WPS6360

Policy Research Working Paper 6360

Impact Evaluation Series No. 83

Impact Evaluation of a Large-Scale Rural Sanitation Project in Indonesia

Lisa Cameron
Manisha Shah
Susan Olivia

The World Bank
Sustainable Development Network
Water and Sanitation Program
February 2013

Policy Research Working Paper 6360

Abstract

Lack of sanitation and poor hygiene behavior cause a tremendous disease burden among the poor. This paper evaluates the impact of the Total Sanitation and Sanitation Marketing project in Indonesia, where about 11 percent of children have diarrhea in any two-week period and more than 33,000 children die each year from diarrhea. The evaluation utilizes a randomized controlled trial but is unusual in that the program was evaluated when implemented at scale across the province of rural East Java in a way that was designed to strengthen the enabling environment and so be sustainable. One hundred and sixty communities across eight rural districts participated, and approximately 2,100 households were interviewed before and after the intervention. The authors found that the project increased toilet construction by approximately 3 percentage points (a 31 percent increase in the rate of toilet construction). The changes were primarily among non-poor households that did not have access to sanitation at baseline. Open defecation among these households decreased by 6 percentage points (or 17 percent). Diarrhea prevalence was 30 percent lower in treatment communities than in control communities at endline (3.3 versus 4.6 percent). The analysis cannot rule out that the differences in drinking water and handwashing behavior drove the decline in diarrhea. Reductions in parasitic infestations and improvements in height and weight were found for the non-poor sample with no sanitation at baseline.

This paper is a product of the Water and Sanitation Program (WPS), Sustainable Development Network. WSP is a multi-donor partnership created in 1978 and administered by the World Bank to support poor people in obtaining affordable, safe, and sustainable access to water and sanitation services. The paper is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The author may be contacted at lisa.cameron@monash.edu.

Impact Evaluation of a Large-Scale Rural Sanitation Project in Indonesia[1]

Lisa Cameron
Manisha Shah
Susan Olivia

JEL codes: I15, J10, O12
Keywords: rural sanitation, child health

[1] Lisa Cameron is a professor in the Department of Econometrics at Monash University and director of the Monash Centre of Development Economics (MCDE). Manisha Shah is an assistant professor of public policy at the University of California, Los Angeles, and a faculty research fellow at the National Bureau of Economic Research (NBER). Susan Olivia is a postdoctoral fellow in the Department of Econometrics at Monash University. The project's global impact evaluation task team is led by Bertha Briceño (in its early stages, it was led by Jack Molyneaux), together with Alexandra Orsola-Vidal and Claire Chase. Professor Paul Gertler has provided guidance and advice throughout the project. Advisors also include Sebastian Galiani, Jack Colford, Ben Arnold, Pavani Ram, Lia Fernald, and Patricia Kariger.The authors are also grateful to Almud Weitz, the regional team leader for WSP-East Asia, and Eduardo Perez, the global task team leader for the Scaling Up Rural Sanitation program. The WSP-Indonesia country technical assistance team was led by Nilanjana Mukherjee, followed by the late Ratna Indrawati Josodipoero, and is now headed by Deviariandy Setiawan. Generous financial support was provided by the Bill & Melinda Gates Foundation.

1. INTRODUCTION

Of the four most important causes of mortality for children under five years old in Indonesia, two (diarrhea and typhoid) are fecal-borne illnesses directly linked to inadequate water supply, sanitation, and hygiene issues (Ministry of Health, 2002). In Indonesia, about 11 percent of children have diarrhea in any two-week period and it has been estimated that more than 33,000 children die each year from diarrhea and 11,000 from typhoid (Curtis, 2004). Inadequate sanitation is associated not only with adverse health effects, but also with significant economic losses. A recent estimate of the economic losses in Indonesia from inadequate sanitation and poor hygiene practices puts them at more than 2.4 percent of the country's gross domestic product (GDP), or approximately US$6.3 billion (Napitupulu and Hutton, 2008).[2]

Despite wide recognition of the consequent health burdens, a large fraction of the world's population still lacks access to basic sanitation. For instance, it is estimated that 110 million people in Indonesia lack access to proper sanitation (WHO/UNICEF, 2012). Of these, 63 million practice open defecation (WHO/UNICEF, 2012). Open defecation is of fundamental importance to development because of the health hazard it poses to anyone living nearby. If some members of a community continue to defecate in the open, the whole community is at greater risk of diarrheal diseases than people living in communities where open defecation is not practiced. Preventable diseases that result from lack of sanitation and poor hygiene behavior cause a tremendous disease burden among the poor. The greatest disease burden falls on infants and children under five years old (Murray and Lopez, 1997). By reducing normal food consumption and nutrient adsorption, diarrheal diseases are also a significant cause of malnutrition, leading to impaired physical growth and cognitive development (Guerrant et al., 1999), reduced resistance to infection (Baqui et al., 1993), and, potentially, long-term gastrointestinal disorders (Schneider et al., 1978). Improvements in sanitation thus can help reduce the transmission of pathogens that cause diarrhea by preventing human fecal matter from contaminating environments.

In January 2007, the Water and Sanitation Program (WSP) launched the Scaling Up Rural Sanitation program[3] to address the precarious sanitation conditions of large rural populations in the developing world. With technical support from WSP, local and national governments are implementing the program in three countries: Indonesia, India, and Tanzania.

The Scaling Up Rural Sanitation program focuses on learning how to combine the approaches of Community-Led Total Sanitation (CLTS), behavior change communications, and social marketing of sanitation to generate sanitation demand and strengthen the supply of sanitation products and services at scale, leading to improved health for people in rural areas. The project

---

[2] A recent four-county study estimate found that the annual economic impact of poor sanitation is approximately US$1.4 billion (1.5 percent of GDP) in the Philippines, US$780 million (1.3 percent of GDP) in Vietnam, and US$450 million (7.2 percent of GDP) in Cambodia (WSP, 2008).

[3] For more information on the Scaling Up Rural Sanitation program, see www.wsp.org/scalingupsanitation.

also seeks to support government efforts to develop strong enabling environments to reform policy and sector institutions to create large-scale sustainable programs.

To test the effectiveness of its approach, the Scaling Up Rural Sanitation program incorporated a randomized controlled trial impact evaluation (IE) in the three project countries as a key component of its monitoring and evaluation framework. This evaluation provides a unique opportunity to learn what health and welfare impacts can be expected from sanitation improvements. The IE uses widely accepted impact evaluation protocols, and was designed to disrupt the planned program as little as possible. Importantly, the IE evaluates an intervention that was implemented at scale and led by the government under real-world conditions. This unusual approach is in contrast to smaller randomized controlled trials that are carried out with the intervention as a pilot and under much more controlled circumstances. By evaluating the project at scale, the IE provides a more reliable estimate of program impact in a scaled up, sustainable program.

The evaluation seeks to examine the program impacts along the causal chain. Poor sanitation leads to diarrhea and lack of nutrients. This adversely affects child height and weight and can lead to anemia. Anemia in young children is associated with poor child development. Thus, key outcome variables of interest are:

- changes in perceptions of consequences of poor sanitation;
- sanitation improvements (toilet construction and access to improved sanitation);
- place of defecation (reduction in open defecation);
- child health outcomes, which include:
    - diarrhea prevalence,
    - intestinal parasite infections,
    - stunting and wasting,
    - iron-deficiency anemia detected through minimally invasive finger-prick tests, and
    - cognitive and motor development.

This report aims to provide endline information for the selected indicators and outcomes of interest included in the study.

2. BACKGROUND AND DESCRIPTION OF THE PROGRAM

In Indonesia, the rural sanitation project is known as Total Sanitation and Sanitation Marketing (TSSM). TSSM aims to improve sanitation practices in rural communities of East Java by generating sanitation demand at scale and increasing the supply of sanitation products and services. The project’s approach differs from the government’s previous established sanitation policies of providing infrastructure and/or subsidies, instead sending facilitators to villages to initiate participatory analysis of existing sanitation practices and the consequences and implications of such practices, thus generating demand for better sanitation services that the market can then respond to.

In Indonesia, the project's programmatic approach consists of three main components:

Community-Led Total Sanitation (CLTS)

This component focuses on stopping open defecation. It aims to trigger the desire for an open-defecation-free (ODF) community. It does this by raising collective awareness of the open defecation problem. Facilitators are sent to communities to initiate analysis and discussions of the sanitation situation. These discussions are held in public places and are open to all. They involve a "walk of shame," during which the facilitator helps people analyze how fecal contamination spreads from the exposed excreta to their living environments and food and drinking water. A map of the village is drawn on the ground and villagers are asked to indicate where they live, where they defecate, and the routes they take there and back. It soon becomes apparent that everyone is ingesting small amounts of each other's feces (to people's horror and embarrassment). This inevitably leads to personal and collective decisions to be free of the hazard by becoming an ODF community. The villagers must forge their own plan for making this happen with only limited follow-up support and monitoring from the program. ODF status is verified by local government agencies. Communities achieving ODF status receive recognition and commendation from local and provincial governments.

Social Marketing of Sanitation

This component focuses on increasing the availability of sanitation-related products (such as latrines or slabs) and services (such as pit emptying). It involves extensive consumer and market research that investigates the sanitation solutions that people desire, the options available to them in the market, and their attitudes and knowledge of sanitation issues. The component develops targeted communications campaigns and enhances the supply of a range of sanitation goods and services that are responsive to preferences and economic capacities of all consumer segments. It also involves partnering with the local private sector—for example, to train local artisans to meet the increased demand for specific products that results from CLTS facilitation sessions.

Strengthening the Enabling Environment

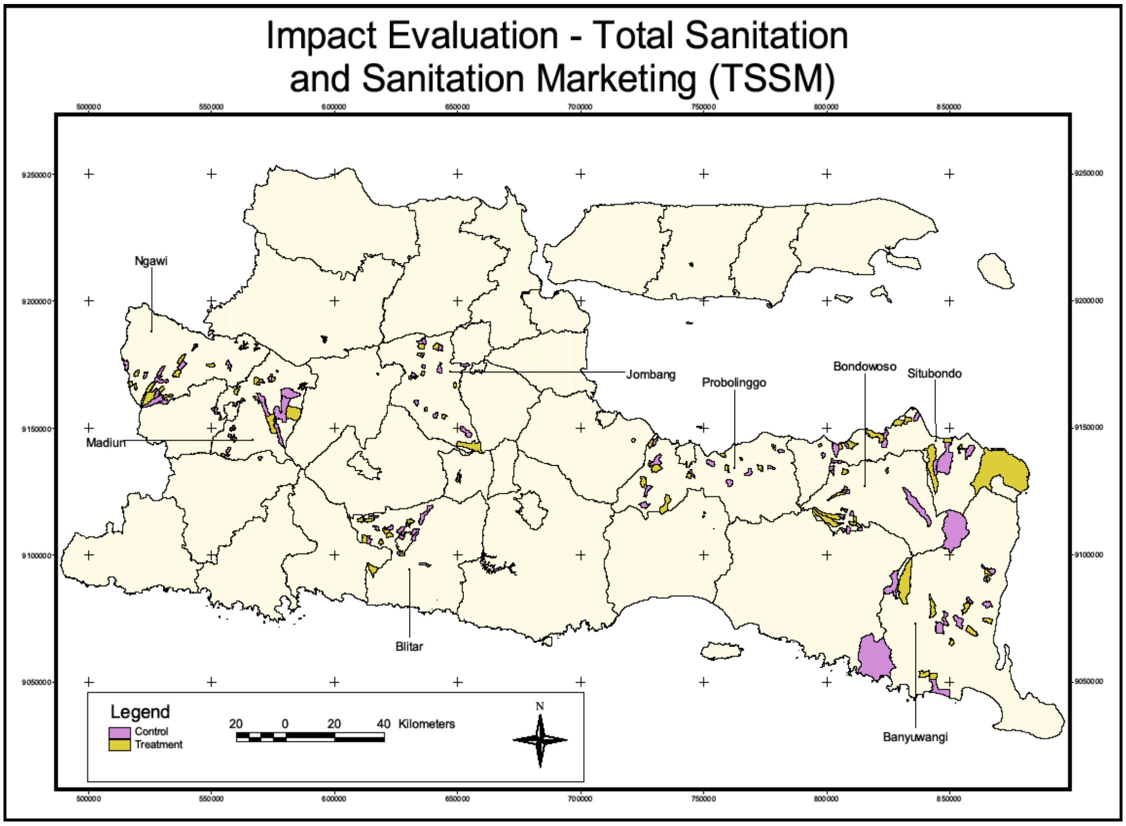
This component aims to support the development of policies and institutional practices that facilitate scaling up, program effectiveness, and sustainability. These include national, state, and local government sanitation policies; sanitation program financing, implementation, and management practices; fiscal rewards for results consistent with policies; training and accreditation of facilitators, masons, and vendors; and regulation and support of local private sector investment in improving sanitation.

3. EVALUATION DESIGN

Accurately measuring the long-term health and welfare impacts of these sanitation interventions requires an IE methodology that establishes the causal linkages between the intervention and the outcomes of interest. To estimate the causal relationship between the project (treatment) and the outcomes of interest, one needs an accurate counterfactual—that is, a comparison group that shows what would have happened to the target group in the absence of the intervention. The IE methodology uses randomization to construct the comparison group. Some communities were randomly selected to receive the treatment while others were randomly selected to serve as controls and not to receive the treatment within the period of the evaluation. If a nonrandom control group was used instead, a comparison of treated and untreated areas could confuse the program impact with pre-existing differences between villages. This is a particular problem if communities are chosen purposely as areas with a high likelihood of success because of favorable local conditions (strong leadership, existing water and sanitation infrastructure, highly educated populations, etc.) or because they differ in terms of hygiene habits, lower motivation, or other factors that are difficult to observe. This is known as selection bias. The use of randomization avoids these difficulties by ensuring that the communities that receive the intervention are no different on average from those that do not.

TSSM in Indonesia was implemented in rural East Java. Eight of 29 rural districts in East Java participated in the evaluation, with a total of 160 communities participating. The map in Figure 1 shows the location of the eight IE districts, which are spread across East Java. They are Probolinggo, Bondowoso, Situbondo, and Banyuwangi in the east of the province, and Ngawi, Madiun, Jombang, and Blitar in the west of the province.

Figure1. Treatment and control villages in East Java



WSP worked with local and national government and the local private sector to implement the project. East Java's 29 rural districts were divided into three groups: Phase 1 districts received the program first, Phase 2 districts received it next, and Phase 3 districts received it last. The evaluation was conducted in Phase 2 districts. Phase 2 was chosen largely on the basis of timing. Evaluating the program in Phase 2 districts provided sufficient time for the baseline survey to be conducted prior to program implementation. Many of the start-up issues confronted in Phase 1 were sorted out by Phase 2 so the evaluation provided an impact estimate, which is more representative of what could be expected from a national scaling up of the program following such large-scale piloting.[4] Districts participating in Phase 2 were asked if they would also participate in the evaluation. All of the 11 original Phase 2 districts responded affirmatively. Eight districts were ultimately chosen, again on the basis of the timing of the interventions, for a total of 160 sub-villages. The representativeness of these districts is discussed in Section 4.1.

[4] However, the social marketing component of TSSM could not be evaluated in Phase 2 given the later commencement of this part of the project.

In each of the participating districts, the IE team randomly selected 10 pairs of villages. Each pair consisted of one treatment village and one comparison village from the same *kecamatan* (sub-district). Villages in Indonesia have various communities or sub-villages, and the project intervention occurred at the sub-village level. At least one community in the treatment village received the full project intervention. No communities in the comparison villages received the project intervention.

4. SAMPLING STRATEGY

4.1 Selecting Sub-villages

From each of the eight districts, 10 treatment and 10 control villages were randomly chosen to participate in the IE. Local government offices from each district gave the IE team a list of at least 30 villages in which the program could be implemented.[5] Most district offices gave the IE teams lists of 40 to 70 villages. These are villages the districts had chosen to participate in the project based on sanitation needs, poverty levels, access to water, and so forth. Using a random number generator in STATA, the IE team randomly selected 10 treatment and 10 control villages from each district list. The IE team then sent the list of 20 villages back to the district government office (without telling them which villages had been selected as control and treatment villages). The reason for this is that the project is actually implemented at the *dusun*, or sub-village level. Villages generally have two to three sub-villages. Wanting the same selection criteria to be used for the selection of sub-village for both the treatment and control villages, the IE team asked each district office to provide the sub-village names for all 20 villages. District offices were told that some would be the treatment and others the control. The IE has internal validity but not external validity in that villages were not randomly chosen from the universe of villages. Different districts chose villages on the basis of different indicators. For example, some districts chose to include villages that had recently participated in water supply programs, whereas other districts explicitly chose to exclude such villages. The sample thus reflects the variety of ways in which government officials generally choose villages for a sanitation program so internal validity is sufficient under these circumstances. That is, the evaluation will provide estimates of the average impact expected given the way governments select villages for such programs. The impact of the different bases for the choices can be examined as part of the evaluation.

Once the IE team received the sub-village lists from the district offices for all 20 villages, they told district offices which villages were in the treatment group and which were in the control group. The district offices promised to do everything possible to make sure the treatment *dusun* were treated and the control *dusun* remained untreated. There was some concern by local

[5] At least two villages were listed for every sub-district on the district lists. This allowed pairing of treatment and control communities with sub-districts as described above.

program implementers that the program might spread like “wildfire” and that it would be difficult to deny it to control villages. However, sample sizes were selected based on this possibility and it does not appear that many control villages were contaminated.

4.2 Selecting Households

Listings were done in each sub-village in the control and treatment villages to gather information on the universe of households with children under the age of two years. These listings were based on information provided by the community health cadre. Thirteen households were then randomly selected from the listing to participate in the baseline survey. These 13 households were given priority rankings so survey teams knew to interview them. If one of the 13 households was unavailable to participate, it was replaced by another household chosen randomly from the listing. Detailed replacement methods are described below. In some of the sub-villages, there were not enough households with children under the age of two years. In those cases, information on households with children under age five was also collected. These households were ranked with priority rankings based on the total number of child under the age of two years, under the age of three years, under the age of four years, and under the age of five years. Households with younger children were given a higher priority.

Households in the sample had at least one living child under the age of two (unless there were not enough households with children under age two in the sub-village). If the child under age two had died or moved since the listing was conducted, the decision-making process was as follows:

1. Are all listed children under age two in this household deceased? If yes, is there another child under age two in this household? If yes, conduct interview. If no, replace the household.
2. If the child under age two is still alive at the listing time, there are three possibilities:
   a. Still alive and at same address for baseline survey: → interview
   b. Household moved but still lives in the same village: → find and interview
   c. Child under age two lives in another household that is in the target household list (and there is no other child under age two in this household):→ interview and add this household as a replacement.
3. If a household with children under age two has moved out of the village, replace the household.
4. Household replacement also applies in these cases:
   a. After four hours, the household still has not completed the interview. This could happen in households that contain only busy adults.
   b. A household with children under age two refuses to be interviewed. The supervisor must pay a visit to the household reported by interviewer and help solve any problems. If after the supervisor visit, the household still refuses, replace it.

c. A household can be a duplicate if the head of household's name, with the same characteristics, shows up more than once on the household list targeted to be interviewed in an enumeration area. In this case, only interview the household with the smallest number and replace the other household.
d. A household cannot be reached after four hours. This could happen if all household members are out of town, or adult household members are too busy to meet.
e. A household on the preprinted data listing is unknown to village authorities and villagers.

All replacements must be authorized by a supervisor.

5. DATA COLLECTION

This section provides information about the datasets used in the evaluation and compares the characteristics of the treatment and control communities. We also examine the data for differential sample attrition across program and comparison villages.

A baseline survey was conducted in both treatment and control communities in August – September 2008. A total of 2,087 households in 160 sub-villages were interviewed. A health nutrition module to record recent illness for children under 5 years of age was also administered, providing a baseline sample of 2,353 children.

The follow-up data collection was conducted approximately 24 months later, between November 2010 and February 2011. The endline survey used the same field methodology as the baseline survey. The final sample size was 2,500 households. Of these, 1,908 were also surveyed in the baseline and so constitute our "panel sample" (see Table 1).

6. VALIDATION OF EVALUATION DESIGN

6.1 Baseline Comparison of Means of Tests for Balance

We use the 2008 baseline survey data to compare characteristics of treatment and control groups. Table 2 compares the means of selected key variables across treatment and control groups. Randomization aims to minimize systematic differences between the two groups. Table 2 shows that the means of the variables are similar in magnitude for the two groups and we cannot reject that they are equal for most of the variables. Thus the randomization successfully created groups comparable along observable dimensions. That is, mean differences between the treatment and control groups are not statistically different from zero. For the key outcome variables (household water and sanitation condition, as well as children's health variables), balance is achieved. The demographic and socio-economic characteristics are also similar across treatment and control groups. The baseline report provides tests of balance on a more extensive set of variables.[6]

6.2 Attrition in the Follow-up Data

A legitimate concern in any follow-up survey involves the extent of sample attrition and the degree to which attrition is nonrandom. In the baseline, a total of 2,087 households were interviewed. Two years later, 1,908 households were successfully re-interviewed. Of the original 2,087 baseline household, 179 could not be contacted (86 households in the control group and 93 households in the treatment group). This figure translates into an overall attrition rate of 8.5 percent.

Despite the low rate of attrition found in the TSSM sample, nonrandom attrition could still lead to biased inferences. In this section, we investigate nonrandom attrition by examining the characteristics of households that left the sample. We examine whether means of outcome and control variables for the attriting households differ across treatment and control communities.

As Table 3 shows, the attriting household characteristics are similar across treatment and control groups. Only one difference is statistically significant at the 5 percent level: the attriting treatment group has fewer household members under five years of age than the attriting control group. A small number of differences are significant at the 10 percent level: the percentage of household heads with senior high school education is higher in the attriting treatment group (p=0.057), as is dwelling ownership (p=0.09) and the use of concrete walls (p=0.05). All of these differences are relatively small and based on a small number of observations.

We also examine whether the households that dropped out of the sample differ from those that were found and surveyed for the endline. Table 4 suggests that the attriting households and panel households differ in some respects. Attriting households were less likely to have owned their dwelling at the baseline (73 versus 83 percent, p=0.001). They were also more likely to have a female head (10.6 versus 3.9 percent, p<0.000), and have more adults (p=0.004) and fewer children between ages six and 10 (but no difference in the number of children under age two). Attriting households were also slightly wealthier (p=0.098) and used more liquid petroleum gas (LPG) as cooking fuel (p<0.05).

As discussed earlier, we added households to the endline sample to attain a sample size of 2,500 households. We did this to increase the power of our endline tests. Table 4 further examines whether the added households differ from the panel households. Columns 3 – 5 report the results of tests of difference in means. Only a small number of differences are statistically significant at the 10 percent level. A lower percentage of the new endline households (20 percent) received

[6] See Annex 1 in *Scaling Up Rural Sanitation: Findings from the Impact Evaluation Baseline Survey in Indonesia*, Water and Sanitation Project Technical Paper, November 2010; http://water.worldbank.org/publications/scaling-rural-sanitation-findings-impact-evaluation-baseline-survey-indonesia.

BLT payments (that is, cash transfers from the government to poor households) than panel households (24 percent). We also observe a greater use of concrete walls and lower use of bamboo for new households. A higher proportion of new endline households had soap available at the handwashing station and had access to improved sanitation, but a lower proportion of them appropriately dispose of children's feces. Thus, Tables 3 and 4 suggest that the attrition was mostly random with respect to treatment assignment so doesn't lead our households in treatment communities to differ systematically from those in the control communities. Further, the panel sample does not appear to differ markedly from the original random sample so should be reasonably representative of the population.

7. PROGRAM EXPOSURE

During the study period, trained facilitators went into the eight project districts and supported the local governments in conducting triggering and follow-up activities in the communities. Of the 80 treatment villages, the endline survey data established that 53 villages (66 percent) were triggered (see Table 5).[7] The table also reports data from an earlier, additional short survey of community heads conducted in May 2010, which generated very similar figures to the endline data. Monitoring data collected as part of the TSSM program reports a higher percentage of treatment villages (83 percent) received the triggering.

Some control villages in the sample did receive the treatment. This was largely as a result of district governments changing their target communities after the randomization plan had been agreed upon. The endline survey data suggests that 13.8 percent of the control villages were exposed to the program (TSSM monitoring data gives a lower figure, at 4 percent).

The extent to which districts adhered to the randomization varied considerably.[8] Jombang, Bondowoso, Blitar, and (to a lesser extent) Madiun performed well in this respect. Banyuwangi and Probolinggo performed less well, with triggering detected in only 50 percent of treatment villages. Situbondo and Ngawi adhered the least to the randomization plan, with treatment detected in only four of the 10 treatment villages (40 percent). The contamination of control communities does not vary markedly across districts.

Next, we look at the extent of exposure to TSSM by households in treatment communities. Table 6 presents evidence of triggering activities as reported by households in the endline survey.[9]

---

[7]This means that one or more of the four community leaders surveyed (village head, community head, health cadre, head of women's organization, or other community leader) reported that the village had received a triggering.

[8]An independent study by Mathematica (2011) found similar results, in which program implementation varied heavily across the six districts they studied (three of the six districts are in our sample), depending upon capacity levels and cross-sectoral commitment to the program.

[9]This is in response to the question, "In the past two years, has there been a community event (such as a triggering/motivational talk) where sanitation was mentioned in this village?" Respondents were subsequently asked

Twenty-five percent of households in treatment villages reported that they knew of triggering activities being held in their community in the last two years, compared to 8.23 percent in control villages. The variation in these figures across districts is similar to that in Table 5. Bondowoso had the most households (42 percent) that were aware of the program in treatment villages. The table also indicates that participation in the triggering sessions differs across the districts.[10] While 30.13 percent of the households in treatment villages in Bondowoso attended the session, only 5 percent attended in Ngawi.

One of the key aims of the TSSM program is to increase awareness of good hygiene practices. To investigate this awareness, households were asked whether they had been exposed to sanitation campaigns through various media.[11] Figure 2a presents the results for any sanitation campaign (not necessarily TSSM). It shows that households received information about sanitation from many sources. Village health staff were an important conduit, as were TV campaigns. Exposure to sanitation messages was higher in treatment communities than in control communities, particularly via village staff (health staff and otherwise), billboards, and printed media.

Figure 2b shows the extent of exposure to messages that are explicitly about TSSM. Not surprisingly, the differences between treatment and control villages are much higher for TSSM-related messages. There was significantly greater exposure to TSSM sanitation messages via TV, printed media, village notice boards, and village personnel. TSSM-related exposure was much lower than the exposure to any sanitation program. For example, in treatment villages, 13 percent of households had heard about TSSM via village health staff whereas 36 percent of households had heard sanitation messages connected with any sanitation campaign in this way.

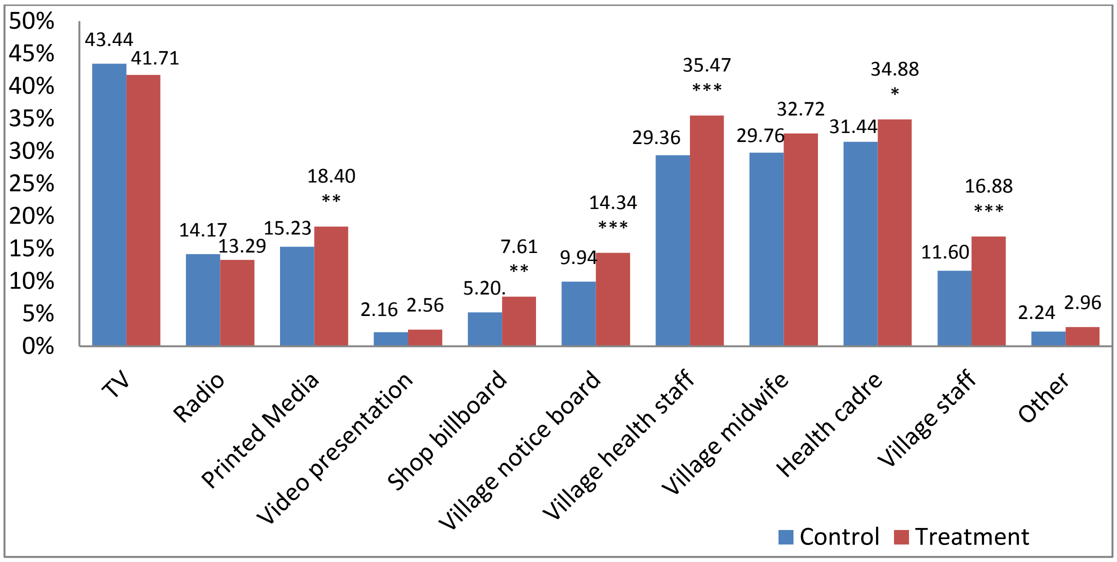
---

whether the triggering was a TSSM/STBM event. (STBM is another name by which TSSM is known.) Almost all respondents (89 percent) who answered that they knew of a triggering reported that it was a TSSM/STBM triggering.

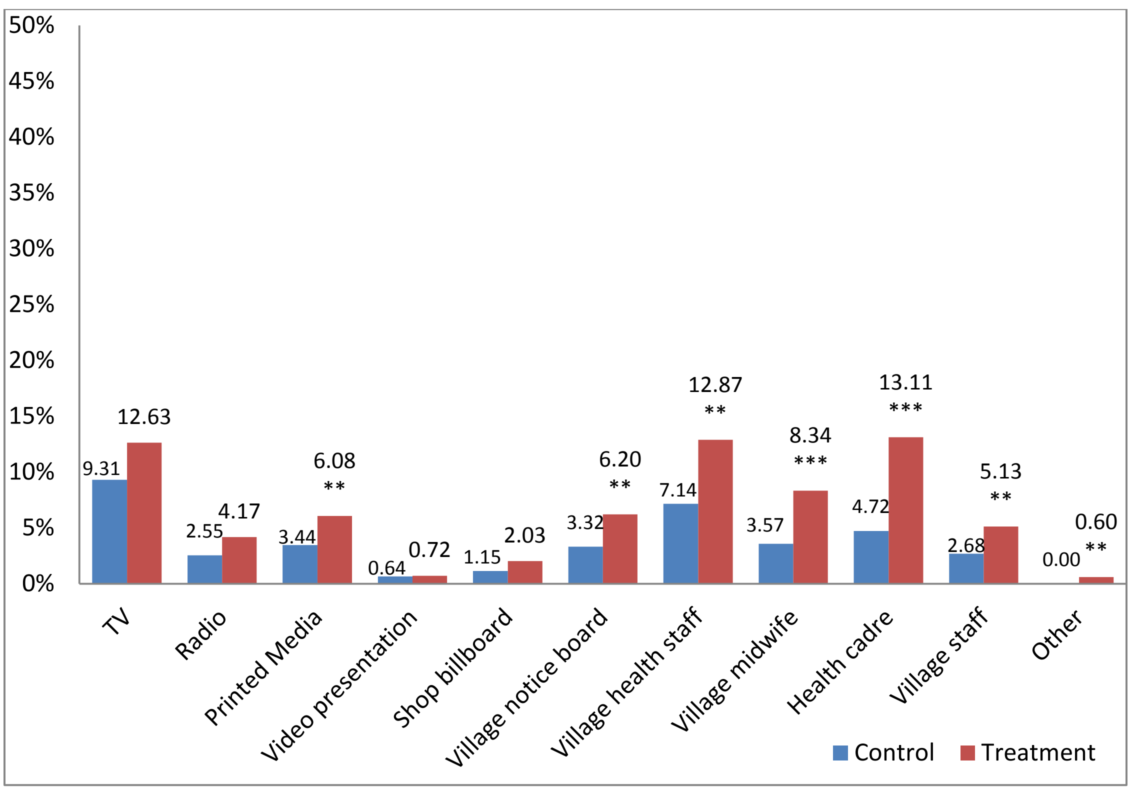
[10]If respondents indicated that there had been a triggering in their village, they were asked, “Did you participate in this community event (triggering/motivational talk) where sanitation was mentioned?”

[11]Specifically, households were asked, “In the past two years do you remember seeing, hearing, or reading anything about an invitation/encouragement to stop open defecation and to use or build yourown toilet/latrine?” They were then asked whether this had been: a) a TV advertisement; b) radio advertisement; c) poster/printed media/wall calender/pocket calender/leaflet of products/merchandise/banner; d) video presentation; e) on a shop nameplate; f) on the village announcement board; g) from a health officer; h) from a triggering officer; i) from a sanitarian; j) from a village midwife; k) from the maternal health cadre; l) from the village/sub-district staff; or m) other source. If they answered yes to any of these categories, they were asked whether the message was from the program called TSSM or STBM.

Figure 2. Household exposure to the sanitation program. (a) Percent of households that know about (any) sanitation campaign through various media and (b) whether it is TSSM related



(a)



(b)

Table 7 presents evidence of program exposure at the community level, in which the respondent is the village head (or in some cases a village staff member who is responsible for sanitation). Seventy percent of respondents in the treatment villages had heard about TSSM and more than 45 percent of the village heads in the treatment group attended the meeting and requested triggering activities. These numbers are all significantly higher than in control communities at the 1 percent level. Table 7 does, however, also show some contamination of control villages, with 11 percent of village heads in control villages reporting they had been to a triggering. Seventy-two percent of the village heads in the treatment villages are satisfied with the program.

The community section of the survey also asked village heads whether they helped households to gain quicker access to sanitation. In the treatment communities, 72 percent of the respondents stated that they did, mainly through building public toilets (45 percent), fundraising (17 percent), and providing lower-cost materials (9 percent).

In addition to surveying village heads, we interviewed community heads (sub-village heads or *kepala dusun)* about program exposure. The lower panel in Table 7 reports the results. In the treatment villages, 53 percent of the *kepala dusun* had heard about TSSM and 35 percent remembered the message (compared to 25 (p=0.000) and 18 (p=0.017) percent, respectively, in the control villages). Half of the treatment villages reported a triggering, with 38 percent reporting that the triggering was related to TSSM, and 38 percent of *kepala dusun* attended the meeting. The figures are much lower for the control villages (statistically significantly different at the 1 percent level). Of the 38 percent of sub-village heads who attended the triggering session, only 13 percent found the facilitator very persuasive, with 77 percent reporting that the facilitator was somewhat persuasive.[12] Sixty-two percent reported that they were satisfied with the program, and 75 percent reported that sanitation has improved in the community post-triggering. In only 11 percent of cases has the community been declared ODF.

8. DESCRIPTIVE STATISTICS COMPARISON BETWEEN BASELINE AND ENDLINE

Tables 8a and 8b provide descriptive statistics for key outcome variables along with control variables before and after program implementation. We first discuss the sanitation-related variables and then examine health outcomes.

8.1 Improved Sanitation and Toilet Construction

Table 8a presents descriptive statistics for sanitation and hygiene-related outcomes. The baseline data indicate that around 43 percent of the households in both the treatment and the control

[12] WSP (2012) found that poor-quality triggering decreased the likelihood of communities reaching ODF status and dampened communities' desire to change behavior toward better hygiene practices.

communities had access to improved sanitation.[13] Access to improved sanitation had increased only very slightly, by about 1 percentage point in both treatment and control communities by endline. Figure 3 presents the percentage of households with no sanitation, unimproved sanitation, and improved sanitation in the baseline and endline by treatment status. It similarly shows small improvements in sanitation in both treatment and control communities.

Figure 3. Households' sanitation status

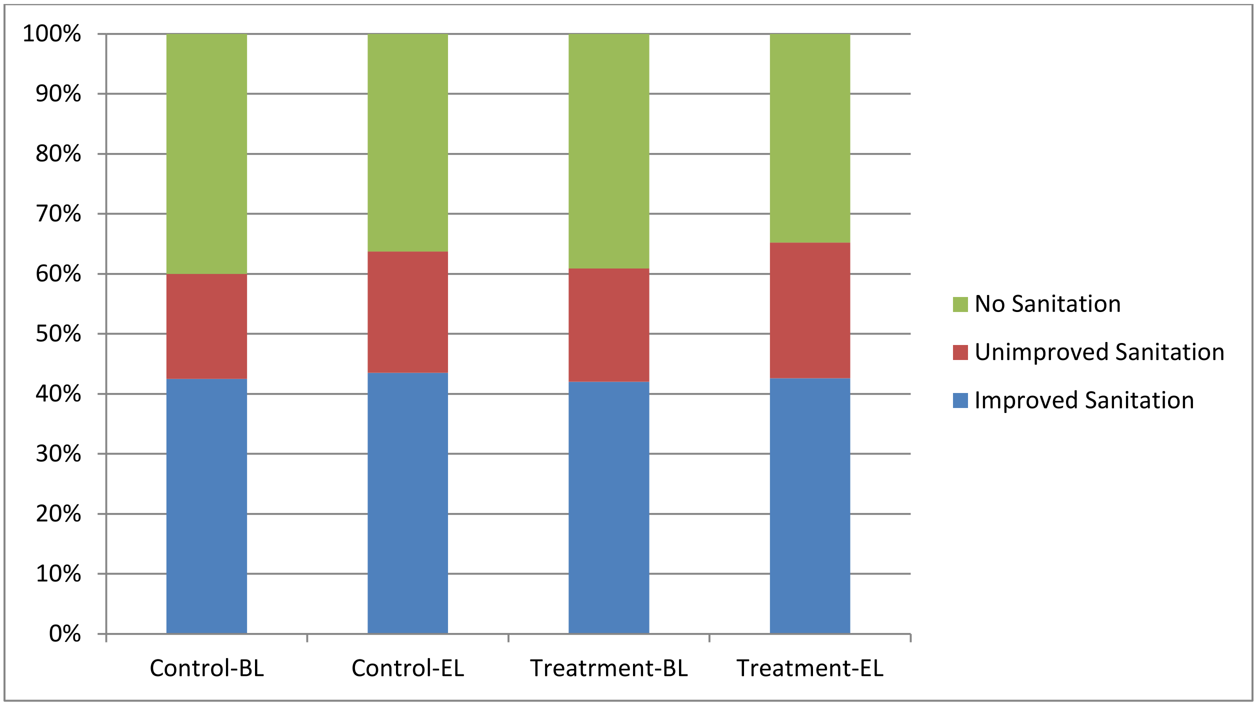


Table 8a also shows that although more than 40 percent of the households in the sample had access to improved sanitation before the program, another 40 percent of households lacked access to any form of sanitation. These households, many of which are particularly poor, defecate in open spaces. The percentage of households with no access to sanitation of any sort declined by about 4 percentage points between the baseline and endline. Again the decline in treatment and control communities appears similar.

The endline data specifically asked households whether they had constructed a toilet in the past two years. This variable is reported in Table 8a and indicates that more toilets were built in treatment communities than in control communities. Approximately 16 percent of households

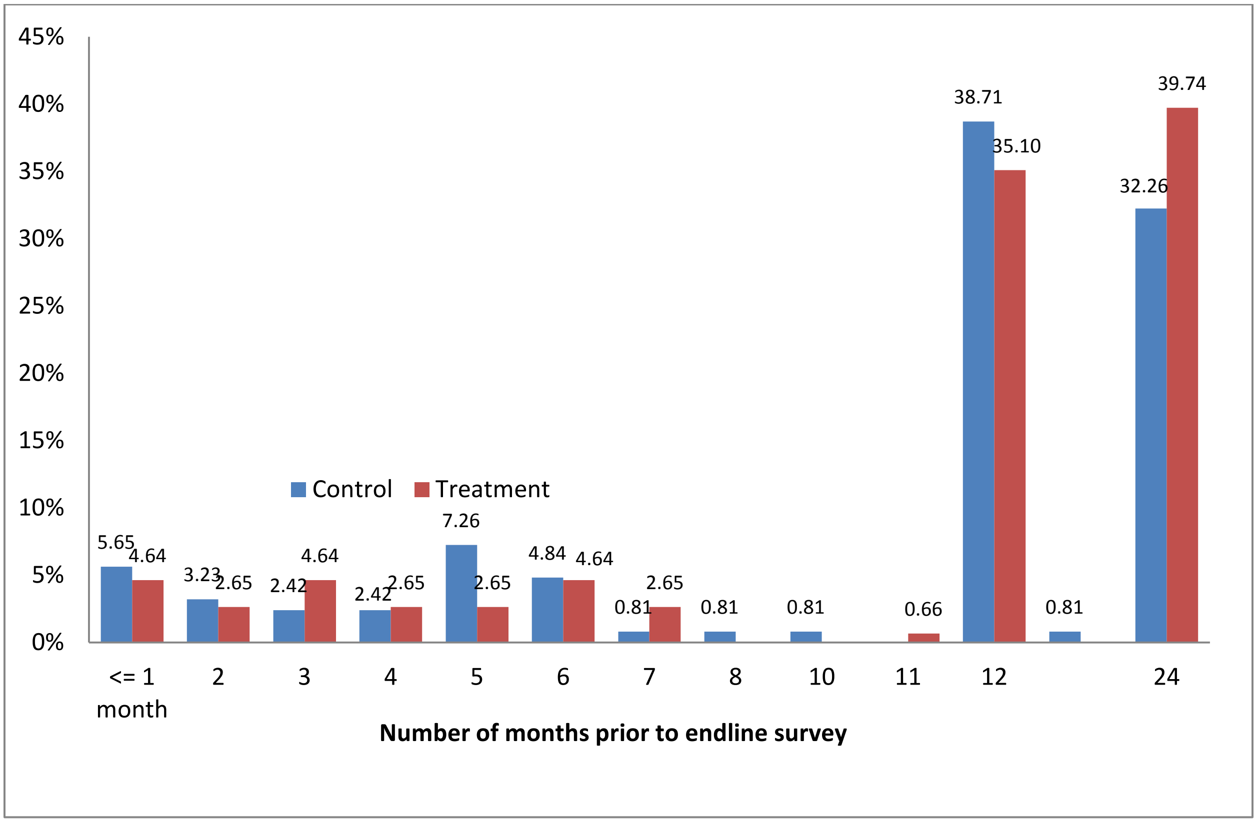
[13] Improved sanitation is defined according to the WHO/UNICEF (2012) Joint Monitoring Programme for Water Supply and Sanitation (JMP; http://www.wssinfo.org). Improved sanitation facilities include: a) a flush toilet or latrine that flushes to a sewer, septic tank, or pit; b) a ventilated improved pit (VIP) latrine; c) a pit latrine with the pit well covered by a slab; or d) composting toilets. Shared and public toilets are considered "unimproved" regardless of their type.

reported having constructed a toilet in the last two years in treatment communities versus 13 percent in control communities. This difference is statistically significant (*p*=0.07).

To reconcile the finding that the percentage of households with access to improved sanitation did not increase in treatment villages relative to control villages with the finding that more toilets were built in treatment villages, Table 9 examines toilet building more closely. It shows that 275 (14.4 percent) of the 1,908 panel households reported constructing a toilet in the last two years, with 151 (16 percent) of these toilets built in treatment communities, and 124 (13 percent) built in control communities. Not all construction resulted in a shift from "unimproved" to "improved" sanitation—102 of these toilets were built by households who already had access to improved sanitation. Only 180 of the households that built toilets improved their sanitation as a result. The number of households that built a toilet and consequently improved their sanitation is only very slightly higher in treatment communities than in control communities (9.7 versus 9.2 percent) and not statistically significant (*p*=0.73). Note, however, that health may improve as a result of toilets being built by households that already had access to "improved sanitation." For example, many of the households that had improved sanitation at baseline and had constructed a toilet by endline had upgraded to a porcelain latrine, which may be easier to keep clean and so more hygienic. Similarly, health benefits may accrue to households that don't have access to "improved" sanitation but have nevertheless gone from defecating in the open to using a pit toilet (without a concrete slab).

The endline survey also asked respondents when the toilet was built. Figure 4 reports the percentage of households that built a toilet prior to the endline survey, which was conducted between November 2010 and February 2011. Overall, a high percentage of households reported that they built their toilet between 12 and 24 months prior to the endline survey—almost 40 percent of the treated households reported that their toilet facility was built 24 months prior to the endline survey. The figure is slightly smaller for households in the control group (32 percent).

Figure 4. When did households build their toilets?



8.2 Open Defecation

Households were also directly asked whether household members ever engage in open defecation (on a daily, seasonal, or occasional basis). This question was asked regardless of the type of sanitation facility households had access to and where they reported they normally defecate. These data were reported separately for male household members, female household members, and children under five years of age, and are presented in Table 10. The table shows that 51 percent of households reported that at least one household member engages in open defecation. Women engage in open defecation slightly less often than men and children. Most open defecation in East Java occurs in rivers (95.36 percent of households that do not have access to sanitation report defecating in rivers). In villages that are within 20 minutes of a river, 54 percent of households had at least one member who defecates in the open at least occasionally. The figure is much lower (13 percent) for households in villages that are more than 20 minutes from a river. The fraction of households reporting that at least one member engages in open defecation is systematically lower in treatment communities than in control communities. The percentage of households reporting that a household member defecates in the open is 4.4 percentage points lower in treatment villages (significant at the 5 percent level). Among communities on rivers, open defecation by any household member is 6 percentage points lower in treatment communities than in control communities. The declines in open defecation are driven by declines in open defecation by people who have access to a toilet. Twenty-four percent of households that have access to a toilet reported sometimes defecating in the open; this figure is 4 percentage points lower in treatment communities than in control communities (26 versus 22 percent).

8.3 Handwashing

Improvements in good hygiene practices, especially handwashing with soap at critical times (before handling food and after defecating) can also help prevent childhood diarrhea (Ejemot et al., 2008) and other illnesses such as acute lower respiratory infection (Aiello et al., 2008). Although TSSM's message focuses on improvements in sanitation as the main vehicle for improving child health, the benefits of handwashing with soap were also discussed. The endline survey collected information on households' handwashing facilities and behavior.

Table 8a shows that almost 100 percent of respondents reported that they wash their hands after defecating. Of these, 67 percent reported having a handwashing station where they usually wash their hands. Observation by the survey enumerators of the handwashing station revealed that water was available at about 96 percent of the stations, and soap and water were available at about 58 percent of stations. These figures had increased from the baseline but there was no significant difference between treatment and control communities.

8.4 Drinking Water

Drinking water quality is another important determinant of child diarrhea and parasitic infection. The WHO defines improved drinking water sources as tap water inside the house and outside the house, protected dug well, boreholes (includes tube wells), protected springs, and rainwater collection (WHO, 2012). Table 8a shows that a large majority (more than 87 percent) of households had access to improved sources of water in both the baseline and the endline. There is no difference between treatment and control communities with respect to access to improved sources of drinking water. Note that we do not expect differences in drinking water outcomes given that this was not part of the treatment; however, it is interesting to note that differences in drinking water cannot be driving the health impacts we observe.

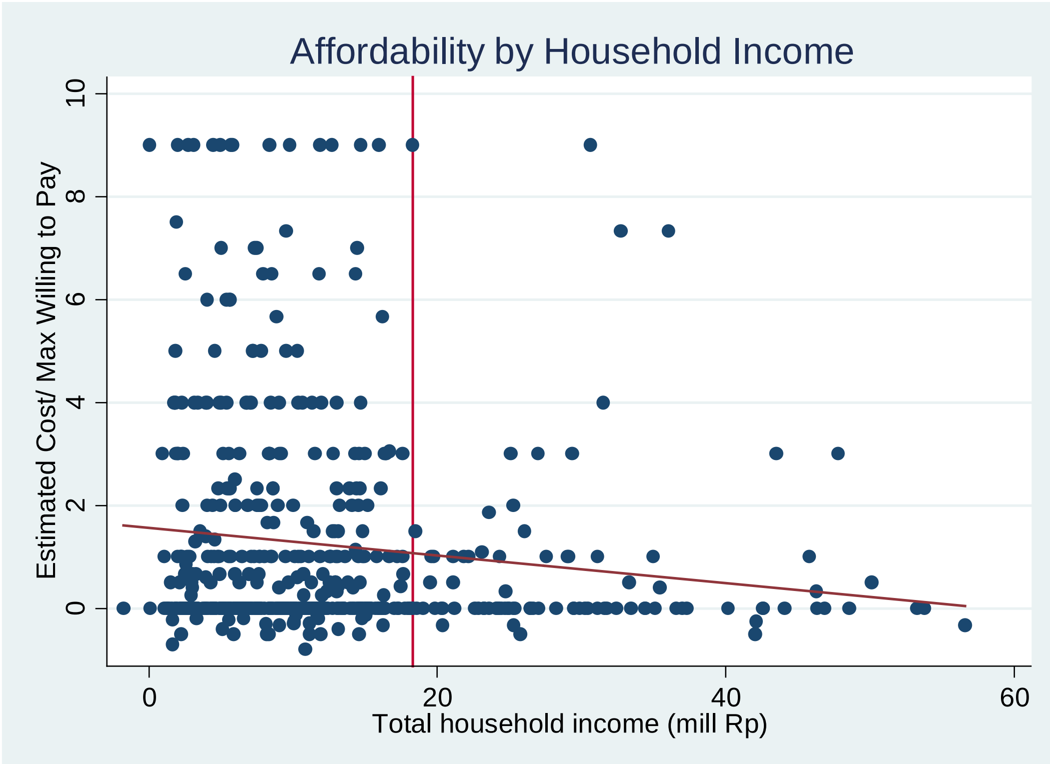
8.5 Attitudinal Change

TSSM seeks to change sanitation-related behavior via attitudinal change. The surveys collected information on people's attitudes about sanitation and the extent of their knowledge of the causes of diarrhea. Table 11 compares attitudes to open defecation in treatment and control communities. These questions were asked only in the endline survey. The vast majority of households (96.1 percent) agreed or strongly agreed with the statement that having a toilet benefits the community as it reduces environmental pollution. Similarly, an overwhelming majority (92 percent) agreed or strongly agreed with the statement that having a toilet protects their family against illness such as diarrhea. Further, 78.4 percent of households reported that having a toilet reduces the probability of being a topic of gossip, and only 6.9 percent of households agreed with the proposition that open defecation is acceptable because their ancestors practiced it. Nevertheless, the figures also show a tolerance of those who need to engage in open defecation. For example, 31.3 percent reported that it is acceptable to practice open defecation if you don't have a toilet, and only 34 percent reported that those who practice open defecation will be socially shunned. Table 11 disaggregates this information by treatment and control communities. There are no significant differences across the communities. Thus, it seems that knowledge of the health consequences of poor sanitation is relatively high across East Java, even in the absence of the program.

Table 12 examines caregivers' perceptions of the causes of diarrhea. These data were collected in both the baseline and endline surveys. Almost all caregivers (approximately 97 percent) thought that diarrhea is caused by eating dirty food or drinking dirty water. Lower proportions thought that it is caused by not washing hands (about 82 percent). About 86 percent reported environmental contamination as a cause, although only 72 percent believed that it is a consequence of others defecating in the open, and a smaller percentage (60 percent) believed it is a result of others defecating in rivers. The belief that diarrhea is caused by unclean food and water increased in both treatment and control communities but by significantly more in treatment communities. This is also the case for the belief that diarrhea is caused by an unclean environment. Increases in treatment villages relative to control villages in the belief that diarrhea is caused by some vaccines and exposure to the sun are more difficult to explain. All of the differences are small, which is perhaps to be expected when so many households seem to have been relatively well-informed about the causes of diarrhea before program implementation.

Households were also asked how much it would cost to build or improve their own toilet and to state the maximum that they would be willing to pay. Fifty-seven percent of households reported a figure equal to the figure they had reported for constructing the cheapest latrine possible. However, 38 percent of households reported a higher figure, indicating that they were only interested in building a more elaborate toilet. Figure 5 shows the relationship between perceived affordability and household income. Many households with household income below the mean reported that the amount they would have to pay is many times the maximum amount they were willing to pay. This is less common among higher income households.

Figure 5. Affordability to build or improve a toilet by household income



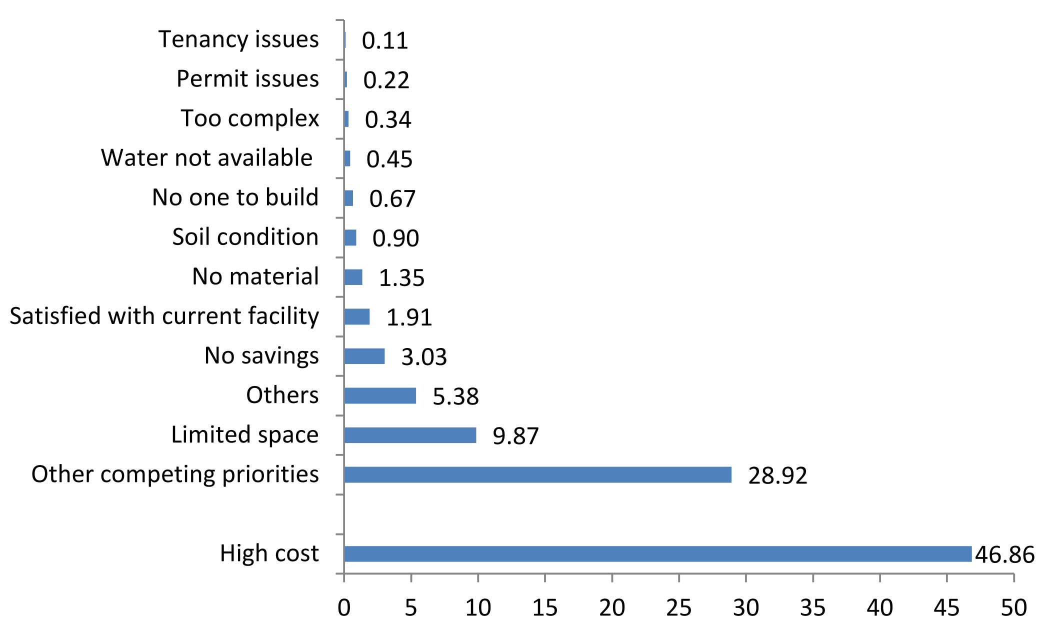
The surveys also collected information on respondents' level of satisfaction with their sanitation facility. One might expect treatment communities to become less satisfied with having no sanitation facility as a result of the program. The data show slightly more dissatisfaction among households without access to a toilet in treatment communities but the difference is small and statistically insignificant (29.2 percent in treatment communities versus 28.1 percent in control communities, p=0.74). Satisfaction with open defecation remains high: 55.6 percent of households that don't have access to sanitation reported that they were satisfied (38.8 percent) or very satisfied (16.8 percent) with their sanitation. This reflects the custom of defecating in rivers, which is thought to be hygienic because the waste is washed away. East Java is rich in rivers and 117 of our 160 communities (73 percent) are within a 10-minute walk from a river. Note that rivers are generally not used as a source of drinking water in Indonesia. Hence, a greater awareness that diarrhea is caused by dirty drinking water is not necessarily inconsistent with high levels of satisfaction with defecating in rivers.

8.6 Obstacles to Latrine Construction

Households that indicated they would like to build a latrine were asked what they consider to be the main obstacle to building one. Figure 6 presents the data from the endline survey. As reported in the baseline survey, the main obstacle was the cost of construction. TSSM's supply-side intervention involved training local masons[14] to build the kind of latrines that the local populations desired. Some districts also instituted schemes to make latrines more affordable (by providing access to credit, instituting rotating saving schemes with savings to be used for latrine construction (*arisan-jamban*), and mobilizing community labor to build latrines).

The implementation of the supply side of TSSM, however, lagged behind that of the demand side and was implemented on a much smaller scale. Table 13 shows how much households estimated it would cost to construct the cheapest possible latrine. On average, households estimated it costs approximately Rp 1.2 million to build a latrine (US$135). This is about one-third of the annual per capita income for the average household in our sample. TSSM estimates of the actual cost to install a latrine with a slab (exclusive of the superstructure) are in the range of US$50 to US$90 (WSP, 2012). There is no statistically significant difference in the costs reported in treatment communities and control communities. The majority of households (64 percent) reported that these costs rose in the last 12 months, in both treatment and control communities. There are also no differences between treatment and control communities with respect to access to materials and tradesmen. Slightly more people in control communities reported that they do not know who to contact to obtain materials and tradesmen (4.3 versus 3.1 percent, $p<0.05$). Slightly more people in treatment communities reported that the costs are affordable, but this difference is not significant.

Figure 6. Main reason for not building a toilet



[14]Later, this was replaced by capacity building of more full-fledged sanitation entrepreneurs who employ masons.

To summarize, the results to this point show that the number of toilets built in treatment communities since the program's implementation was higher (3 percentage points) than in control communities. This is a 31 percent increase in the rate of toilet construction in treatment communities relative to control communities. There is evidence of declines in occasional open defecation by households that have access to a sanitation facility. Significantly fewer households in treatment communities reported that one or more household members defecate in the open (4.4 percentage points fewer). Knowledge of the health benefits of improved sanitation seems to be high across East Java, even in control communities. And the program is associated with increased acknowledgement that unclean food and water leads to diarrhea. However, defecating in rivers is still an accepted practice. The main obstacle to constructing facilities as reported by households seems to be the cost.

8.7 Prevalence of Diarrhea, Acute Lower Respiratory Infection (ALRI), Helminth Infections, and Acute Respiratory Infection (ARI)

Now we turn our attention to health outcomes. Table 8b presents the summary statistics for the health variables. Diarrhea was defined using the WHO definition (three or more stools per day and the stools were loose or watery, or blood and/or mucus is visible in the stool (Baqui et al., 1991) based on symptoms reported by the child's caregiver. There was a decline in diarrhea prevalence between the baseline and endline surveys.[15] Diarrhea prevalence was balanced in the baseline. In the endline, diarrhea prevalence in the treatment group was lower than that in the control group, regardless of the recall period used. Using a recall period of seven days, 2.4 percent of children under the age of five in the treatment group reported having had diarrhea in comparison to about 3.8 percent in the control group. The same pattern is observed if we use the shorter recall period of two days. These differences are statistically significant.[16]

Soil transmitted helminthes (*Ascaris lumbricoides, Trichuris trichuria,* hookworm) are estimated to infect about one-third of the world's population, and a significant share of infection is born by preschool-aged children (Albonico et al., 2008). Helminth infections can lead to iron-deficiency anemia, protein-energy malnutrition, abdominal pain, and listlessness (Adams et al., 1994). To examine this, stool specimens were collected and analyzed at endline from a random subsample of 2,330 children in the 80 villages. Each child in the parasitological sample was given a plastic container (a stool specimen container, with a spoon attached to the lid) and asked to provide a stool sample. The container had 15 ml of preservative (formalin) in it, unless the child was reported to have diarrhea, in which case there was 30 ml of formalin in it. Caregivers were told

---

[15] This decrease at least in part reflects the aging of the children as these are figures for children who were present during both the baseline and endline surveys. Diarrhea prevalence is higher in children under the age of two. It may also be due to the different timing of the two surveys. The endline survey was conducted in the rainy season, whereas the baseline was conducted at the end of the dry season.

[16] Shorter recall periods may be a more accurate measure of diarrhea prevalence because they involve less recall error.

that specimens must not be contaminated with urine or water. They were also told that the preferred time for specimen collection was early in the morning. Caregivers were instructed to scoop enough stool to fill approximately one-third of the container (about 5 grams for solid stool and 10 ml for liquid stool, which was about two full scoops), screw the cap on tightly, and shake well. The sample was then transferred to a lab. Following Katz (1972), reported infection intensity was proxied by worm eggs per gram of stool.

Table 8b reports the results from the parasitology analysis. Worm infection was relatively low in our sample, with only 4 percent of the surveyed children testing positive for at least one of the following: worms (*Ascaris lumbricoides, Trichuris trichuria*) and hookworms (*Ancylostoma duodenale* and *Necator americanus*). We found no significant difference in the probability of having any eggs between the children in the treatment and control groups. Children in the treatment groups appear to have lower infection intensity than those in control groups. However, the difference is not statistically significant. Table 8b also provides information on the prevalence and intensity of infection by organism. Prevalence is highest for *Ascaris* (3 percent). Hookworm prevalence is 0.5 percent, and *trichuris* prevalence is 0.1 percent. We compare infection prevalence categorized by low, moderate, and high infection intensity using the WHO (2002) thresholds for fecal egg count per gram (epg).[17] As before, we find no significant difference in infection intensity for each organism in the treatment and control groups.

Acute respiratory infections (ARI) are among the leading causes of death in children under five years old. Following the WHO definition, a child is identified as having ARI when he or she presents with a persistent cough or difficulty breathing. A meta-study conducted by Roth et al. (2008) found that throughout the world, 1.9 million children died from ARI in 2000; 70 percent of these deaths were in Africa and Southeast Asia. Acute lower respiratory infection (ALRI) is a more serious condition and is diagnosed when a child has a constant cough or difficulty breathing (ARI) and a raised respiratory rate. A number of factors affect ARI and ALRI rates in young children, including malnutrition, lack of breastfeeding, and the incidence of other diseases that affect susceptibility. The child's environment (such as crowding and air pollution) can also affect the risk of getting ALRI. Better sanitation in the form of frequent handwashing can reduce the spread of the infection from person to person.

As Table 8b shows, using either a seven- or two-day recall period, ARI and ALRI prevalence are slightly higher for children in the treatment group than for children in the control group, although the differences are small and not statistically significant. ALRI prevalence in the sample is low (around 2 percent of children were found to have had ALRI in the past seven days).

[17] The infection thresholds for *Ascaris lumbricoides* are low: 1–4,999 eggs per gram (epg); moderate: 5,000–49,999 epg; and heavy: >50,000 epg. For *trichuris trichuria*, infection thresholds are low: 1–999 epg; moderate: 1,000–9,999 epg; and heavy: >10,000 epg. For hookworm, thresholds are low: 1–1,999; moderate: 2,000–3,999; and heavy: >4,000 epg (WHO, 2002).

Table 8b shows similar comparisons for a range of other symptoms of child illness. Changes in health symptoms that are unlikely to be affected by the treatment can be used to confirm that the observed differences between the treatment and control communities are not attributable to systematic reporting bias. Bruising/abrasions and itchy skin/scalp are two symptoms that are unlikely to be affected by the intervention. Table 8b shows that the differences in these symptoms between control and treatment communities are small and statistically insignificant.

Indeed, most of the reported symptoms show no difference between the control and treatment groups in the endline. However, the prevalence of mucus or blood in the stool is significantly lower for children in the treatment group (0.8 percent *cf.* 2.0 percent). Refusal to eat is also lower in the treatment group. Apart from these symptoms, only the prevalence of nasal congestion differs between the control and treatment communities. It is higher in the treatment communities.

8.8 Anemia, Anthropometric, and Developmental Measures

Poor sanitation leads to diarrhea and diarrhea can increase the prevalence of anemia because it reduces the absorption of nutritional intake, including the absorption of iron, which is used in the production of hemoglobin. Any child with hemoglobin levels below 110 g/L is considered anemic as per WHO guidelines for children under five years of age (UNICEF, WHO, and United Nations University, 2001; WHO, 2011). Hemoglobin levels rose during the study period, consistent with the children having aged between baseline and endline. This translates into a much lower proportion of children with anemia at the end of the study period than at the beginning (44.2 versus 71.2 percent). However, there is no evidence of great declines in treatment communities vis-à-vis control communities. There is no significant difference in hemoglobin levels and anemia between the control and treatment groups.

Anthropometric measurements (weight, height, arm and head circumference, body mass index) were taken for all children under five years old at the time of the baseline in the surveyed households. Child anthropometric measurements provide useful information about nutritional intake and so are adversely affected by illnesses such as diarrhea. Weight (conditional on height and gender) typically varies in the short-term so is used as a measure of current health status whereas height, given age and gender, is an indicator of longer-term health and welfare (Thomas, et al., 1991).

To assess the child's growth and general nutritional status, we used a standardized age- and gender-specific growth reference based on the WHO (2006, 2007) standard to calculate *z*-scores for a) weight-for-height, b) height-for-age, c) weight-for-age, d) body-mass-index, e) head circumference-for-age, and f) arm circumference-for-age.[18] All of the child growth measures

[18] For example, a *z*-score for height subtracts from the child's height, the median height in the reference population, for a child of the same gender and age in months, and divides by the standard deviation of height in the reference population, also for a child of the same gender and age in months. A weight-for-height *z*-score is defined in an analogous manner, except that the standardization is done using the reference population median and standard

(with the exception of upper arm circumference in the baseline) tended to be lower than the population mean using the WHO standard (i.e., have negative *z*-scores). All of these measures decreased between the baseline and the endline, reflecting the aging of the cohorts and the growth faltering that occurs during the first two years of life (Victora et al., 2010). There are no statistically significant differences in the anthropometric measures between the treatment and control groups.

8.9 Child Development

Both the baseline and endline surveys included an "Ages and Stages" questionnaire, which measures child development across five domains: communication, gross motor skills, fine motor skills, problem solving, and personal-social development. At the baseline this questionnaire was administered to parents or caregivers of children ages 3 – 24 months. In the endline survey it was extended to children up to 60 months of age. This report focuses on the communication, motor skills, and personal-social development. These measures are of interest because a child's health status might affect his or her development. The questions asked about each child were selected according to the child's age in months. To make a comparison, the child development index for each skill is standardized by calculating *z*-scores.[19] Table 8b shows greater improvement in communication (significant at the 10 percent level) for children in treatment communities.

To summarize, a comparison of raw outcomes in treatment and control communities shows greater toilet construction in treatment communities and declines in open defecation. In line with this, there is evidence of decreases in diarrhea prevalence in treatment communities, relative to control communities. The prevalence of mucus or blood in the stool has also decreased. Health outcomes further along the causal chain appear to have been largely unaffected by the TSSM program.

9. ESTIMATION STRATEGY AND RESULTS

In this section, we estimate the "intention-to-treat" (ITT) of program impact. The ITT approach estimates the program impact from a comparison of those communities to which the program has been offered, regardless of whether they actually receive it, with control communities. The ITT allows us to determine a program's average impact on the population it is targeting and is often used in the context of randomized controlled trials (Duflo et al., 2008). The ITT estimator is computed by comparing the mean outcome for the treatment group to that for the control group, which can be estimated using the following regression equation:

---

deviation of weight for children of a given gender and height. The WHO standards use an international reference population.

[19] *Z*-scores standardize the measures so they have a mean of zero and a standard deviation of 1. The different measures are thus expressed in the same scale and so can be compared across different age groups.

$$Y_{ij} = \alpha + \beta T_{ij} + \gamma_j + \varepsilon_{ij} \quad (1)$$

where Y_{ij} is the outcome for observation *i* in village *j* conditional on treatment; T_{ij} is the treatment dummy, which is 1 for the households or children in the treatment group, and 0 otherwise; and ε_{ij} is the error term that is correlated within villages, given the cluster-RCT design.

If one only controls for treatment status on the right-hand side of the equation, ITT generates the same results as the differences in means presented in the previous section. The ITT regression approach, however, also allows us to control for other variables. In the specifications reported next, we add a set of pairwise sub-district dummy variables (γ_j) to control for sub-district effects given that the randomization was assigned such that in each of the participating districts, 10 (treatment and control) pairs of villages from the same *kecamatan* (sub-district) were chosen.

In some specifications, we also add lagged dependent variables (i.e., outcomes as observed in the baseline $Y_{ij[BL]}$) and a vector of household characteristics known to influence sanitation practices and health outcomes (X_i) as additional right-hand side controls.

$$Y_{ij} = \alpha + \beta T_{ij} + \tau Y_{ij[BL]} + \delta X_i + \gamma_j + \varepsilon_{ij} \quad (2)$$

Here, β is the parameter of interest and represents the effect of the TSSM program on sanitation/health outcomes, conditional on **X** and the baseline outcome variables; and τ reflects the degree to which health/sanitation outcomes are correlated over time. All specifications also allow for village-level clustering of the standard errors.

Finally, in some specifications we estimate the model conditional on whether households had sanitation at baseline. Doing so enables us to examine the effect of behavioral change on those that already have toilets (using them more); as well as those who do not have toilets (using public facilities more). In addition, we further disaggregate households at baseline who did not have sanitation facilities by poor and non-poor households. This can help us better understand whether there was differential construction of toilets and/or behavioral change due to income differences. We can also detect any health improvements resulting from the behavioral changes toward better hygiene practice.

Results

9.1 Program Effects on Sanitation and Hygiene Behavior

Table 14 presents the estimated impact of the TSSM program on sanitation. It presents means of the dependent variable in the control group in column 1 as an aid to interpreting the magnitude of the effects. Column 2 reports the result of estimating Equation (1) with pairwise sub-district fixed effects and shows a moderate increase in the rate of toilet building in the last two years in the treatment households (4 percentage points higher for households residing in treatment districts compared to the control group). Column 3 restricts the sample size to panel households. The point estimate still shows greater toilet building in treatment communities, although it is now slightly smaller. The point estimate remains similar in magnitude and significant once we control for baseline sanitation (column 4)[20] and various exogenous controls (columns 5 and 6). Columns 7 and 8 show the estimate conditioning on whether households had a toilet at baseline. TSSM is associated with a 4 percentage point significant increase in toilet construction among households that did not have a toilet at baseline (column 7). The increase in toilet construction is smaller (1 percentage point) for households that already had a toilet at baseline and is not statistically significant (column 8).

The other rows in the table show the same analysis for improved sanitation, where a household normally defecates, whether any householder defecates in the open even occasionally, whether the household correctly disposes of children's feces, and whether the household has access to improved drinking water. We find that access to improved sanitation increased significantly in treatment communities for those who did not have a toilet at baseline. Open defecation also decreased significantly in treatment communities among households that had no sanitation at baseline (5.8 percentage point reduction from a 48 percent base). We see no such decrease among those who had access to a toilet at baseline. We also find a significant improvement in the disposal of children's feces among households that had no sanitation at baseline. We find no improvements in behavior among those who did have sanitation at baseline. This suggests that the behavioral changes were driven by toilet construction.

Table 15 reports the results for handwashing behavior. Results are presented in the same order as in Table 14. There is no evidence of improvements in handwashing behavior, which is not surprising given that handwashing was not a key component of the intervention. The only significant coefficient suggests that treatment is associated with a decline in water availability at handwashing stations.

The main obstacle to constructing a toilet reported by households was the cost. Credit constraints likely limited households' ability to build a toilet. To investigate this, we estimate the model separately for poor and non-poor households. A household is defined as poor if they are in the bottom quintile in terms of the value of their non-land assets (less than Rp. 1,975,000). These households are actually "particularly poor." Note that at baseline 36 percent of our sample fell below the 2008 rural Indonesian poverty line of Rp 161,831 per capita per month. Our results suggest that the treated non-poor households are building toilets, not the poor (Table 16). This suggests that provision of credit or subsidies to the particularly poor households may work to increase the project impact. Table 16 also shows that the program increased access to improved

[20] For the building toilet variable, we used access to improved sanitation in the baseline as our baseline control variables.

sanitation for non-poor households that didn't have access to any sanitation at baseline. This group also showed improvements in the disposal of children's feces.[21]

9.2 Program Effects on Child Health Outcomes

Table 17 reports the ITT effects for child health outcomes. Consistent with the data presented in Table 8b, we find a statistically significant reduction in diarrhea prevalence (seven- and two-day prevalence). Specifically, being in a treatment community is associated with diarrhea prevalence being lower by approximately 1.3 percentage points from a 4.6 percent base. This is the case across all specifications on the panel observations (columns 3 – 6) and for seven- and two-day prevalence. This is a large effect—about a 30 percent (44 percent) reduction in seven-day (two-day) diarrhea prevalence. The fact that the estimated program effect changes only slightly in magnitude as we add controls is unsurprising given that the TSSM survey provides a balanced sample.[22] Columns 7 and 8 show that this is being driven by decreases in diarrhea among households that did not have sanitation at baseline.

We also find some evidence that parasitic infection intensity for both *Ascaris* and *Trichuris*, as proxied by number of egg worms per milligram, is lower in treatment communities. However, this result is not statistically significant over the whole sample.[23] Having blood or mucus in the stool is also significantly lower in treatment communities by about 1 percentage point (57 percent) as well as refusal to eat for children in treatment communities.

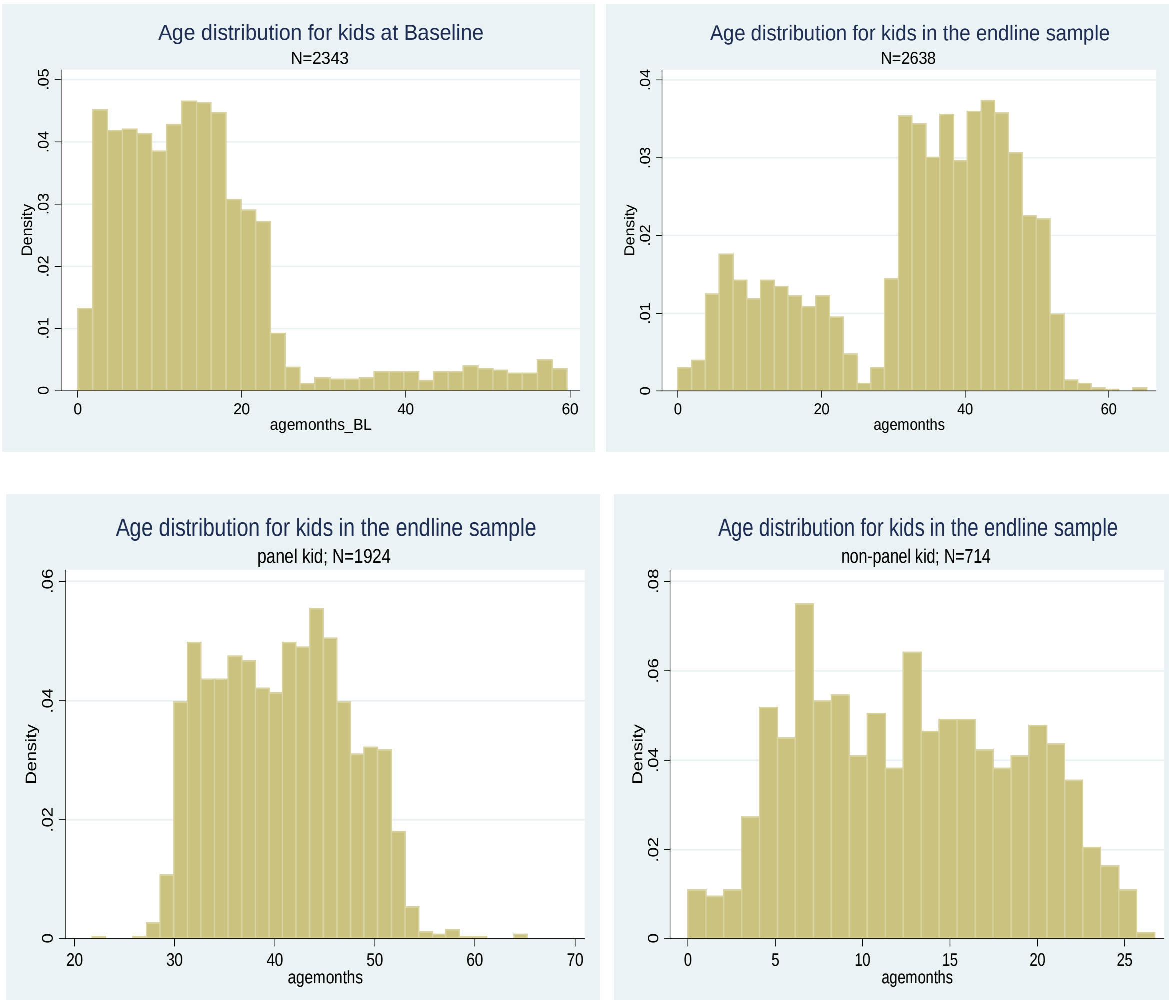
---

[21] Additional one-tailed tests of significance (not reported in Table 16) show a significant decrease at the 10 percent level in the practice of open defecation for the treated non-poor households with no sanitation at baseline.

[22] The estimate in column 2 uses the entire endline sample (not restricted to panel households) and although the point estimate is negative, it is weakly significant for seven-day prevalence ($p<0.1$), while it is statistically insignificant for two-day prevalence. Note that the use of the full sample involves adding children under the age of two, whereas all children in the panel sample are over age two as they had to have been born at the time of the baseline. This is shown in Figure 7. If we re-estimate these regressions only on children under age two, we find no treatment effect on diarrhea prevalence. Hence, the weakly significant and insignificance of the full sample results may be being driven by the younger age of the children. It may be that the treatment has a reduced impact on diarrhea of very young children because children under age two are unlikely to use a new facilitation facility themselves. In contrast, children between ages two and five would be taking themselves to the toilet so may be more likely to directly benefit from new facilities.

[23] We also estimated the model with the dependent variable being the prevalence of each of *Ascaris, Trichuris*, and hookworm categorized using the WHO (2002) thresholds into low, moderate, and heavy infection. We find no significant treatment effect.

Figure 7. Age distribution comparisons for panel and non-panel children



Columns 8b and 8c estimate the health impacts separately for poor and non-poor households that did not have sanitation at baseline. Diarrhea prevalence decreased in both of these subgroups, although by more for poor households. This is curious because non-poor households constructed more toilets. It may be that wealthier households are partially protected by other aspects of their environment that reduce the probability of getting diarrhea—for example, cleaner food-preparation areas. Note, however, that a one-tailed test of significance (not reported in the table) shows a weakly significant (at the 10 percent level) reduction in the seven days of diarrhea prevalence for non-poor household with no sanitation at baseline. A one-tailed test of significance also finds a decrease in the occurrence of mucus or blood in the stool samples for this group. Further, parasitic infection (*Alascaris*) decreased significantly in treatment communities for the non-poor who had no sanitation at baseline.[24]

Congestion and breathing difficulty are significantly higher in treatment communities for non-poor households that did not have sanitation at baseline, for which we have no explanation.

When examining anthropometric and developmental impacts, we do not use *z*-scores because we are concerned about the propensity for measurement error in age (which is considerable in Indonesia) to affect the results. Instead, we use the raw measures and include age dummy variables (defined in terms of three-month age blocks) on the right-hand side. Increases in weight, height, and weight-for-height were detected for non-poor households that had no sanitation at baseline (again, these are the households that were most likely to build a toilet as a result of the program). Although Table 17 does not indicate significant effects for poor households with no sanitation at baseline in terms of weight and height, it shows that treatment is associated with a higher BMI for this group. This may be an artifact of lower average height in treatment communities (insignificantly so). The estimates for head circumference are insignificant. Arm circumference indicates a negative treatment effect but this becomes insignificant once we break the sample down by whether or not households had sanitation at baseline.

Communication skills are positively impacted and social skills negatively impacted in some specifications but overall we find no systematic program impact on the development indicators.

9.3 Discussion of Results

We find evidence of increased toilet construction and reduced open defecation in treatment communities relative to control communities. These improvements were largely among households that had no sanitation at baseline and were not “particularly poor” and so could afford to construct a toilet. We also find improvements in health among these households, although some health benefits seem to extend to poorer households that also had no sanitation at baseline.

The estimated reductions in diarrhea are large—1.4 percentage points or 30 percent—whereas the increases in toilet construction and reduction in open defecation practices are more moderate. Two questions that naturally arise are how plausible are the diarrhea results and are they truly a consequence of the program. The results disaggregated by sanitation status support the argument that the health improvements are a consequence of the program. The health improvements largely accrue to those who did not have sanitation at baseline and so benefitted the most from toilet construction.

[24]Note that if we disaggregate by proximity of the village to the river as in Table 10, we find that the decreases in diarrhea are driven by decreases in treatment villages that are on rivers. This is consistent with Table 10’s finding that open defecation decreased most in villages on rivers.

We cannot rule out that the difference in caregiver-reported diarrhea prevalence between groups could be an artifact of biased outcome reporting. However, we do not observe differences between groups in falsification symptoms that could be subject to similar bias. Moreover, the results from the fecal samples (arguably more objective data than caregiver-reported diarrhea prevalence) show benefits accruing even more narrowly to non-poor households that had no sanitation at baseline, which are the group most likely to have built toilets as a result of the program. The weight and height improvements are also restricted to this group.

Positive externalities are another potential explanation for the large estimated impact on diarrhea prevalence. Perhaps only small improvements in sanitation are needed to give rise to large health effects given the benefits that flow across the community as a result of a less contaminated environment. The benefits of a household's new toilet extend beyond that household's. This is an area we have not fully explored in this report but is an area for future research.

10. CONCLUSION

In conclusion, knowledge of the link between poor sanitation and health seems to be high in East Java even in control communities. Nevertheless, open defecation is still tolerated and many households believe that other people defecating in the open has no bearing on their own household's health. Defecating in rivers is still a relatively attractive option for many households. A qualitative study conducted by WSP (2012) found that communities located near rivers are less likely to achieve ODF status. Households in these areas considered defecating in the river superior to using a pit latrine. This suggests that specific strategies focused directly at communities that are in close proximity to water bodies might be needed.

Nevertheless, we find that TSSM is associated with sanitation improvements. More households built toilets in treatment communities (3 percentage points higher) than control communities. This is a 31 percent increase in the rate of toilet construction in treatment communities relative to control communities. Further, fewer households reported that household members ever defecate in the open in treatment villages (4.4 percentage points lower than in control communities). The sanitation improvements were largely driven by toilet construction by non-poor households. Sanitation improvements were more limited for "particularly poor" households. Poor households are more likely credit-constrained and are thus less able to build toilets and so improve their access to sanitation.

In terms of health outcomes, we detect significant declines in the prevalence of diarrhea among young children (between two and five years old). Diarrhea prevalence decreased by 1.5 percentage points more (a decline of 30 – 45 percent depending on whether we examine seven- or two-day prevalence) in treatment communities than in control communities. This is a large impact. We also find decreases in the intensity of parasitic infestation, and increases in height and weight among non-poor households that had no sanitation at baseline.

REFERENCES

Adams, E., Stephenson, L., Latham, M., and Kinoti, S. (1994). Physical activity and growth of Kenyan school children with hookworm, *Trichuris trichiura* and *Ascaris lumbricoides* infections are improved after treatment with Albendazole. *Journal of Nutrition* 124:1199-1206.

Aiello, A., Coulborn, R., Perez, V., and Larson, E. (2008). Effect of hand hygiene on infectious disease risk in the community setting: A meta-analysis. *American Journal of Public Health* 98:1372-1381.

Albonico, M., Allen, H., Chitsulo, L., Engels, D., Gabrielli, A.-F., and Savioli, L. (2008). Controlling soil-transmitted *helminthiasis* in pre-school-age children through preventive chemotherapy. *PLoS Neglected Tropical Diseases* 2: e126.

Baqui, A., Black, R., Yunus, M., Hoque, A.R., Chowdhury, H., and Sack, R. (1991). Methodological issues in diarrheal diseases epidemiology: Definition of diarrheal episodes. *International Journal of Epidemiology* 20:1057-1063.

Baqui, A., Black, R., Sack, R., Chowdhury, H., Yunus, M., and Siddique, A. (1993). Malnutrition, cell-mediated immune deficiency, and diarrhea: A community based longitudinal study in rural Bangladeshi children. *American Journal of Epidemiology* 137: 355-365.

Cameron, L. and Shah, M. (2010). *Scaling Up Rural Sanitation: Findings from the Impact Evaluation Baseline Survey in Indonesia*. Washington, DC: Water and Sanitation Program, World Bank.

Curtis, V. (2004). Handwashing, hygiene and health: Proposals for strengthening WSLIC-2's health component and a national hand-wash campaign. London: Hygiene Centre, London School of Hygiene and Tropical Medicine.

Duflo, E., Glennerster, R., and Kremer, M. (2008). Using randomization in development economics research: A toolkit. *Handbook of Development Economics.* Vol. 4. Schultz, T. and Strauss, J., eds. Amsterdam: Elsevier: 3896-3903, 3911-3917, 3928-3933.

Ejemot, R.I., Ehiri, J.E., Meremikwu, M.M., and Critchley, J.A. (2008). Handwashing for preventing diarrhea, *Cochrane Database of Systematic Reviews* 1. Article CD004265; doi: 10.1002/14651858. CD004265.pub2.

Guerrant, D., Moore, S., Lima, A., Patrick, P., Schorling, J., and Guerrant, R. (1999). Association of early childhood diarrhea and cryptosporidiosis with impaired physical fitness and cognitive function four-seven years later in a poor urban community in Northeast Brazil. *American Journal of Tropical Medical and Hygiene* 61: 707-713.

Katz, N., Chaves, A., and Pellegrino, J. (1972). A simple device for quantitative stool thick-smear technique in *Schistosomiasis mansoni*. *Revista do Instituto de Medicina Tropical de São Paulo* 14: 397-400.

Mathematica (2011), *Improving Sanitation at Scale: Lessons from TSSM Implementation in East Java, Indonesia*. New Jersey: Mathematica Policy Research.

Medley, G., and Anderson, R. (1985). Density-dependent fecundity in *Schistosoma mansoni* infections in man. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 79: 532-534.

Miguel, E., and Kremer, M. (2004). Worms: Identifying impacts on education and health in the presence of treatment externalities. *Econometrica* 72: 159-217.

Ministry of Health. (2002). *Indonesia Health Profile*. Ministry of Health, Jakarta.

Mukherjee, N., with Robiarto, A., Saputra, E., and Wartono, D. (2012). *Achieving and Sustaining Open Defecation Free Communities: Learning from East Java*. Jakarta: Water and Sanitation Program, World Bank.

Murray, C., and Lopez, A. (1997). Global mortality, disability and the contribution of risk factors: Global burden of disease study. *Lancet* 349: 1436-1442.

Napitupulu, L., and Hutton, G. (2008). *Economic Impacts of Sanitation in Indonesia*. Jakarta: Water and Sanitation Program, World Bank.

Roth, D., Caulfield, L., Ezzati, M., and Black, R. (2008). Acute lower respiratory infections in childhood: Opportunities for reducing the global burden through nutritional interventions. *Bulletin of the World Health Organization* 86: 356-364.

Schneider, R.E., Shiffman, M., and Faigenblum, J. (1978). Potential effect of water on gastrointestinal infections prevalent in developing countries. *American Journal of Clinical Nutrition* 31: 2089-2099.

Thomas, D., Strauss, J., and Henriques, M. (1991). How does mother's education affect child height? *Journal of Human Resources* 26: 183-211.

United Nations Children's Fund, World Health Organization, and United Nations University. (2001). *Iron Deficiency, Anemia: Assessment, Prevention and Control, a Guide for Programme Managers*. Geneva: World Health Organization.

Victora, C. G., de Onis, M., Hallal, P.C., Blossner, M., and Shrimpton, R. (2010). Worldwide timing of growth faltering: Revisiting implications for interventions. *Pediatrics* 125: e473-e480.

WHO Multicentre Growth Reference Study Group. (2006). *WHO Child Growth Standards: Length/Height-for-age, Weight-for-Age, Weight-for-Length, Weight-for-Height and Body Mass Index-for-Age: Methods and Development*. Geneva: World Health Organization.

WHO Multicentre Growth Reference Study Group. (2007). *WHO Child Growth Standards: Head Circumference-for-Age, Arm Circumference-for-Age, Triceps Skinfold-for-Age and Subscapular Skinfold-for-Age: Methods and Development*. Geneva: World Health Organization.

WHO. (2002). *Prevention and Control of Schistosomiasis and Soil-Transmitted Helminthiasis. Report of a WHO Expert Committee*. Geneva: World Health Organization.

WHO. (2011). *Hemoglobin Concentrations for the Diagnosis of Anemia and Assessment of Severity*. *Vitamin and Mineral Nutrition Information System*. Geneva: World Health Organization. http://www.who.int/vmnis/indicators/haemoglobin.pdf, accessed December 2012.

World Health Organization and UNICEF. (2012). *Progress on Drinking Water and Sanitation: 2012 Update*. New York: UNICEF.

Table 1. Distribution of Sample Households by District and Survey Period

District Name	Endline			Baseline			Panel		
	Treatment	Control	Total	Treatment	Control	Total	Treatment	Control	Total
Banyuwangi	155	155	310	130	130	260	115	117	232
Bondowoso	156	155	311	131	130	261	122	125	247
Jombang	160	160	320	130	130	260	121	118	239
Blitar	155	155	310	130	130	260	121	121	242
Madiun	156	155	311	129	131	260	111	114	225
Probolinggo	155	155	310	130	130	260	120	126	246
Ngawi	156	157	313	131	130	261	114	115	229
Situbondo	157	158	315	130	135	265	124	124	248
ALL	**1,250**	**1,250**	**2,500**	**1,041**	**1,046**	**2,087**	**948**	**960**	**1,908**

Table 2.Test for Randomization of the TSSM Sample

	Total			Treatment			Control			Difference	
Variables	**N**	**Mean**	**Std Error**	**N**	**Mean**	**Std Error**	**N**	**Mean**	**Std Error**	**T-C**	**p-value**
Water and Sanitation											
Improved sanitation[25]	2,087	0.425	0.011	1,042	0.420	0.015	1,045	0.430	0.015	0.010	0.630
HH practices open defecation	2,087	0.395	0.011	1,042	0.389	0.015	1,045	0.402	0.015	–0.013	0.270
Improved water source	2,086	0.873	0.007	1,042	0.872	0.010	1,044	0.874	0.010	–0.002	0.444
% HH with water available at the handwashing station	2,087	0.620	0.011	1,042	0.631	0.015	1,045	0.609	0.015	0.022	0.150
% HH with water and soap available at the handwashing station	2,087	0.470	0.011	1,042	0.478	0.015	1,045	0.462	0.015	0.016	0.225
Selected Child Health Outcomes											
Diarrhea											
Incidence in the past:											
48 hours	2,344	0.043	0.004	1,182	0.038	0.006	1,162	0.048	0.006	–0.010	0.119
7 days	2,344	0.073	0.005	1,182	0.070	0.007	1,162	0.075	0.008	–0.005	0.319
14 days	2,344	0.084	0.006	1,182	0.080	0.008	1,162	0.088	0.008	–0.008	0.240
ALRI											
Incidence in the past:											
48 hours	2,344	0.020	0.003	1,182	0.019	0.004	1,162	0.022	0.004	–0.003	0.298
7 days	2,344	0.026	0.003	1,182	0.025	0.005	1,162	0.027	0.005	–0.002	0.389
14 days	2,344	0.029	0.003	1,182	0.028	0.005	1,162	0.029	0.005	–0.001	0.444
Anemia and anthropometric											
Anemia, Hb<110 (g/L)	1,592	0.709	0.011	788	0.714	0.016	804	0.703	0.016	0.011	0.313
Weight (in kg)	2,095	8.22	1.85	1,047	8.22	0.057	1,048	8.236	0.057	–0.018	0.815
Height (in cm)	2,334	70.17	7.85	1,176	69.99	0.226	1,158	70.45	0.233	–0.454	0.162
Body mass index (weight/height^2)	2,072	16.053	0.05	1,039	16.086	0.064	1,033	15.983	0.75	–0.104	0.295
Anthropometric z-score:											

[25] Note that the definition of improved sanitation used in this report is different from that used in the baseline report. Here we define improved sanitation to exclude those who used a sanitation facility that was “improved” in terms of the quality of the infrastructure but that was a shared or public facility. This is in line with the WHO definition of improved sanitation.

Weight-for-height	2,082	–0.776	0.025	1,041	–0.759	0.036	1,041	–0.793	0.036	0.034	0.252
Height-for-age	2,090	–0.878	0.030	1,045	–0.888	0.043	1,045	–0.869	0.042	–0.019	0.376
Body mass index-for-age	2,072	–0.349	0.029	1,039	–0.327	0.040	1,033	–0.371	0.041	0.044	0.221
Weight-for-length	2,077	–0.429	0.028	1,042	–0.404	0.039	1,035	–0.455	0.041	0.051	0.184
Head circumference-for-age	2,079	–0.400	0.025	1,037	–0.427	0.035	1,042	–0.372	0.036	–0.055	0.137
Arm circumference-for-age	1,922	0.029	0.027	965	0.059	0.039	957	–0.001	0.037	0.060	0.132
Child development z-score											
Mobility	1,762	0.000	0.024	884	0.010	0.033	878	-0.010	0.034	0.020	0.336
Communication	1,761	0.000	0.024	883	0.013	0.003	878	-0.013	0.035	0.026	0.294
Social	1,762	0.000	0.024	884	0.022	0.034	878	-0.022	0.034	0.044	0.180
Household characteristics											
Average age, HH head	2,087	39.278	0.254	1,042	39.341	0.361	1,045	39.215	0.357	0.126	0.402
% Male, HH head	2,087	0.956	0.004	1,042	0.955	0.006	1,045	0.957	0.006	-0.002	0.407
Household head											
Less than elementary	2,087	0.010	0.002	1,042	0.010	0.003	1,045	0.010	0.003	0.000	0.500
Elementary school, MI	2,087	0.526	0.011	1,042	0.535	0.015	1,045	0.518	0.015	0.017	0.211
General/vocational junior high	2,087	0.198	0.009	1,042	0.190	0.012	1,045	0.207	0.013	-0.017	0.168
General/vocational senior high	2,087	0.175	0.008	1,042	0.170	0.012	1,045	0.181	0.012	-0.011	0.258
University (S1,S2,S3)	2,087	0.040	0.004	1,042	0.039	0.006	1,045	0.041	0.006	-0.002	0.407

Table 3. Comparison of Control and Treatment Groups That Were Not in the Endline Survey

	Treatment	Control	*p-value*
Household Characteristics			
Household size	4.462	4.756	*0.129*
Age of household head	39.763	40.884	*0.565*
HH head with less than elementary education	0.000	0.012	*0.230*
HH head completed elementary	0.527	0.535	*0.915*
HH head completed junior high school	0.172	0.233	*0.316*
HH head completed senior high school	0.194	0.093	*0.057**
HH head with university degree	0.011	0.035	*0.278*
Female-headed household	0.108	0.105	*0.951*
Number of children under 5 in HH	1.097	1.209	*0.036***
Number of children between 6 and 10 years of age	0.290	0.244	*0.541*
Number of adults in the household	2.849	3.116	*0.101*
Mean per capita income (Rp per year)	3,771,601	3,201,643	*0.554*
HH received BLT fund in 2007	0.258	0.244	*0.832*
Dwelling Characteristics			
Own current dwelling	0.677	0.791	*0.088**
Using LPG as cooking fuel	0.078	0.093	*0.671*
Using wood as cooking fuel	0.742	0.674	*0.323*
Dwelling Type (% HHs)			
Detached single-story house	0.914	0.884	*0.504*
Detached multistory house	0.022	0.058	*0.209*
Connected, single-story house	0.054	0.058	*0.899*
Connected, multistory house	0.011	0.000	*0.338*
Wall Materials			
Brick	0.108	0.058	*0.236*
Concrete	0.591	0.733	*0.005**
Wood/logs	0.161	0.140	*0.687*
Bamboo	0.118	0.070	*0.271*
Household Sanitation/Hygiene Practice			
Household with improved sanitation	0.419	0.477	*0.443*
Household correctly disposes of child's feces	0.516	0.547	*0.686*
Household with improved drinking water source	0.828	0.860	*0.552*
Household practices open defecation	0.366	0.419	*0.471*
Household treats drinking water prior to consumption	0.785	0.733	*0.415*
Soap available at the handwashing station	0.548	0.512	*0.625*
Water available at the handwashing station	0.896	0.931	*0.489*
Soap and water available at the washing station	0.527	0.488	*0.609*
Sample Size	**86**	**93**	

Note: ***, **, * denote statistically significant at 1, 5, and 10 percent, respectively.

Table 4. Comparisons of Sample Present at Baseline and Endline

	Present at Baseline (Baseline values of variables)			Present at Endline (Endline values of variables)		
	Tracked at Endline (1)	Lost to follow-up (2)	*p-value*	Additional Sample (3)	Panel Sample (4)	*p-value*
	(N=1,908)	**(N=179)**	[(1)-(2)]	**(N=592)**	**(N=1,908)**	[(3) - (4)]
Household Characteristics						
Household size	4.573	4.603	*0.760*	4.878	4.868	*0.874*
Age of household head	39.194	40.302	*0.222*	41.395	40.479	*0.083**
HH head with less than elementary education	0.007	0.006	*0.848*	0.002	0.004	*0.457*
HH head completed elementary	0.526	0.531	*0.908*	0.514	0.500	*0.566*
HH head completed junior high school	0.198	0.201	*0.923*	0.199	0.205	*0.747*
HH head completed senior high school	0.178	0.145	*0.275*	0.193	0.192	*0.991*
HH head with university degree	0.032	0.022	*0.479*	0.035	0.044	*0.393*
Female headed household (dummy=1; 0 otherwise)	0.039	0.106	*0.000****	0.022	0.035	*0.125*
Number of children under 5 in HH	1.189	1.151	*0.241*	1.139	1.133	*0.730*
Number of children between 6 and 10 years of age	0.348	0.268	*0.050**	0.318	0.323	*0.810*
Number of adults in the household	2.742	2.978	*0.004****	3.091	3.016	*0.179*
Mean per capita income (Rp per year)	2,918,884	3,497,766	*0.098**	3,694,026	3,843,229	0.512
HH received BLT fund (dummy=1; 0 otherwise)	0.254	0.251	*0.947*	0.204	0.240	*0.073**
Dwelling Characteristics						
Owned current dwelling	0.834	0.732	*0.001****	0.860	0.860	*0.987*
Using LPG as cooking fuel	0.048	0.084	*0.035***	0.453	0.426	*0.245*
Using wood as cooking fuel	0.742	0.709	*0.342*	0.527	0.550	*0.320*
Dwelling Type (% HHs)						
Detached single-story house	0.873	0.899	*0.305*	0.890	0.884	*0.679*
Detached multistory house	0.044	0.039	*0.754*	0.025	0.024	*0.870*
Connected, single-story house	0.081	0.056	*0.227*	0.081	0.088	*0.591*
Connected, multistory house	0.002	0.006	*0.241*	0.003	0.004	*0.916*
Wall Materials						
Brick	0.074	0.084	*0.617*	0.037	0.051	*0.170*
Concrete	0.611	0.659	*0.208*	0.709	0.673	*0.096**

Wood/logs	0.172	0.151	*0.477*	0.152	0.156	*0.820*
Bamboo	0.111	0.095	*0.516*	0.064	0.089	*0.054**
Household Sanitation/Hygiene Practice						
Household with improved sanitation	0.447	0.423	*0.535*	0.470	0.431	*0.097**
Household correctly disposes of child's feces	0.530	0.531	*0.983*	0.525	0.569	*0.061**
Household with improved drinking water source	0.876	0.844	*0.217*	0.877	0.890	*0.355*
Household reported wash hands after going to the toilet	0.990	0.983	*0.395*	0.986	0.985	*0.833*
Household practices open defecation	0.396	0.391	*0.903*	0.331	0.355	*0.280*
Household treat drinking water prior to consumption	0.782	0.760	*0.493*	0.644	0.645	*0.962*
Soap available at the handwashing station	0.484	0.531	*0.235*	0.637	0.588	*0.033***
Water available at the handwashing station	0.897	0.912	*0.604*	0.966	0.964	*0.854*
Soap and water available at the handwashing station	0.466	0.508	*0.283*	0.622	0.575	*0.046***

*Note:****, **, * denote statistically significant at 1, 5, and 10 percent, respectively.

Table 5. Triggering Status of Villages (in percent)

	Treatment Villages			Control Villages		
	WSP (Nov. 2009)	IE Longitudinal (May 2010)	IE Endline (Jan. 2011)	WSP (Nov. 2009)	IE Longitudinal (May 2010)	IE Endline (Jan. 2011)
Jombang	100	100	90	0	0	0
Blitar	70	80	100	0	10	0
Bondowoso	100	100	100	10	20	20
Madiun	80	40	60	0	10	20
Banyuwangi	90	30	50	0	30	10
Probolinggo	80	40	50	0	10	20
Situbondo	80	70	40	20	30	20
Ngawi	60	40	40	0	10	20
AVERAGE	**83%**	**62.5%**	**66.3%**	**4%**	**15%**	**13.8%**

Table 6. Household Knowledge and Attendance at Triggering Session

	Percent of households with knowledge of triggering		**Percent who attended**	
	Treatment (N=286)	**Control (N=100)**	**Treatment (N=171)**	**Control (N=58)**
Jombang	38.2	8.4	18.1	5.0
Blitar	23.2	8.4	11.6	1.9
Bondowoso	41.7	14.4	30.1	9.0
Madiun	20.1	5.7	7.7	2.6
Banyuwangi	18.5	4.0	9.7	1.3
Probolinggo	20.8	3.3	12.3	0.6
Situbondo	20.8	18.8	14.7	15.8
Ngawi	14.1	2.6	5.1	0.6
ALL	**24.6**	**8.2**	**13.7**	**4.6**

Table 7. Program Exposure at the Community Level

A. Respondent: Village Head/ Staff Responsible for Sanitation

	Treatment	Control	All	*p-value*
Heard about SToPS[26]	70.00	35.00	52.5	*0.000****
Attended SToPS kecamatan meeting	45.00	11.25	28.13	*0.000****
Request triggering to camat and/or PUSKESMAS	46.25	25.00	35.63	*0.005****
Sample size	**80**	**80**	**160**	
Of those who said they helped households gain quicker access to sanitation, through … *	72.41	58.82	69.33	*0.292*
Offering credit	4.76	10.00	5.77	*0.533*
Providing lower-cost material	9.52	10.00	9.62	*0.964*
Fundraising for households that don't have toilet	16.67	10.00	15.38	*0.608*
Building a public toilet	45.24	80.00	51.92	*0.049*
Other	52.38	40.00	50	*0.491*
Sample size	**42**	**10**	**52**	
Satisfied with SToPS performance (excellent, good, and satisfactory)	72.41	58.82	69.33	*0.292*
Sample size	**58**	**17**	**75**	

Note: This question not necessarily restricted to SToPS related triggering

B. Respondent: Sub-village Head

	Treatment	Control	All	*p-value*
Heard about SToPS	52.5	25		*0.000****
Remembered the message	35.00	17.5		*0.017***
There's triggering in the village	50	15		*0.000****
- SToPS triggering	37.5	10		*0.000****
Attended the triggering session [N[Treatment]=30; N[Control]=6]	37.5	7.5		*0.000****
- Facilitator is very persuasive	13.33			
- Facilitator is somewhat persuasive	76.67			
Thinks that sanitation has improved post-triggering	75.00			
Satisfied with the success of SToPS (very satisfied, satisfied, and somewhat satisfied)	62.50			
Village has been declared open defecation free (ODF)	11.25	1.25		*0.009***

*Note:****, **, * denote statistically significant at 1, 5, and 10 percent, respectively.

[26] StoPs is the Indonesian acronym for Total Sanitation and Sanitation Marketing (TSSM).

Table 8A. Descriptive Statistics of Pre- and Post-Intervention Sanitation and Demographic Variables

	Endline (post-intervention)			Baseline (pre-intervention)		
	Treatment	Control	*p-value*	Treatment	Control	*p-value*
Sanitation and Hygiene Behavior						
Where household members normally defecate:						
a. Improved sanitation facility	0.426	0.435	*0.702*	0.420	0.425	0.841
b. In fields, river, or on beach.	0.348	0.363	*0.486*	0.391	0.400	0.692
Toilet built in the past 2 years	0.159	0.130	*0.072**	N/A	N/A	N/A
Improved drinking water	0.886	0.895	*0.489*	0.878	0.875	0.854
Wash hands after going to the toilet	0.982	0.988	*0.252*	0.988	0.992	0.484
Soap is available at place of washing hands	0.589	0.586	*0.876*	0.491	0.479	0.615
Water is available at place of washing hands	0.961	0.966	*0.583*	0.907	0.888	0.243
Water and soap is available	0.575	0.576	*0.937*	0.474	0.461	0.584
Food is fully covered	0.829	0.845	*0.366*	0.733	0.738	0.809
Household Characteristics						
Age of HH head	40.485	40.472	*0.979*	39.313	39.001	0.553
HH head with less than primary education	0.004	0.003	*0.701*	0.007	0.006	0.775
HH head completed primary school	0.516	0.484	*0.170*	0.534	0.516	0.431
HH head completed junior high school	0.201	0.210	*0.603*	0.192	0.205	0.482
HH head completed senior high school	0.180	0.205	*0.160*	0.168	0.187	0.271
HH head with university degree	0.044	0.043	*0.895*	0.031	0.034	0.709
Female-headed household	0.039	0.030	*0.308*	0.039	0.039	0.985
Annual per capita household income (‘000 Rp)	3,914	3,775	*0.543*	2,784	3,053	0.170
Household size	4.916	4.820	*0.142*	4.608	4.603	0.245
Max N	**952**	**956**		**952**	**956**	

*Note: ***, **, * denote statistically significant at 1, 5, and 10 percent, respectively.*

Table 8b. Descriptive Statistics of Pre- and Post-Treatment Child Health Outcomes

	Endline (post-intervention)			Baseline (pre-intervention)		
	Treatment	Control	*p-value*	Treatment	Control	*p-value*
Diarrhea, 7-day prevalence	0.024	0.038	*0.070**	0.076	0.083	*0.569*
Acute respiratory infection, 7-day prevalence	0.181	0.167	*0.399*	0.144	0.116	*0.063**
Acute lower respiratory infection, 7-day prevalence	0.021	0.019	*0.729*	0.018	0.011	*0.250*
Diarrhea, 2-day prevalence	0.016	0.031	*0.025***	0.044	0.054	*0.298*
Acute respiratory infection, 2-day prevalence	0.152	0.137	*0.273*	0.213	0.217	*0.809*
Acute lower respiratory infection, 2-day prevalence	0.018	0.017	*0.847*	0.021	0.019	*0.736*
Child's fecal sample contains any worm[27]	0.040	0.039	*0.889*	N/A	N/A	N/A
Prevalence of children infected with *Al Ascaris*	0.034	0.033	*0.881*	N/A	N/A	N/A
Prevalence of low intensity	0.027	0.021	*0.423*	N/A	N/A	N/A
Prevalence of moderate intensity	0.006	0.012	*0.197*	N/A	N/A	N/A
Average infection intensity (eggs per gram in stool)[28]	103.2	159.6	*0.224*	N/A	N/A	N/A
Prevalence of children infected with *TT Trichuris*	0.000	0.001	*0.319*	N/A	N/A	N/A
Prevalence of low intensity	0.000	0.001	*0.319*	N/A	N/A	N/A
Average infection intensity (eggs per gram in stool)	0.000	0.711	*0.319*	N/A	N/A	N/A
Prevalence of children infected with *CT hook*	0.006	0.005	*0.733*	N/A	N/A	N/A
Prevalence of low intensity	0.006	0.005	*0.733*	N/A	N/A	N/A
Average infection intensity (eggs per gram in stool)	0.437	0.434	*0.995*	N/A	N/A	N/A
Symptoms reported:						
Fever, 7-day prevalence	0.155	0.154	*0.953*	0.279	0.331	*0.013***
Cough, 7-day prevalence	0.163	0.151	*0.493*	0.260	0.263	*0.868*
Congestion/runny nose, 7-day prevalence	0.284	0.244	*0.046***	0.469	0.420	*0.033***

[27]Dummy=1 if any of these worms detected in the fecal sample: *Al ascaris, TT trichuris, and Ct hook.*

[28]Figures reported here are based on the mean compared across the whole sample. If we restrict the sample to children with positive helminthes *Al Ascaris, TT Trichuris*, and *CT Hook*, the averages across observations with positive *helminthes* are 1,714 epg, 600 epg, and 81.5 for *Al Ascaris, TT Trichuris*, and *CT Hook*, respectively.

Breathing difficulty, 7-day prevalence	0.050	0.044	*0.498*	0.065	0.065	*0.986*
Stomach pain/cramps, 7-day prevalence	0.046	0.036	*0.288*	0.035	0.036	*0.913*
Nausea, 7-day prevalence	0.035	0.024	*0.133*	0.028	0.033	*0.517*
Vomiting, 7-day prevalence	0.041	0.040	*0.978*	0.074	0.076	*0.878*
3+ bowel movements in the last 24 hours, 7-day prevalence	0.038	0.042	*0.580*	0.092	0.089	*0.857*
Watery or soft stool, 7-day prevalence	0.059	0.061	*0.875*	0.111	0.112	*0.904*
Mucus or blood in stool, 7-day prevalence	0.008	0.020	*0.034***	0.019	0.018	*0.857*
Refusal to eat, 7-day prevalence	0.052	0.075	*0.043***	0.143	0.145	*0.919*
Abrasion, scrapes, or bruising, 7-day prevalence	0.117	0.133	*0.293*	0.062	0.063	*0.864*
Skin itching, 7-day prevalence	0.090	0.088	*0.902*	0.075	0.071	*0.713*
Anthropometric and Anemia:						
Hemogloblin level (g/L)	110.877	110.417	*0.623*	100.838	101.692	*0.271*
Anemic (Hb < 110 g/L)	0.447	0.437	*0.884*	0.722	0.703	*0.475*
Weight (in kg)	12.673	12.684	*0.910*	8.23	8.26	*0.745*
Height (in cm)	91.565	91.708	*0.609*	71.13	71.45	*0.355*
Body mass index (weight/height^2)	15.073	15.194	*0.499*	16.086	15.983	*0.298*
Weight-for-age (*z*-score)	–1.364	–1.389	*0.618*	–0.756	–0.818	*0.367*
Height-for-age (*z*-score)	–1.716	–1.724	*0.258*	–0.932	–0.988	*0.333*
Body mass index (*z*-score)	–0.370	–0.324	*0.731*	–0.334	–0.393	*0.225*
Weight-for-height (*z*-score)	–0.566	–0.626	*0.337*	–0.417	–0.481	*0.560*
Head circumference (*z*-score)	–1.194	–1.219	*0.393*	–0.550	–0.443	*0.357*
Mid-upper-arm circumference (*z*-score)	–0.206	–0.090	*0.667*	0.108	0.063	*0.429*
Development Indicators:						
Communication skills-for-age (*z*-score)	0.043	–0.039	*0.084**	–0.016	–0.024	*0.861*
Mobility skills-for-age (*z*-score)	–0.002	–0.002	*0.999*	0.000	–0.014	*0.787*
Social-personal skills-for-age (*z*-score)	0.028	–0.021	*0.370*	–0.002	–0.022	*0.694*
Child's gender (1=male, 0=female)	0.497	0.513	*0.495*	0.499	0.513	*0.539*

Child's age (in month)	40.492	40.787	*0.324*	11.974	12.327	*0.230*
Max N	**959**	**956**		**959**	**956**	

Note: ***, **, * denote statistically significant at 1, 5, and 10 percent, respectively. Fecal sample only available in endline.

Table 9. Toilet Construction Since Baseline

	All		Treatment		Control		Diff	*p*-value
Households who built a toilet in past 2 years	275	14.4%	151	15.9%	124	13.0%	3.0%	*0.072**
Who previously had:								
a. no sanitation	61	8.16%	37	10.0%	24	6.3%	3.7%	*0.065**
b. public or shared sanitation	38	20.9%	26	23.4%	12	16.9%	6.5%	*0.29*
c. private unimproved sanitation	69	42.6%	32	47.8%	37	39.0%	7.9%	*0.26*
d. improved sanitation	102	12.8%	53	13.4%	49	12.2%	1.2%	*0.614*
Whose sanitation changed to "improved" as a result of the construction	180	9.4%	92	9.7%	88	9.2%	0.5%	*0.73*
N		**1,908**		**956**		**952**		

Note: ***, **, * denote statistically significant at 1, 5, and 10 percent, respectively.

Table 10. Household Reports of Open Defecation

	Any householder		Men		Women	Children		N
All	0.511		0.428		0.394	0.454		2,500
Control	0.532		0.444		0.405	0.475		1,250
Treatment	0.488		0.411		0.384	0.432		1,250
Diff	–0.044		–0.033		–0.021	–0.043		
p-value	*0.025*	**	*0.098*	*	*0.2883*	*0.030*	**	
On River								
All	0.537		0.451		0.418	0.478		2,339
Control	0.566		0.472		0.434	0.509		1,152
Treatment	0.508		0.431		0.402	0.449		1,187
Diff	–0.058		–0.041		–0.032	–0.06		
p-value	*0.005*	***	*0.047*	**	*0.125*	*0.004*	***	
Not on River								
All	0.130		0.081		0.050	0093		161
Control	0.143		0.112		0.061	0.082		98
Treatment	0.111		0.032		0.032	0.111		63
Diff	–0.032		–0.08		–0.029	0.029		
p-value	*0.563*		*0.068*	*	*0.404*	*0.532*		

Note: ***, **, * denote statistically significant at 1, 5, and 10 percent, respectively.

Table 11. Attitudes to Open Defecation

	All	Treatment	Control	*p-value*
% strongly agreeing or agreeing with:				
Having a toilet/latrine at home will protect my family from being the target of gossip.	78.4	78.7	78.1	*0.70*
Toilet facilities in the village will benefit the community as environmental pollution is lessened.	96.1	96.3	95.8	*0.54*
Most of the people I know normally defecate in a toilet/latrine.	73.3	74.2	72.5	*0.34*
Having a toilet facility will protect me and my family members from illnesses such as diarrhea.	92.0	91.8	92.2	*0.71*
Those who practice open defecation will not be accepted by the community in which they live.	34.0	32.6	35.5	*0.12*
If our ancestors could practice open defecation, then we can do so now.	6.9	7.0	6.7	*0.75*
Defecating in the river is a common occurrence and others do it too.	35.0	34.6	35.4	*0.65*
It is alright for children to practice open defecation.	22.0	21.4	22.5	*0.53*
It is acceptable to practice open defecation if you don't have a toilet/latrine.	31.2	31.9	30.7	*0.52*
N	**2,500**	**1,250**	**1,250**	

Table 12. Caregivers' Perception about Causes of Diarrhea

	Endline		Baseline		Diff-in-Diff	
	Treatment	Control	Treatment	Control		
% caregivers reported diarrhea caused by:						
Eating stale food	93.4	92.7	96.2	96.8	*1.8*	
Eating street food	81.1	81.9	81.4	82.4	*0.3*	
Eating food that had been touched by flies	95.4	94.3	95.9	96.4	*1.5*	
Eating unclean food	97.2	96.0	97.4	97.8	*1.8*	*
Drinking unclean water	95.8	94.4	95.8	97.0	*2.6*	**
Using dirty latrine/toilet	83.2	83.5	86.5	85.9	*–1.0*	
Not washing hand with water	88.1	89.1	87.7	87.0	*–1.7*	
Not washing hands with soap and water	82.8	83.3	81.7	82.8	*0.5*	
Change of climate	56.8	55.5	67.9	68.1	*1.5*	
Exposure to sun	31.9	28.8	28.6	31.2	*5.6*	**
Teething	16.7	16.1	18.9	19.7	*1.5*	
Getting several types of vaccines	10.4	9.3	11.2	13.4	*3.3*	*
Dirty house	79.3	75.7	81.5	80.5	*2.6*	
Unclean surrounding	82.2	79.9	85.5	86.8	*3.5*	*
Others defecating in the river	58.3	56.9	60.4	60.0	*1.0*	
Others practicing open defecation	65.0	66.3	72.8	71.6	*–2.4*	
Diarrhea can be prevented	98.1	97.6	99.2	98.5	*–0.3*	
N	**972**	**981**	**972**	**981**		

Note: ***, **, * denote statistically significant at 1, 5, and 10 percent, respectively.

Table 13. Households Knowledge about Building a Latrine

	All (N=1,908)	Treatment (N=952)	Control (N=956)	
Cheapest cost to build latrine (in Rupiah)	1,196,655	1,269,935	1,124,547	
Costs of building latrine in comparison to two years ago (% HHs)				
Far more expensive	25.00	24.58	25.42	
More expensive	39.31	39.5	39.12	
Much the same	15.04	14.71	15.38	
Cheaper	15.78	15.86	15.69	
Much cheaper	2.78	2.52	3.03	
Know who to contact to get materials and tradesmen in order to build toilet (% HHs)				
Yes	93.97	94.43	93.51	
Maybe	0.58	0.74	0.42	
No	3.09	1.89	4.29	***
Easy to get materials and tradesmen for building a latrine (% HHs)				
Yes	95.28	95.06	95.5	
Maybe	1.73	1.47	1.99	
No	2.36	2.63	2.09	
Costs for building materials and tradesmen for latrine building is affordable (% HHs)				
Yes	68.92	67.65	70.19	
Maybe	9.07	9.35	8.79	
No	21.02	21.95	20.08	

Note: ***, **, * denote statistically significant at 1, 5, and 10 percent, respectively.

Table 14. Estimated Effects of TSSM Program on Sanitation and Hygiene Behavior

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Control Means	All	All	All	All	All	No Sanitation at Baseline	Sanitation at Baseline
Toilet built in the past two years (since baseline)	0.128	0.037***	0.029**	0.029**	0.029**	0.032**	0.037**	0.010
	(0.334)	(0.011)	(0.013)	(0.013)	(0.013)	(0.013)	(0.016)	(0.021)
JMP defined improved sanitation	0.443	–0.011	–0.014	–0.008	–0.006	0.0000	0.046**	–0.044
	(0.497)	(0.026)	(0.027)	(0.018)	(0.017)	(0.016)	(0.022)	(0.026)
Households normally defecate in the open	0.362	–0.022	–0.015	–0.009	–0.012	–0.017	–0.058**	0.015
	(0.481)	(0.025)	(0.026)	(0.013)	(0.012)	(0.012)	(0.025)	(0.011)
One or more householders defecate in the open at least occasionally	0.533	–0.040	–0.029	–0.029	–0.034	–0.048**	–0.044**	–0.054**
	(0.499)	(0.027)	(0.029)	(0.029)	(0.026)	(0.024)	(0.022)	(0.027)
JMP defined improved drinking water source	0.895	–0.019	–0.012	–0.012	–0.012	–0.008	0.009	–0.041**
	(0.306)	(0.012)	(0.013)	(0.013)	(0.013)	(0.013)	(0.022)	(0.017)
Household correctly disposes of child's feces	0.539	0.039	0.033	0.014	0.019	0.034*	0.081**	0.009
	(0.499)	(0.048)	(0.052)	(0.036)	(0.035)	(0.019)	(0.028)	(0.017)
Number of observations	1,250	2,500	1,908	1,908	1,908	1,908	938	969
Panel sample	N	N	Y	Y	Y	Y	Y	Y
Control for outcome in baseline	n.a.	N	N	Y	Y	Y	Y	Y
Controls	n.a.	N	N	N	A	B	B	B
Pairwise sub-district fixed effects	n.a.	Y	Y	Y	Y	Y	Y	Y

Note: Robust standard-errors in parentheses. ***, **, * denote statistically significant at 1, 5, and 10 percent, respectively. Controls A=, HH's gender, HH's education; Controls B= Controls A+ household size, use wood to cook, recipient of government cash transfer, access to piped water, dwelling with a dirt floor; wall made of either brick, concrete, adobe, or logs; and whether the community is within 10 minutes of a river. Data for one or more householders defecate in the openat least occasionally only available in endline, the coefficients reported for this variable excluded outcome in baseline.

Table 15. Estimated Effects of the TSSM Program on Handwashing Behavior

	(1) Control Means	(2) All	(3) All	(4) All	(5) All	(6) All	(7) No Sanitation at BL	(8) Sanitation at BL
Household reported washing hands after defecation	0.988	–0.003	–0.006	–0.006	–0.005	–0.005	–0.001	–0.003
	(0.109)	(0.003)	(0.004)	(0.005)	(0.004)	(0.005)	(0.007)	(0.003)
Water and soap available at the handwashing station	0.590	–0.008	–0.006	–0.005	–0.001	0.008	0.037	–0.033
	(0.492)	(0.019)	(0.022)	(0.020)	(0.019)	(0.018)	(0.025)	(0.023)
Water available at the handwashing station	0.970	–0.013**	–0.007	–0.009	–0.009	–0.009	0.004	–0.019**
	(0.170)	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)	(0.014)	(0.009)
Soap available at the handwashing station	0.598	–0.000	–0.000	0.000	0.004	0.014	0.034	–0.023
	(0.490)	(0.019)	(0.022)	(0.020)	(0.019)	(0.018)	(0.025)	(0.022)
Number of observations	1,250	2,500	1,908	1,908	1,908	1,908	938	969
Panel sample	n.a.	N	Y	Y	Y	Y	Y	Y
Control for outcome in baseline	n.a.	N	N	Y	Y	Y	Y	Y
Controls	n.a.	N	N	N	A	B	B	B
Pairwise sub-district fixed effects	n.a.	Y	Y	Y	Y	Y	Y	Y

Note: Robust standard errors in parentheses. ***, **, * denote statistically significant at 1, 5, and 10 percent, respectively. Controls A= HH's gender, HH's education; Controls B= Controls A+ household size, use wood to cook, recipient of BLT cash transfer, piped water, dwelling with dirt floor, and walls made of either brick, concrete, adobe, or logs; community on a river.

Table 16. Estimated Effects of the TSSM Program on Sanitation and Hygiene Behavior for Poor and Non-poor Households

	(1) Control Means	(2) Poor	(3) Non-poor	(4) No Sanitation at Baseline	
				Poor (4a)	Non-poor (4b)
Toilet built in the past two years (since baseline)	0.128	–0.031	0.052***	0.022	0.042*
	(0.334)	(0.025)	(0.015)	(0.023)	(0.022)
JMP defined improved sanitation	0.443	0.042*	0.018	0.018	0.055*
	(0.497)	(0.022)	(0.029)	(0.029)	(0.033)
Households normally defecate in the open	0.362	–0.024	–0.031	–0.051	–0.098**
	(0.481)	(0.033)	(0.022)	(0.051)	(0.039)
One or more householders defecate in the open at least occasionally	0.533	–0.073*	–0.028	–0.044	–0.045
	(0.499)	(0.039)	(0.020)	(0.042)	(0.032)
JMP defined improved drinking water source	0.895	0.014	–0.021	0.004	–0.012
	(0.306)	(0.026)	(0.015)	(0.037)	(0.030)
Household correctly disposes of child's feces	0.539	0.051	0.055	0.055	0.070*
	(0.499)	(0.036)	(0.043)	(0.043)	(0.038)
Number of observations	1,250	454	1,436	329	601
Panel sample	N	Y	Y	Y	Y
Control for outcome in baseline	n.a.	Y	Y	Y	Y
Controls	n.a.	B	B	B	B
Pairwise sub-district fixed effects	n.a.	Y	Y	Y	Y

Note: Robust standard-errors in parentheses. ***, **, * denote statistically significant at 1, 5, and 10 percent, respectively. Controls A= head's gender, head's education; Controls B= Controls A+ household size, use wood to cook, recipient of BLT cash transfer, piped water, dwelling with dirt floor and wall materials made of either brick, concrete, adobe or logs; community on a river.

Table 17. Estimated Effects of the TSSM Program on Child Health Outcomes

	(1) Control Means	(2) All	(3) All	(4) All	(5) All	(6) All	(7) Sanitation at BL	(8) No Sanitation at BL		
								All (8a)	Poor (8b)	Non-poor (8c)
Diarrhea, 7-day prevalence	0.046	–0.009*	–0.013***	–0.014***	–0.013***	–0.013**	–0.005	–0.031***	–0.055*	–0.020
	(0.211)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.009)	(0.028)	(0.012)
Diarrhea, 2-day prevalence	0.034	–0.006	–0.014***	–0.014***	–0.014***	–0.014***	–0.003	–0.030***	–0.049*	–0.018*
	(0.180)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.010)	(0.028)	(0.011)
Al ascaris infection (# eggs per gram)	128.25	–39.405	–60.232	n.a.	n.a.	n.a.	–71.920	–54.668	38.241	–146.082**
	(29.88)	(27.548)	(37.299)				(75.649)	(35.477)	(32.874)	(70.934)
Trichuris infection (# eggs per gram)[29]	0.520	–0.512	–0.703	n.a.	n.a.	n.a.	–0.991	–	–	–
	(0.520)	(0.367)	(0.507)				(0.767)			
Ct hook infection (# eggs per gram)[30]	0.318	0.056	0.089	n.a.	n.a.	n.a.	0.270	–0.109	–	0.047
	(0.180)	(0.176)	(0.250)				(0.296)	(0.415)		(0.625)
Acute respiratory infection, 7-day prevalence	0.177	0.014	0.015	0.014	0.003	0.003	–0.009	0.002	0.042	0.006
	(0.382)	(0.020)	(0.021)	(0.021)	(0.013)	(0.013)	(0.022)	(0.022)	(0.060)	(0.029)
Acute respiratory infection, 2-day prevalence	(0.144)	0.013	0.013	0.013	0.006	0.006	–0.007	0.000	0.019	0.007
	(0.351)	(0.011)	(0.012)	(0.012)	(0.013)	(0.013)	(0.022)	(0.019)	(0.049)	(0.028)
Acute lower respiratory infection, 7-day prevalence	0.024	–0.001	0.002	0.003	–0.001	–0.001	–0.013	0.001	–0.034**	0.005
	(0.152)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.010)	(0.007)	(0.016)	(0.009)
Acute lower respiratory infection, 2-day prevalence	0.024	–0.002	0.001	0.001	–0.003	–0.003	–0.013	–0.001	–0.039**	0.008
	(0.145)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.010)	(0.008)	(0.015)	(0.009)
Symptoms reported (7 day-prevalence):										
Fever, 7-day prevalence	0.184	-0.002	0.004	0.006	0.005	0.005	–0.006	0.007	0.052	0.021
	(0.387)	(0.011)	(0.013)	(0.013)	(0.015)	(0.015)	(0.024)	(0.020)	(0.048)	(0.025)
Cough, 7-day prevalence	0.161	0.008	0.004	0.005	0.002	0.002	0.002	–0.007	0.032	–0.011
	(0.367)	(0.012)	(0.012)	(0.012)	(0.013)	(0.014)	(0.020)	(0.021)	(0.060)	(0.027)
Congestion/runny nose, 7-day prevalence	0.277	0.039***	0.034**	0.032**	0.031*	0.032*	0.009	0.054*	0.110	0.065**

[29] We do not have enough observations to estimate the model once we restrict the sample to households with access to sanitation at baseline.

[30] We do not have enough observations to estimate the model once we restrict the sample to poor households with no access to sanitation at baseline.

	(0.448)	(0.015)	(0.015)	(0.015)	(0.017)	(0.017)	(0.025)	(0.028)	(0.073)	(0.032)
Breathing difficulty, 7-day prevalence	0.050	0.004	0.007	0.007	0.009	0.009	–0.011	0.025*	0.008	0.043**
	(0.219)	(0.007)	(0.007)	(0.007)	(0.008)	(0.008)	(0.013)	(0.013)	(0.035)	(0.018)
Stomach pain/cramps, 7-day prevalence	0.036	0.008	0.009	0.009	0.007	0.007	0.007	0.006	0.027	0.001
	(0.186)	(0.005)	(0.006)	(0.006)	(0.007)	(0.007)	(0.008)	(0.012)	(0.031)	(0.017)
Nausea, 7-day prevalence	0.025	0.015***	0.011**	0.011**	0.009	0.009	0.011	0.001	–0.001	0.003
	(0.157)	(0.005)	(0.006)	(0.006)	(0.006)	(0.007)	(0.009)	(0.011)	(0.023)	(0.014)
Vomiting, 7-day prevalence	0.043	0.006	–0.001	–0.001	–0.004	–0.004	–0.008	–0.004	0.029	–0.014
	(0.204)	(0.006)	(0.007)	(0.007)	(0.008)	(0.009)	(0.013)	(0.012)	(0.033)	(0.012)
3+ bowel movements in the last 24 hours, 7-day prevalence	0.055	–0.003	–0.005	–0.006	0.001	0.002	–0.006	–0.013	–0.030	0.000
	(0.228)	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)	(0.010)	(0.011)	(0.028)	(0.015)
Watery or soft stool, 7-day prevalence	0.076	–0.004	–0.002	–0.002	0.005	0.006	–0.004	0.010	0.004	0.024
	(0.265)	(0.007)	(0.008)	(0.008)	(0.009)	(0.009)	(0.012)	(0.016)	(0.029)	(0.024)
Mucus or blood in stool, 7-day prevalence	0.018	–0.005	–0.011***	–0.011***	–0.006*	–0.006*	–0.005	–0.009	–0.011	–0.011
	(0.134)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.006)	(0.013)	(0.009)
Refusal to eat, 7-day prevalence	0.070	–0.009	–0.022**	–0.024**	–0.022**	–0.022**	–0.034**	–0.011	0.046	–0.016
	(0.255)	(0.008)	(0.010)	(0.010)	(0.011)	(0.011)	(0.017)	(0.016)	(0.042)	(0.017)
Abrasion, scrapes, or bruising, 7-day prevalence	0.116	–0.016*	–0.016	–0.016	–0.014	–0.011	–0.024	–0.006	–0.031	0.003
	(0.320)	(0.009)	(0.011)	(0.011)	(0.014)	(0.015)	(0.019)	(0.023)	(0.042)	(0.034)
Skin itching, 7-day prevalence	0.090	0.004	0.001	0.001	–0.006	–0.007	–0.021	–0.000	0.013	0.000
	(0.286)	(0.010)	(0.010)	(0.010)	(0.012)	(0.012)	(0.015)	(0.020)	(0.045)	(0.026)
Anemia and Anthropometrics:										
Anemic (Hb < 110 g/L)	0.499	–0.004	0.011	0.021	0.019	0.020	0.052	–0.003	–0.038	0.001
	(0.500)	(0.016)	(0.018)	(0.022)	(0.022)	(0.022)	(0.035)	(0.038)	(0.069)	(0.052)
Weight (in kg)	11.564	–0.063	0.030	0.075	0.070	0.068	0.021	0.241***	0.280	0.277**
	(2.748)	(0.073)	(0.073)	(0.054)	(0.054)	(0.053)	(0.087)	(0.077)	(0.178)	(0.109)
Height (in cm)	86.668	–0.335	–0.053	0.090	0.058	0.085	–0.106	0.465	–0.437	0.802*
	(10.759)	(0.246)	(0.226)	(0.189)	(0.164)	(0.167)	(0.218)	(0.316)	(0.589)	(0.432)
Weight for height [31]	0.132	0.013	0.043	0.054	0.063	0.058	0.042	0.179***	0.267	0.211**
	(0.001)	(0.040)	(0.049)	(0.046)	(0.046)	(0.045)	(0.075)	(0.068)	(0.167)	(0.106)
Body mass index (weight/height^2)	15.425	–0.140	–0.131	–0.141	–0.119	–0.143	–0.008	–0.242	0.502**	–0.317
	(0.128)	(0.113)	(0.133)	(0.135)	(0.133)	(0.142)	(0.138)	(0.326)	(0.248)	(0.357)
Head circumference (in cm)	46.628	–0.107	–0.013	0.036	0.017	0.031	0.090	0.088	0.033	0.177
	(0.096)	(0.101)	(0.087)	(0.086)	(0.077)	(0.077)	(0.117)	(0.108)	(0.320)	(0.161)
Mid-upper-arm circumference (in cm)	15.699	–0.261**	–0.251*	–0.243*	–0.224*	–0.267*	–0.622	–0.097	0.146	–0.193

[31]The first column for weight for height is derived using the ratio of weight to height. The estimates reported in columns 2–8c based on regressing child's weight controlling for his/her height.

	(0.156)	(0.119)	(0.133)	(0.130)	(0.133)	(0.156)	(0.381)	(0.109)	(0.222)	(0.160)
Development Indicators:										
Communication skills-for-age	399.086	–3.427	0.936	1.726***	0.837	0.941*	1.278*	0.473	0.120	1.111
	(2.784)	(2.199)	(0.905)	(0.621)	(0.552)	(0.551)	(0.760)	(0.832)	(2.735)	(1.178)
Mobility skills-for-age	329.994	–1.948	–0.989	0.213	–0.242	–0.195	0.469	0.447	1.991	0.266
	(3.734)	(3.192)	(1.300)	(0.975)	(0.947)	(0.940)	(1.438)	(1.609)	(2.952)	(2.270)
Social-personal skills-for-age	327.503	–3.011	–2.386	–0.863	–1.885**	–1.884**	0.757	–3.472**	3.952	–2.279
	(3.595)	(3.074)	(2.003)	(0.924)	(0.897)	(0.889)	(1.501)	(1.487)	(3.164)	(2.225)
Number of observations	1,313	2,639	1,925	1,919	1,919	1,919	980	939	332	607
Panel Sample	N	N	Y	Y	Y	Y	Y	Y	Y	Y
Control for outcome in baseline	n.a.	N	N	Y	Y	Y	Y	Y	Y	Y
Controls	n.a.	N	N	N	A	B	B	B	B	B
Pairwise sub-district fixed effects	n.a.	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note: Robust standard-errors in parentheses. ***, **, * denote statistically significant at 1, 5, and 10 percent, respectively. Controls A=child’s age, child’s gender, HH’s gender, HH’s education; Controls B= Controls A+ household size, use wood to cook, recipient of BLT cash transfer, piped water, dwelling with dirt floor, and wall materials made of either brick, concrete, adobe or logs; community on a river. Results for worm eggs in columns 7 – 9 are restricted to panel children, with no controls in the baseline as these data are only available in the endline