reveal a U-shaped effect of the intervention. We do not observe any significant difference at the lower end of WTP between the treatment and control arms. For example, there is no significant difference in the likelihood of paying Rs. 500/ Rs. 800 or more between the two groups. However, as the threshold value increases, we observe a disproportionate and significant increase in the likelihood of willingness to pay the amount in the treated wards. Nevertheless, as this value increases, we find that the difference in the likelihood between the two groups starts decreasing and becomes statistically insignificant at the value of Rs. 5000 or more. This is intuitive as it suggests that the intervention has been successful in increasing how people value ODF surroundings but not beyond a certain threshold.

Figure 1: Results from the discrete choice experiment



Note: The marginal effects from LPM model are plotted. 90% Confidence Interval calculated by the standard errors, clustered at the ward level are also plotted.

Robustness check

(i) Bias adjusted treatment effect

To estimate the impact of the intervention, we utilize its random implementation across the wards. The comparison of socio-economic and demographic characteristics along with indicators related to toilet usage across the treated and control wards as shown in table 1 indicate the two groups are well-balanced. Nevertheless, we find one variable (years of education) is not balanced with the difference significant at a 5 percent level. The difference in gender and marital status is significant at a 10 percent level. In absence of the indicators on WTP in the baseline wave, we are not sure about the balance between the two arms in terms of these indicators. Further, with the onset of the COVID-19 pandemic, the attrition of our sample from the baseline to the endline has been substantial, which is also potentially non-random. Despite selecting replacement households randomly from the vicinity of the missing sample, any potential bias arising because of these factors should be accounted for.

To address this issue, we make use of the "selection on unobservables versus observable" approach to ensure that the treatment effects account for these potential biases (Altonji et al., 2005; Oster, 2019). The underlying assumption is that the extent of selection on unobservables is related proportionally to that on observables, the ratio of which is given by a parameter δ. Another parameter (R^2_{max}), which is the hypothetical explained variation in a regression that includes all possible observed and unobserved characteristics. Oster (2019) simulated a number of existing randomized experiments to propose that the $R^2_{max} = 1.3 * R^2_0$, where, R^2_0 indicates the *R*-squared value of the regression model that uses only the observed control variables. Next, with a given R^2_0 and R^2_{max}, the value of δ is calculated such that the marginal effect is zero. A value of $|\delta|$, which is greater than 1 implies that the explanatory power of the unobservables is higher than that of the controlled covariates. In a setting like ours, where the intervention has been randomly implemented, the selection bias through the unobservables is likely to be low if not negligible. Because of this, if we find the $|\delta| > 1$, it is likely to be improbable. The values of δ, calculated for each of the two main regressions as shown in table 3 (columns 5 and 6) are found to be 92.9 and 88.3 which is much higher than the recommended benchmark of 1. This

confirms additionally that our estimates capture the treatment effects and are not confounded by other factors.[6]

(ii) Inverse Probability Weighting

To account for the potential bias resulting from this non-random attrition, we make use of the treatment effects estimators with Inverse Probability Weighted (IPW) regression, which can control for the potential selection bias at the treatment stage (Wooldridge 2007). Table 3 presents the estimates from IPW. The findings indicate a significant increase in WTP because of the intervention.

Table 4: Bias-adjusted treatment effects and Inverse Probability Weighting

	Bias-adjusted treatment effect		Inverse Probability weighting	
	WTP (full OD)	WTP (half OD)	WTP (full OD)	WTP (half OD)
	(1)	(2)	(4)	(5)
Delta (δ)	159.441	74.904		
Treatment			0.330***	0.232***
			(0.042)	(0.031)
Observations	2562	2562	2562	2562

Note: The marginal effects from the LPM model are presented in columns 3 and 4. The WTP is divided by 1000 and then used in the regression. The standard errors clustered at the ward level are presented in the parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

(iii) Excluding wards affected by CAA-NRC protest

Back in December 2019, the Government of India enacted the Citizenship Amendment Act (CAA) Bill, which amends the Indian citizenship act to accept non-Muslim migrants from Afghanistan, Pakistan, and Bangladesh who entered India before 2014 following religious persecutions. The National Register of Citizens (NRC) will officially record every legal citizen of India, which requires a set of prescribed documents.[7] Following these two, major protests dominated by Muslims across the country broke out and continued from December 2019 to March 2020.[8] Because of this, in 3 out of the 76 Muslim-dominated wards, many households did

[6] This method has been used by a number of studies to examine the possibility of selection on unobservables in different contexts (Alesina et al. 2016; Michalopoulos and Papaioannou, 2016).

[7] More information on CAA can be accessed from https://egazette.nic.in/WriteReadData/2019/214646.pdf. More information on NRC can be obtained from https://blogs.lse.ac.uk/humanrights/2020/08/10/the-national-register-of-citizens-and-indias-commitment-deficit-to-international-law/ (accessed on September 7, 2022).

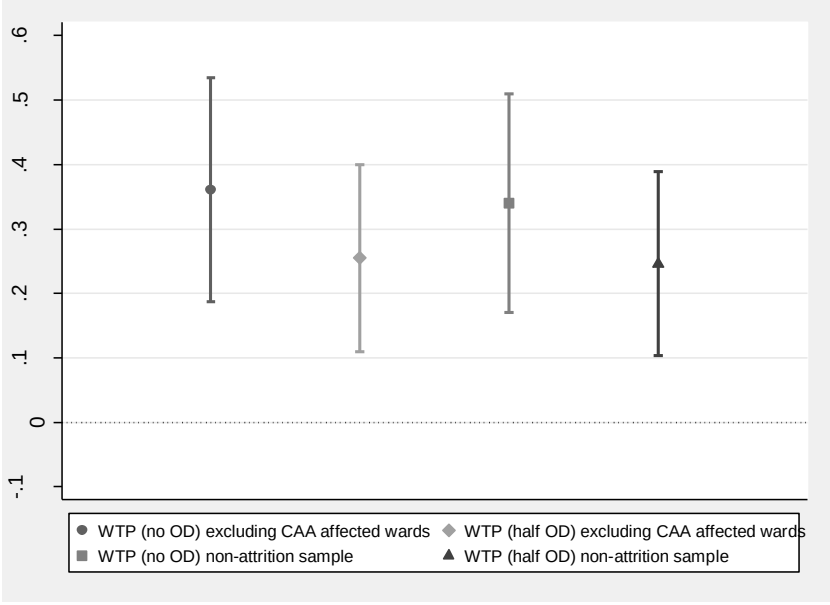
[8] Information on CAA-NRC protest can be obtained from https://indianexpress.com/about/caa-protest/ (accessed on September 7, 2022).

not give consent and we had to replace them with Hindu households within the same ward during the baseline wave. This compromised the random selection of households for the survey in these three wards. Nevertheless, even if we drop these three wards and re-run the main regressions, our inference on the intervention effects holds (figure 2).

(iv) Non-attrition sample

In addition to the above robustness exercises, we also check if our results hold for the non-attrition sample. As mentioned above, we were able to re-survey 1874 households, which were originally chosen before the onset of the pandemic. We re-ran the regression on this set of households and present the results in figure 2. The findings remain the same: a statistically significant increase in WTP is observed among households exposed to the intervention for moving from an area with high OD to another with zero or 50 percent prevalence. The effect size as well remains similar. Of note is the fact that the sample after attrition is also balanced between the treated and control arm at the baseline (table 1) and hence our estimates lend credence to the causal link.

Figure 2: Impact on willingness to pay excluding wards affected by CAA protests and non-attrition households



Note: The marginal effects from the LPM model are plotted. 90% Confidence Intervals calculated by the standard errors, clustered at the ward level are also plotted. The WTP is divided by 1000 and then used in the regression.

Mechanisms: Social Expectations

As discussed, we hypothesize that the "common is moral" theory is one potential mechanism that can drive up the valuation attached to ODF environment. In our case, this works through EE- one of the channels through which the observed increase in WTP is observed is the associated changes in how people perceive about the prevalence of toilet usage after the intervention. Further, if one perceives that the people around her approve of toilet usage, then the increase in WTP is also likely to be influenced by the associated changes in NE because of the intervention.

To study this, we first examine if the intervention can be linked to changes in EE and NE. To measure EE, we use a number of variables: *(i) "Out of ten members in your ward, how many do you think use a toilet every time they need to defecate?"; (ii) "Out of ten households in your ward, how many do you think own a toilet at home?"; (iii) "Many people in my ward built new toilets in the past six months"; and (iv) "I think more people expect others to use toilets compared to six months ago"*. For NE, the following is used: "*Out of ten members of your ward, how many do you think believe one should use a toilet to defecate?"*. We use these as our outcome variables and estimate the same regression as outlined in equation (1). The regression estimates are presented in table 5. We observe a significant increase in both, EE and NE on average because of the intervention. Not only are the respondents from the intervention wards more likely to report higher toilet usage in their community but also they think that more households around them are constructing toilets. In addition, we also find that they are more likely to think that their community approves of toilet usage for defecation.

Table 5: Effect on social expectations

	Empirical expectations				Normative expectations
	How many out of 10 use a toilet to defecate?	How many out of 10 own a toilet?	Many people are constructing a toilet	Many people are using a toilet	How many out of 10 think one should use a toilet to defecate?
Treatment	0.030	0.680***	0.122***	0.113***	0.615***
	(0.286)	(0.195)	(0.035)	(0.033)	(0.231)
All controls	Yes	Yes	Yes	Yes	Yes
Town Panchayat FE	Yes	Yes	Yes	Yes	Yes
Observations	2,562	2,562	2,562	2,562	2,562

Note: The marginal effects from the LPM model are presented. The standard errors clustered at the ward level are presented in the parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

Next, we use standard mediation analysis to explore how much of the treatment effect on WTP is through the changes in EE and NE. For this, we consider all those indicators of EE and NE for which a statistically significant intervention effect is found. We control for these variables in our main regression along with the other covariates and examine the changes in the marginal effects of the outcome variables if the regression is run without this variable and then with it. The findings reveal a considerable drop in the marginal effects for both the outcome variables that are considered (Table 6). Importantly this drop is observed when the hypothesized variables that measure EE and NE are considered. Nevertheless, it appears that the changes in NE had a lower mediating effect in comparison to EE.

Table 6: Effects after controlling for EE and NE

	WTP (no OD)					WTP (half OD)				
		Controlling for EE			Controlling for NE		Controlling for EE			Controlling for NE
	No EE or NE controls	How many out of 10 own a toilet?	Many people are constructing a toilet	Many people are using a toilet	How many out of 10 think one should use a toilet to defecate?	No EE or NE controls	How many out of 10 own a toilet?	Many people are constructing a toilet	Many people are using a toilet	How many out of 10 think one should use a toilet to defecate?
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment	0.336***	0.329***	0.295***	0.296***	0.325***	0.237***	0.229**	0.211**	0.213**	0.231**
	(0.105)	(0.105)	(0.099)	(0.101)	(0.105)	(0.085)	(0.087)	(0.082)	(0.084)	(0.088)
All controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Town Panchayat FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,562	2,562	2,562	2,562	2,562	2,562	2,562	2,562	2,562	2,562

Note: The marginal effects from LPM model are presented. The WTP is divided by 1000 and then used in the regression. The standard errors clustered at the ward level are presented in the parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

To explore this more formally, we use the "*medeff*" command in STATA developed by Hicks and Tingley (2011) to gauge the mediation effects. In particular, we use the framework developed by Imai et al. (2010) that provides the estimation results of the average direct effect (ADE), Average Causal Mediation Effect (ACME), and total effect. The findings from the mediation analysis are presented. 12-14 percent of the treatment gains on WTP for relocating to ODF area is found to be mediated by the indicators of EE that we consider. For relocation in areas with 50 percent OD, about 5 percent of about 5 percent and less than 2 percent respectively for the two types of outcomes variables considered. Importantly, the mediating effect of EE and NE is statistically significant at the 5% level for the WTP on zero OD. While, the mediating effects of EE on WTP for relocating to areas with half OD are also significant, that for NE is statistically indistinguishable from 0.

Table 7: Mediation effects

	WTP (no OD)			WTP (half OD)		
	Mean	95% confidence interval		Mean	95% confidence interval	
EE: Many people are constructing a toilet						
ACME 1	0.045	0.029	0.065	0.029	0.018	0.042
ACME 0	0.045	0.029	0.065	0.029	0.018	0.042
Direct Effect 1	0.274	0.190	0.361	0.203	0.143	0.266
Direct Effect 0	0.274	0.190	0.361	0.203	0.143	0.266
Total effect	0.320	0.224	0.422	0.232	0.164	0.305
EE: Many people are using a toilet						
ACME 1	0.045	0.029	0.065	0.027	0.016	0.040
ACME 0	0.045	0.029	0.065	0.027	0.016	0.040
Direct Effect 1	0.275	0.191	0.361	0.206	0.145	0.268
Direct Effect 0	0.275	0.191	0.361	0.206	0.145	0.268
Total effect	0.320	0.224	0.422	0.232	0.164	0.304
NE: How many out of 10 think one should use a toilet to defecate?						
ACME 1	0.015	0.002	0.030	0.004	-0.006	0.013
ACME 0	0.015	0.002	0.030	0.004	-0.006	0.013
Direct Effect 1	0.305	0.220	0.392	0.228	0.167	0.291
Direct Effect 0	0.305	0.220	0.392	0.228	0.167	0.291
Total effect	0.320	0.231	0.415	0.232	0.170	0.298

medeff command in STATA is used to estimate the ACME

Further Analysis

(i) Effect on toilet ownership and usage

If the intervention is able to change the social expectations through EE and NE as discussed above, which in turn increases individual demand or WTP for ODF environment, it should affect their own toilet usage behavior. To explore this further, we gauge the intervention effects on toilet usage and ownership separately.

We use the following questions asked during the survey and use them as outcome variables*: (i) "Some people defecate in the open and some people use a toilet. Where did you defecate the last time you needed to?"; (ii) "During the last two days, where was your primary place of defecation?"; (iii) "During the last two days, did you only use a toilet for defecation?"; and (iv) "During the last seven days, including today, did you use a toilet for defecation when you needed to defecate?"*. Please note that these variables are categorized as binary and we estimate an LPM regression model through equation (1) to elicit the impact of the intervention on toilet ownership and its usage for defecation purposes. As in the last section, we also use equation (2) to re-estimate the treatment effects after controlling for the baseline value using the non-attrition sample.

The estimates are presented in table 8. The findings indicate that the intervention had a statistically significant effect on increasing the probability of toilet ownership. Additionally, the prevalence of toilet usage also increased significantly because of the intervention. Importantly, exclusive toilet usage for defecation purposes is found to be higher when we use two-day and seven-day periods. Further, we also found that the likelihood of all household members to use latrines is higher in treated wards and hence we are able to confirm the treatment gains at the household level as well. Therefore, to sum up, we find an increase in EE and NE because of the intervention which led to an increase in toilet and ownership usage, thereby also raising the demand or value for improved sanitation practices.

Table 8: Impact on toilet ownership and usage

	Last usage: toilet	Last two days primary defecation place: toilet	Exclusive toilets usage in last two days	Exclusive toilets usage in last seven days	All members have access to toilet
Treatment	0.075**	0.076***	0.075**	0.082***	0.098**
	(0.029)	(0.028)	(0.029)	(0.028)	(0.041)
All controls	Yes	Yes	Yes	Yes	Yes
Town Panchayat FE	Yes	Yes	Yes	Yes	Yes
Observations	2,562	2,562	2,562	2,562	2,562

Note: The marginal effects from the LPM model are presented. The standard errors clustered at the ward level are presented in the parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

It must be noted that, for some of these variables measuring social expectations, toilet ownership, and usage, baseline information has also been collected. For these indicators, we re-estimate the treatment effects controlling for the baseline value for the non-attrition sample as elucidated in equation (2). We find that the intervention led to a significant increase in EE and NE. In addition, toilet ownership and usage for defecation purposes appears to have increased significantly (table 9).

Table 9: Treatment effects controlling for baseline values

	Social Expectations			Toilet ownership and usage				
	How many out of 10 own a toilet?	Many people are constructing a toilet	How many out of 10 think one should use a toilet to defecate?	Last usage: toilet	Last two days primary defecation place: toilet	Exclusive toilets usage in last two days	Exclusive toilets usage in last seven days	All members have access to toilet
Treatment	-0.018	0.160**	0.631**	0.078**	0.081***	0.077**	0.088***	0.104**
	(0.284)	(0.061)	(0.256)	(0.030)	(0.029)	(0.032)	(0.031)	(0.043)
All controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Town Panchayat FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,852	1,860	1,862	1,870	1,870	1,870	1,870	1,869

Note: The marginal effects from LPM model are presented. The standard errors clustered at the ward level are presented in the parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

(ii) Non-compliance

One of the key ingredients of our intervention is the provision of information dissemination through posters, wall painting, and household stickers. Therefore it is likely that people from the control wards who stepped into the treated ones would have seen these. Further, it is also likely that family relatives, or friends may stay in the treated wards through whom information from the intervention would have been flown and hence contaminate the control areas. While we have taken measures to avoid spillover of information from the treated to the control wards by ensuring a buffer space between these wards, practically it may not be sufficient to ensure zero contamination. Our data shows that 66 percent of the respondents within the treated wards have heard of the *Nam Nalavazhvu* intervention. However, 26 percent of the control wards have also

heard of the intervention. Therefore, because of this contamination, the impact estimates of the intervention might be biased.

To overcome this, we estimate the LATE to estimate the impact of the actually reported treatment status as reported during the survey on WTP. Because the actual treatment status might be endogenous to the outcome variables, we use the original randomized treatment status, which is exogenous as the IV. Here, we estimate a two-stage least square (2SLS) regression (Angrist and Pishcke, 2008).

We use whether the respondents have heard of *Nam Nalavazhvu*. First, we run naïve LPM regressions of our outcome variables on these two measures separately. The estimates from these regressions are presented in table 10 (columns 1 and 2). As one would expect, there is a statistically significant association between the treatment status and the outcome variables. Next, we use the original random assignment of the intervention as an IV and estimate a 2SLS model, the estimates of which are presented in columns 3 and 4 of the same table. We observe that the WTP for relocating to an area with zero OD increases by INR 842 and by INR 596 in an area with a 50 percent prevalence of OD for respondents who are exposed to the intervention. These LATE estimates are statistically significant at 1 percent level. Therefore, the estimates without adjusting for the potential contamination is possibly an under-estimate of the treatment effect.

Table 10: LATE estimates to account for contamination

	WTP (no OD)		WTP (half OD)	
	Naive	IV	Naive	IV
Heard of *Nam Nalavazhvu*	0.286***	0.165**	0.842***	0.596***
	(0.086)	(0.064)	(0.269)	(0.231)
All controls	Yes	Yes	Yes	Yes
Town Panchayat FE	Yes	Yes	Yes	Yes
Observations	2,562	2,562	2,562	2,562
F-stats		54.43		54.43

Note: The marginal effects from LPM model are presented. The WTP is divided by 1000 and then used in the regression. The standard errors clustered at the ward level are presented in the parenthesis. *ivreg2* command in STATA is used to estimate the IV regression. *** p<0.01, ** p<0.05, * p<0.1.

Discussion and Conclusion

Inadequate access to improved sanitation can lead to a number of infectious diseases and is also associated with lower productivity. Despite rigorous economic growth, OD in India has been a perpetual problem and many do not consider it to be real threat to health and well-being. In this

paper, we assess the impact of a randomly implemented norm-centric intervention on how people value residence in ODF areas. To elicit this valuation or demand, we use the WTP for relocating in OD surroundings or those with a lower prevalence of OD. The intervention comprised of activities like signalling and counselling sessions among others that aim to change people's perception about the prevalence and approval of toilet usage in the intervention areas. The objective was to shift social beliefs about latrine usage within the community, which can motivate people to use toilets for defecation and start valuing ODF surroundings.

The experiment was implemented in peri-urban areas of Tamil Nadu in India for about one year. The findings indicate a significant increase in WTP for relocation in OD areas among the respondents in the treated areas. The estimates are robust to response bias due to local political events or potential bias due to unwarranted attrition in our sample. We are also able to ensure that the treatment gains are intact even after considerable non-compliance because of the nature of the intervention. We observe a significant increase in respondents' EE in the treated areas, which is what they perceive about the prevalence of toilet usage in their community. The NE, which is the perceived approval of latrine usage is also found to be higher. We argue that these changes in EE and NE through the intervention are the key drivers to the observed increase in the valuation of residence in an area with improved sanitation behavior. Important to note here is that we also find an increase in toilet ownership and usage because of the intervention.

The study indicates that norm-centric intervention can not only change latrine usage behavior but also can be influential in raising the value of residing in areas with improved sanitation practices. Arguably, this is instrumental for long-term change which can be sustained. With regards to the target set up by India on being ODF, similar interventions that leverage social beliefs and expectation can be implemented. One potential limitation of the study is the lack of external validity attached to the experiment. However future research can comprise of studying the local context and devising and evaluating similar norm-centric centric interventions that comply with the local settings. Policy instruments can take a cue from the research to devise customized interventions to not only ensure higher toilet usage but also sustain that behavior.

References

Abebe, T. A., & Tucho, G. T. (2020). Open defecation-free slippage and its associated factors in Ethiopia: a systematic review. *Systematic reviews*, 9(1), 1-15.

Alesina, A., Giuliano, P., & Nunn, N. (2013). On the origins of gender roles: Women and the plough. *The quarterly journal of economics*, 128(2), 469-530.

Alesina, A., Harnoss, J., & Rapoport, H. (2016). Birthplace diversity and economic prosperity. *Journal of Economic Growth*, 21(2), 101-138.

Altonji, J. G., Elder, T. E., & Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. *Journal of political economy*, 113(1), 151-184.

Angrist, J. D., & Pischke, J. S. (2009). Mostly harmless econometrics: An empiricist's companion. Princeton university press.

Ashraf, N., Karlan, D., & Yin, W. (2006). Tying Odysseus to the mast: Evidence from a commitment savings product in the Philippines. The Quarterly Journal of Economics, 121(2), 635-672.

Ashraf, S., Bicchieri, C., Delea, M. G., Das, U., Chauhan, K., Kuang, J., ... & Thulin, E. (2021). Norms and Social Network–Centric Behavior Change Intervention (Nam Nalavazhvu) for Improved Toilet Usage in Peri-Urban Communities of Tamil Nadu: Protocol for a Cluster-Randomized Controlled Trial. JMIR research protocols, 10(5), e24407.

Ashraf, S., Kuang, J., Das, U., Shpenev, A., Thulin, E., & Bicchieri, C. (2022). Social beliefs and women's role in sanitation decision making in Bihar, India: An exploratory mixed method study. Plos one, 17(1), e0262643.

Aunger, R., & Curtis, V. (2016). Behaviour Centred Design: towards an applied science of behaviour change. Health psychology review, 10(4), 425-446.

Barnard, S., Routray, P., Majorin, F., Peletz, R., Boisson, S., Sinha, A., & Clasen, T. (2013). Impact of Indian Total Sanitation Campaign on latrine coverage and use: a cross-sectional study in Orissa three years following programme implementation. PloS one, 8(8), e71438.

Bicchieri, C. (2005). The grammar of society: The nature and dynamics of social norms. Cambridge University Press.

Bicchieri, C. (2016). Norms in the wild: How to diagnose, measure, and change social norms. Oxford University Press.

Bicchieri, C., & Xiao, E. (2009). Do the right thing: but only if others do so. Journal of Behavioral Decision Making, 22(2), 191-208.

Bicchieri, C., Ashraf, S., Das, U., Delea, M., Kohler, H. P., Kuang, J., ... & Thulin, E. (2018). Phase 2 project report. social networks and norms: Sanitation in Bihar and Tamil Nadu, India.

Bicchieri, C., Das, U., Gant, S., & Sander, R. (2022)a. Examining norms and social expectations surrounding exclusive breastfeeding: Evidence from Mali. World Development, 153, 105824.

Bicchieri, C., Dimant, E., Gächter, S., & Nosenzo, D. (2022)b. Social proximity and the erosion of norm compliance. Games and Economic Behavior, 132, 59-72.

Brauer, M., & Chaurand, N. (2010). Descriptive norms, prescriptive norms, and social control: An intercultural comparison of people's reactions to uncivil behaviors. European Journal of Social Psychology, 40(3), 490-499.

Bursztyn, L., & Jensen, R. (2015). How does peer pressure affect educational investments?. The quarterly journal of economics, 130(3), 1329-1367.

Bursztyn, L., Egorov, G., & Fiorin, S. (2017). From extreme to mainstream: How social norms unravel (No. w23415). National Bureau of Economic Research.

Bursztyn, L., González, A. L., & Yanagizawa-Drott, D. (2020). Misperceived social norms: Women working outside the home in saudi arabia. American economic review, 110(10), 2997-3029.

Cameron, L., Chase, C., Haque, S., Joseph, G., Pinto, R., & Wang, Q. (2021). Childhood stunting and cognitive effects of water and sanitation in Indonesia. Economics & Human Biology, 40, 100944.

Coffey, D., Gupta, A., Hathi, P., Khurana, N., Spears, D., Srivastav, N., & Vyas, S. (2014). Revealed preference for open defecation. Econ Polit Wkly, 49(38), 43.

Coffey, D., Spears, D., & Vyas, S. (2017). Switching to sanitation: Understanding latrine adoption in a representative panel of rural Indian households. Social Science & Medicine, 188, 41-50.

Das, U., Paul, A., & Sharma, M. (2021). Can information campaigns reduce last mile payment delays in public works programme?. UNU WIDER Working Paper 2021/21

DellaVigna, S., List, J. A., & Malmendier, U. (2012). Testing for altruism and social pressure in charitable giving. The quarterly journal of economics, 127(1), 1-56.

DellaVigna, S., List, J. A., Malmendier, U., & Rao, G. (2016). Voting to tell others. The Review of Economic Studies, 84(1), 143-181.

Eriksson, K., Vartanova, I., Ornstein, P., & Strimling, P. (2021). The common-is-moral association is stronger among less religious people. Humanities and Social Sciences Communications, 8(1), 1-8.

Friedrich, M., Balasundaram, T., Muralidharan, A., Raman, V. R., & Mosler, H. J. (2020). Increasing latrine use in rural Karnataka, India using the risks, attitudes, norms, abilities, and self-regulation approach: A cluster-randomized controlled trial. Science of The Total Environment, 707, 135366

Gertler, P., Shah, M., Alzua, M. L., Cameron, L., Martinez, S., & Patil, S. (2015). How does health promotion work? Evidence from the dirty business of eliminating open defecation (No. w20997). National Bureau of Economic Research.

Gertler, P., Shah, M., Alzua, M. L., Cameron, L., Martinez, S., & Patil, S. (2015). How does health promotion work? Evidence from the dirty business of eliminating open defecation (No. w20997). National Bureau of Economic Research.

Geruso, M., & Spears, D. (2018). Neighborhood sanitation and infant mortality. American Economic Journal: Applied Economics, 10(2), 125-62.

Giné, X., Karlan, D., & Zinman, J. (2010). Put your money where your butt is: a commitment contract for smoking cessation. American Economic Journal: Applied Economics, 2(4), 213-35.

Giuliano, P. (2007). Living arrangements in western europe: Does cultural origin matter?. Journal of the European Economic Association, 5(5), 927-952.

Hammer, J., & Spears, D. (2016). Village sanitation and child health: effects and external validity in a randomized field experiment in rural India. Journal of Health Economics, 48, 135-148.

Haushofer, J., Chemin, M., Jang, C., & Abraham, J. (2020). Economic and psychological effects of health insurance and cash transfers: Evidence from a randomized experiment in Kenya. Journal of Development Economics, 144, 102416.

Hicks, R., & Tingley, D. (2011). Causal mediation analysis. The Stata Journal, 11(4), 605-619.

Hidrobo, M., Peterman, A., & Heise, L. (2016). The effect of cash, vouchers, and food transfers on intimate partner violence: evidence from a randomized experiment in Northern Ecuador. American Economic Journal: Applied Economics, 8(3), 284-303.

Hillenbrand-Gunn, T. L., Heppner, M. J., Mauch, P. A., & Park, H. J. (2010). Men as allies: The efficacy of a high school rape prevention intervention. Journal of Counseling & Development, 88(1), 43-51.

Imai, K., Keele, L., & Yamamoto, T. (2010). Identification, inference and sensitivity analysis for causal mediation effects. Statistical science, 25(1), 51-71.

Jayachandran, S. (2015). The roots of gender inequality in developing countries. Annual Review of Economics, 7(1), 63-88.

Kar K, Chambers R. IDS and Plan UK. London, UK: 2009. [2019-12-04]. Handbook on Community-Led Total Sanitation (CLTS) http://www.communityledtotalsanitation.org/sites/communityledtotalsanitation.org/files/cltshandbook.pdf.

Kelley, H. H. (1971). Moral evaluation. American Psychologist, 26(3), 293.

Kremer, M., & Levy, D. (2008). Peer effects and alcohol use among college students. Journal of Economic perspectives, 22(3), 189-206.

Michalopoulos, S., & Papaioannou, E. (2016). The long-run effects of the scramble for Africa. American Economic Review, 106(7), 1802-48.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. Journal of Business & Economic Statistics, 37(2), 187-204.

Perkins, H. W. (Ed.). (2003). The social norms approach to preventing school and college age substance abuse: A handbook for educators, counselors, and clinicians. John Wiley & Sons.

Petit, V., & Zalk, T. (2019). Everybody Wants to Belong: A Practical Guide to Tackling and Leveraging Social Norms in Behavior Change Programming. New York, NY: UNICEF.

Routray, P., Torondel, B., Jenkins, M. W., Clasen, T., & Schmidt, W. P. (2017). Processes and challenges of community mobilisation for latrine promotion under Nirmal Bharat Abhiyan in rural Odisha, India. BMC Public Health, 17(1), 1-15.

Saleem, Mahrukh, Teresa Burdett, and Vanessa Heaslip. "Health and social impacts of open defecation on women: a systematic review." BMC public health 19.1 (2019): 1-12.

Spears, D. (2020). Exposure to open defecation can account for the Indian enigma of child height. Journal of Development Economics, 146, 102277.

Spears, D., & Lamba, S. (2016). Effects of early-life exposure to sanitation on childhood cognitive skills: Evidence from India's Total Sanitation Campaign. Journal of Human Resources, 51(2), 298-327.

Steg, L., & Vlek, C. (2009). Encouraging pro-environmental behaviour: An integrative review and research agenda. Journal of environmental psychology, 29(3), 309-317.

Strunz, E. C., Addiss, D. G., Stocks, M. E., Ogden, S., Utzinger, J., & Freeman, M. C. (2014). Water, sanitation, hygiene, and soil-transmitted helminth infection: a systematic review and meta-analysis. PLoS medicine, 11(3), e1001620.

Thaler, R. H., & Sunstein, C. R. (2009). Nudge: Improving decisions about health, wealth, and happiness. Penguin.

Welch, M. R., Xu, Y., Bjarnason, T., Petee, T., O'Donnell, P., & Magro, P. (2005). “But everybody does it…”: The effects of perceptions, moral pressures, and informal sanctions on tax cheating. Sociological Spectrum, 25(1), 21-52.

WHO/UNICEF (2014) Progress on drinking water and sanitation, 2014 Update. Joint Monitoring Programme for Water Supply and Sanitation (JMP)

Wooldridge, J. M. (2007). Inverse probability weighted estimation for general missing data problems. Journal of econometrics, 141(2), 1281-1301.