

Understanding the Links and Interactions between Low Sanitation and Health Insurance in India Baseline report*

May 1, 2015

Orazio Attanasio, Britta Augsburg, Felipe Brugues, Bet Caeyers,
Bansi Malde, Borja Perez-Viana

Institute for Fiscal Studies



*This research was made possible through the generous support of the Strategic Impact Evaluation Fund (SIEF). Any errors are the responsibility of the authors.

Contents

1	Executive summary	4
2	Introduction	18
3	Project description	19
3.1	Project background	19
3.2	Project geographical focus	21
3.3	Description of the intervention	23
3.3.1	Sanitation loans	24
3.3.2	Sanitation awareness creation activities	25
3.4	Intervention objectives	27
4	Evaluation design	30
4.1	Randomised evaluation approach	30
4.2	Potential risk factors	32
4.3	Sample selection strategy	34
4.3.1	Selection and randomisation of study GPs	34
4.3.2	GP segmentation and listing	36
4.3.3	Matching listing dataset to GK clients database	38
4.3.4	Sample selection and sample size	39
4.3.5	Randomisation of branch order	42
4.4	Instruments for data collection	44
4.4.1	Listing questionnaire	44
4.4.2	Community (village) questionnaire	45
4.4.3	Household questionnaire	45
4.4.4	Women questionnaire	46
4.4.5	Men questionnaire	47
4.4.6	Monitoring data and rapid assessments	47
4.5	Problems in data collection	47
5	Findings	50
5.1	Study population (listing data)	52
5.2	Village profile (community survey)	56
5.3	Households in study sample (household survey)	66
5.3.1	General household characteristics	66
5.3.2	Household member characteristics	68
5.3.3	Income generating activities	70
5.3.4	Assets	72
5.3.5	Consumption	72
5.3.6	Savings, credit and insurance	75
5.3.7	Shocks	80
5.3.8	Sanitation	81
5.3.9	Water	88
5.3.10	Health care utilisation	89

5.4	Women versus men in study sample (individual woman and man survey)	93
5.4.1	Background and status	93
5.4.2	Women empowerment	95
5.4.3	Credit and savings	97
5.4.4	Social networks, group membership and political activity	98
5.4.5	Sanitation	99
5.4.6	Personal hygiene	110
5.4.7	Health	111
5.5	Children in study sample (individual woman survey)	117
5.5.1	General characteristics children	117
5.5.2	Health	118
5.5.3	Nutrition	120
5.5.4	Hygiene	123
5.5.5	Anthropometrics	125
6	Conclusions	127
A	Power analysis	130
B	More details on the segmentation procedure	131
C	More details on the baseline sampling strategy	132
D	Details study GPs	134

1 Executive summary

This document reports on the baseline data collection for the project titled “Understanding the Links and Interactions between Low Sanitation and Health Insurance in India”, funded through the Strategic Impact Evaluation Fund (SIEF). The overall purpose of this project is to shed light on (i) innovative ways of increasing the uptake and usage of safe sanitation practices and (ii) provide evidence on the links and interactions between improved sanitation and health insurance. It does so by studying two distinct but topically-linked projects: The smaller of these two projects is designed to explore the potential of providing primary community health insurance for free to communities that reduced open defecation conditional on sustaining this tendency. This component of the project is still in the development phase and will hence not be covered in this report. The second project, which includes a full randomised controlled trial impact evaluation, analyses two variants of an intervention, which in achieving sustainable improvements in household and community sanitation, aims to improve the health and reduce health expenditures of the poor in rural India – potentially reflected in lower health care claims volumes.

This report discussed the activities and findings of the baseline data collection for this RCT component. The two overarching aims are (1) to provide an interesting snapshot of our study population, serving as a useful tool to understand the context in which the intervention is taking place, and (2) to formally test whether we see any systematic differences between the treatment and control group prior to the intervention starting. We see this document as an important reference for processes followed, decisions made and their rationale, and related outcomes for everything relevant to the Impact Evaluation (IE) design and hope that it will serve as a useful guide for anyone interested in using the project’s data or understanding the analysis we will undertake going forward.

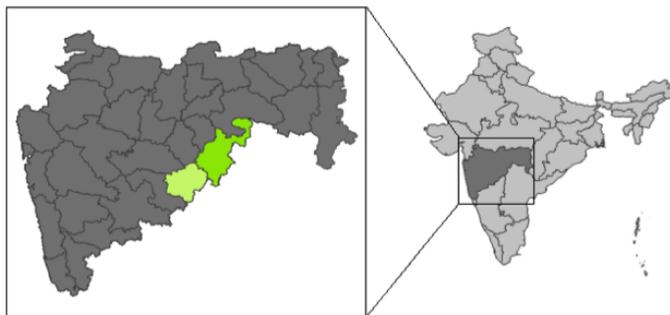
The ultimate aim of this project component is to use (primary) health insurance claims data as an innovative measure for health impacts of sanitation. The study is designed to test whether improvements in sanitation lead to lower (primary) health insurance claims, and thereby assess the feasibility of underwriting health insurance contracts based on sanitation ownership.

To answer this question, the first step in the project is to assess the effectiveness of a sanitation intervention in improving sanitation outcomes. Only if this is achieved, can we go to the next step and measure the impact of improved sanitation on health insurance claims. This report focuses on the baseline data collection for the sanitation intervention impact evaluation, which is implemented by the microfinance institution Grameen Financial Services Pvt. Ltd. and its NGO arm Navya Disha. These two organisations have different focuses, which we explore in the evaluation design, which includes two treatment arms and one control group. In the first treatment arm, potential credit constraints in sanitation uptake are alleviated by providing micro-loans for toilet construction. The second treatment arm is exposed to the same financial intervention plus awareness creation and other sanitation related activities. These interventions are described in detail in Section 3.3 of this report.

Overview of data collection

The study covers two districts in the Indian state of Maharashtra. These districts are Latur and Nanded and are depicted in Figure 1.

Figure 1: Geographical focus of the study



Within these districts, we cover 120 gram panchayats (GPs), the smallest administrative unit by the Government of India. These GPs were identified in collaboration with our implementing partner based on two dominant criteria: (i) they should fall within currently active operational areas and (ii) neither sanitation loans nor health insurance products should have been offered at any point in time to community members by our implementing partner. The process of sample identification and subsequent randomisation to treatment arms is comprehensively described in Section 4.3.

In each GP we then set out to conduct two data collection exercises: a household listing and the full baseline survey.

The listing survey started on 14 September 2014 and was completed within a month, on 12 October 2015. As can be seen in Table 1, 38,868 households were interviewed during the listing exercise, achieving a response rate of 95%. The main reason for non-response was non-availability of households when the survey took place. We discuss this exercise in detail in Section 4.3.2.

The baseline survey (which took place between 24 November 2014 and 26 January 2015) targeted a sample of 30 respondents per GP, implying an overall sample size of 3,600 households. The achievement is very close to this target, falling just short by 5 household interviews (see Table 1). These summary statistics hide an important detail that significantly complicated data collection activities: our evaluation design makes use of a stratified sampling approach, where households were stratified by (i) whether a client of our implementing partner lives in the household and (ii) whether a child under the age of two years lives in the household. The main complication was driven by the fact that, in order to avoid the survey contaminating the study design, the listing survey did not explicitly ask the respondent whether or not any household member

was a GK client. Instead, the identification of GK households was to be achieved ex-post listing, by matching the names of the (only female) clients in the GK clients database to the names of household members provided during the listing survey. The process and challenges faced are described in Section 4.3.3. Breakdowns on response rate achievements by strata are provided in Section 4.3.

Table 1: Study units, survey instruments and corresponding response rates

	Control	GK	GK + ND	Total
Randomisation unit - GPs	41	40	39	120
<i>Data collection activities: Listing</i>				
# of HHs visited during listing (total)	14,277	12,830	13,960	41,067
# of HHs available for interview (total)	13,482	12,337	13,049	38,868
Listing data response rate	94%	96%	94%	95%
<i>Data collection activities: Household survey</i>				
# of targeted BL HH respondents	1,230	1,200	1,170	3,600
# of achieved BL HH interviews	1,238	1,187	1,170	3,595
BL HH survey response rate	101%	99%	100%	99.9%
<i>Data collection activities: Individual survey - male</i>				
Male individual interview conducted	1,176	1,146	1,139	3,461 [†]
% of HH interviews with man survey	95%	96%	97%	96%
<i>Data collection activities: Individual survey - female</i>				
Female individual interview conducted	1,232	1,179	1,156	3,567 [†]
% of HH interviews with man survey	99%	99%	98%	99%
<i>Data collection activities: Community survey</i>				
# of targeted surveys	41	40	39	120
# of achieved surveys	40	40	39	119
BL HH survey response rate	98%	100%	100%	99.2%

[†]This number excludes observations for which no parent household record is available. 14 such records were excluded for the female sample and 13 for the male one.

We chose our four different survey instruments during the baseline survey (household, individual - male, individual - female, community) for the following reasons, which are also outlined in Sections 4.4.3, 4.4.4, 4.4.5, and 4.4.2:

1. **Household survey:** The household questionnaire, which is described in Section 4.4.3, was designed to (a) provide us with information on the baseline levels of the outcomes of interest for the study and (b) to collect characteristics of the household that provide a good description of the study population, poverty levels and wealth, to be used when investigating heterogeneous impacts and to help to improve power of our impact analysis. The household questionnaire was hence the most extensive module, covering socioeconomic characteristics of the household (assets, income, savings, credit, consumption expenditures), household member information (age, gender, education, etc) and a detailed section on sanitation and hygiene infrastructure, practices and beliefs. Our main outcomes of interest covered through this survey instruments can be summarized as follows:

- (a) **Primary Outcomes:** Sanitation Uptake, Uptake of safe sanitation, Usage of safe sanitation, health insurance claims;
 - (b) **Secondary Outcomes:** Health (diarrhea, child nutritional status), awareness about sanitation, changes in perceptions of costs and benefits of safe sanitation (for the Navya Disha intervention); uptake of credit, awareness about health insurance, uptake of health insurance, household income and consumption (affected by credit);
 - (c) **Tertiary outcomes:** Productivity, and schooling, among other.
2. **Individual woman survey:** The woman survey had four key purposes: (i) to collect information on individual sanitation behaviour (one of our primary outcomes) reported by the woman herself; (ii) to collect information on individual sanitation preferences and beliefs to understand women’s perceptions of the costs and benefits of sanitation, and also to identify what they value about it; (iii) to collect information on child health, child care practices and child nutrition for children aged < 5 years and (iv) to collect information on women’s status in the household.
 3. **Individual man survey:** The main purpose for the man questionnaire is in line with points 1 and 2 for the individual woman survey. On the second, sanitation preferences and perception, this is driven by anecdotal evidence that suggests that men and women value sanitation differently, and that this explains the slow take-up rates of sanitation.
 4. **Community survey:** The community survey was designed to collect information on factors that are expected to facilitate or potentially constrain the uptake and success of the interventions we study. Accounting for characteristics of the environment in which the intervention is implemented will be crucial to assess its success or understand failures. The instrument therefore covers information on the size, location, access/ connectivity, typical income generating activities, infrastructure, NGOs and services, the political economy, community activities, sources of water, shocks and prices.

Except for the community survey, these instruments were fielded using CAPI methodology. We describe the process, including piloting, in the respective sections mentioned above.

IE Design validation Table

As explained in Section 4.1, the evaluation methodology will be based on comparing the outcomes between the different treatment groups, i.e. control, GK (sanitation loans only) and 'GK + ND' (sanitation loans + awareness creation). In order to be able to attribute any effects to the sanitation interventions, it is imperative that the three groups being compared are similar in all respects at the outset of the intervention. To

test whether randomisation was properly done, we will compare the observable (pre-treatment) characteristics and test that there are no significant differences in their distribution between the different treatment arms.

Section 5 provides a detailed analysis of data collected during the listing and baseline. The sections focus on providing insight into the study context, while at the same time validating our IE design. We provide here a few summary tables of the key outcome and background characteristics.

We present tables showing the average values of different variables for each of the treatment groups. The key balance test will be based on the statistical joint 'F-test' to test whether overall there are any differences between any of the three evaluation groups. The advantage of this test is that it controls for the fact that we are making multiple comparisons, which a normal t-test fails to do. The disadvantage of this test is that if we reject the balance test, i.e. if the test suggests at least one difference, we do not know which treatments can be said to be significantly different from each other. We therefore also show the results of two-way comparisons between control and each of the treatment groups (as ultimately these will be the comparisons made in the impact evaluation), to see if any observed differences between the means are statistically significant at conventional levels. The box presented in Figure 13 in the main body of the document explains key statistical concepts that we use in the report.

Before proceeding, note that in all of the tables that follow, we use the following format: The first column gives information on which variable is concerned. We then show the mean for the whole sample (treatment groups and control combined). The following three columns show the mean of the control and the treatment groups separately. The fifth column shows the F-statistic of the test of statistical differences between any of the treatment groups and the second last column shows the associated p-value. The last column shows the total number of observations over which the whole sample mean is calculated. Statistical differences based on two-way comparisons with the control group, if any, are indicated with asterixes (*).¹

General Household Characteristics

Table 2 gives general information on our study households. Our typical study household is hindu (76%) and consists of 5 household members of which, for almost one in two households, one member is a child under the age of two years. The vast majority of households (93%) are headed by a male, with an average age of 44 years and 6 years of education.

Most households (97%) live in a dwelling they own, which is for 60% of households a semi-pucca construction² and for 23% of households a kucha building.

¹Note that throughout, the tests account for clustering of the standard errors at the gram panchayat level.

²*Pucca* stands for "strong", meaning made of materials like cement, concrete, oven burnt bricks, stone, timber etc. A semi-pucca dwelling has either the walls or the roof but not both, made of pucca materials.

Table 2: General Household Characteristics - Summary

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
HH religion: Hinduism	75.8 (1.634)	76.7 (2.414)	72.6 (3.428)	77.9 (2.481)	0.81	0.45	3595
Nr of HH members	5.43 (0.0465)	5.45 (0.0830)	5.36 (0.0738)	5.48 (0.0830)	0.67	0.51	3595
HHs with children <2 years	43.9 (0.858)	45.2 (1.503)	42.1 (1.276)	44.2 (1.619)	1.33	0.27	3595
Nr of children <2 years	0.47 (0.00962)	0.49 (0.0166)	0.44* (0.0147)	0.47 (0.0179)	1.84	0.16	3595
Gender HH head (fraction male)	92.5 (0.566)	92.2 (0.826)	92.5 (0.992)	92.9 (1.115)	0.11	0.89	3595
Age HH head	44.5 (0.314)	44.1 (0.631)	44.2 (0.485)	45.2 (0.475)	1.37	0.26	3595
Years of education HH head	6.02 (0.115)	6.09 (0.161)	6.34 (0.199)	5.63* (0.219)	2.97	0.055*	3433
Dwelling owned by HH	97.1 (0.353)	96.4 (0.654)	98.0** (0.413)	96.8 (0.699)	2.47	0.089*	3595
Dwelling structure: Semi-pucca house	60.3 (1.425)	61.8 (2.327)	57.4 (2.264)	61.6 (2.738)	1.14	0.32	3595
Dwelling structure: Kutcha House	23.4 (1.317)	24.2 (2.202)	21.4 (2.161)	24.5 (2.441)	0.60	0.55	3595
HH owns BPL card	13.9 (0.937)	14.3 (1.537)	12.4 (1.653)	15.1 (1.657)	0.73	0.49	3595

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Household Questionnaire.

Even though almost everyone in the sample holds a government benefit card (not shown), only 14% own one of the Below Poverty Line (BPL) kind, given only to the poor belonging to a vulnerable section of the society. Provided that the sanitation subsidy scheme of the SBM government program primarily targets BPL households (see Section 3.1), most households in our study sample may not be eligible for government support. This gives space for sanitation loans to potentially play an important role in tackling financial barriers to sanitation uptake.

While these and other household characteristics presented in Section 5.3 are for the large part balanced, we note that some imbalances are presented in this Table. Specifically, the F-test suggests a slight imbalance (significant at the 10% level) for dwelling ownership and years of education of the household head. Since both of these variables are expected to be determining factors in sanitation uptake decisions, we will have to ensure that we control for these variables in our impact analysis.

Household Economic Status

The next set of variables presented relate to the households' economic status. We present information on income, consumption expenditures, assets and related to credit, savings and insurance in Table 3.

We start by presenting overall descriptives of the total household income: Households in our sample earned in the last year an average of Rs. 60,365 (~USD 970).³ Employing some back-on-the-envelope calculations indicates that our study households

³This figure includes two hundred households who report not having received any income (in-kind or in cash) over the last year.

live on about US\$1.69 per person per day, putting our households slightly above the internationally accepted poverty line of US\$1.25 a day.⁴

Earnings come to a very large extent from agriculture-related activities, with 48% of the sample reporting to receive wages from agricultural labour and 34% deriving income from farming.⁵ Another important source of income are wages from employment outside the agriculture and allied sector, benefitting 30% of the households sampled. In line with these occupational patterns we find that a significant proportion of households (44%) own agricultural land. The average plot size is 4.6 acres. Amount of land owned is the only variable of those just discussed which displays some imbalances, with households in the GK treatment arm owning slightly more land. However, also this imbalance is only significant at the 10% level.

We ask households detailed questions on what they spend their earnings on. Based on some assumptions (all outlined in Section 5.3.5) we can calculate total yearly household expenditures. The mean value of this aggregate measure is Rs. 110,128 (~USD 1,769), almost one third of which is food expenditures (not shown). Note that the total estimated value of these consumables is higher than the reported annual income of Rs.60,365. This discrepancy is primarily driven by the fact that the average household produces at home or receives as gifts almost 20% of all food items it consumes.

Table 3: Household Economic Status - Summary

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Total HH income	60365 (3460.6)	60365 (4398.7)	56257 (5016.8)	63158 (8049.4)	0.43	0.65	3595
Income source: Wages agriculture	48.1 (1.459)	49.3 (2.563)	46.3 (2.329)	48.7 (2.653)	0.42	0.66	3595
Income source: Wages non-agriculture	30.2 (1.353)	28.9 (2.240)	28.9 (2.337)	32.9 (2.403)	0.96	0.39	3593
Income source: Business/Farm	34.0 (1.316)	35.5 (2.483)	33.0 (2.123)	33.4 (2.170)	0.31	0.74	3595
Agricultural land owned by HH - Acres	4.61 (0.142)	4.44 (0.218)	5.06* (0.259)	4.34 (0.249)	2.37	0.098*	1570
Total consumption expenditure (last year)	110128 (1775.9)	110128 (3196.1)	106506 (2724.2)	110985 (3186.5)	1.23	0.29	3595
HH knows credit source	61.0 (1.341)	61.3 (2.046)	62.6 (2.461)	59.1 (2.427)	0.52	0.60	3593
HH taken loan last year	22.0 (1.491)	22.2 (2.565)	18.8 (2.179)	25.1 (2.866)	1.59	0.21	3592
Amount outstanding debt	49668 (5536.5)	49668 (8291.3)	41653 (6090.2)	53820 (11681.4)	0.71	0.49	792
HH has savings	24.9 (1.633)	25.0 (2.680)	26.2 (2.741)	23.2 (3.044)	0.27	0.77	3593
HH has insurance	19.8 (0.968)	17.2 (1.519)	22.0** (1.651)	20.3 (1.779)	2.35	0.100*	3592
HH insurance type: Health	12.0 (1.404)	14.1 (2.740)	8.08* (2.008)	14.3 (2.457)	2.53	0.084*	711
HH insurance type: RGJAY	8.02 (1.211)	8.45 (1.966)	6.54 (2.027)	9.24 (2.229)	0.44	0.64	711

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Household Questionnaire.

⁴The calculated value adjusts for the purchasing power parity conversion factor. Without doing so, we get to US\$0.76 per person per day. Details are provided in Section 5.3.3

⁵This category also includes other type of businesses. However, for 94% of the households this business is a farm (see Table 42).

The lower half of Table 3 focusses on financial access of our study population. One can see that while households are aware of possible credit sources, only 22% have taken a loan of more than Rs 500 in the last year. The percentage of households with a loan outstanding (not shown) is comparable and if they do, the average outstanding amount is Rs.49,668 (~USD 798), which amounts to a bit more than 80% of the sample average yearly household income.

Just like income and assets, consumption expenditures are nicely balanced across treatment arms, suggesting that the experimental design was successful in randomly allocating households of varying income groups to the different samples.

The percentage of households with savings stands with 24% slightly higher than those who took credit in the last year. And, finally, just about 20% of households have some type of insurance policy. The most common insurance type is life insurance (not shown), held by almost 80% of those with a policy. We show in this summary table the average percentages for private health insurance policies and policies under the Rajiv Gandhi Jeevandayee Arogya Yojana (RGJAY) health insurance scheme (sponsored by the state government of Maharashtra), which are 12% and 8% respectively. We observe that households residing in communities allocated to the GK treatment arm are more likely to own insurance. While the F-test is significant at the 10% level only, it will be important to take these baseline values into account in future analysis, particularly when getting to the second part of this project where we assess the impact of improved sanitation on health insurance claims.

Sanitation and Health

The next table (Table 4) provides an overview of sanitation infrastructure and practices of our study households. As with the two tables presented so far, all variables are constructed with data collected in the household survey. Hereafter, we will present a number of indicators collected at the individual level.

We find that less than a third of households (31%) own a toilet, out of which 96% are reported to be currently in use. Toilet ownership at baseline (November 2014 - January 2015) is only slightly higher than the 28% coverage that was reported during the listing survey that took place two months prior to the baseline survey (see Section 5.1). These figures about existing sanitation facilities were validated through direct observation by the data collection team, which we find to be highly correlated (90%) with the self-reported measure. Section 5.3.8 goes into details on the type of toilet, construction materials used as well as other features of the toilet. In general, they are improved toilets, reflected in an average value of Rs. 26,527 (~USD 417), which came predominantly from the households' own savings. Few households had access or made use of external funding to cover the costs of the toilet and only 12% benefitted from any type of subsidy.

This reported construction cost is significantly higher than either the sanitation loan offered by our implementing partner (Rs. 15,000) or the subsidy provided by the government (Rs. 12,000). These results are suggestive that only those with already the means to have a toilet built undertook its construction. ⁶Despite this high cost, only

⁶Supporting this hypothesis are statistics presented in [Raman and Tremolet \[2010\]](#), which reports

13% of households that do not own a toilet think that it is too expensive. At the same time though, it is clear that financial constraint are a major hurdle to uptake: When asking the same households (those without a toilet) why they do not own one, the vast majority (83%) responds that they are not able to afford a toilet. This is despite the universal government scheme which subsidises sanitation construction, and hence implying a role for sanitation loans as considered in this study. Further encouragement for the intervention under consideration is provided by the fact that more than half of those households without a toilet at least theoretically support the idea of taking a loan for the construction of sanitation facilities.

All of these variables related to sanitation in our study communities are very nicely balanced. While we see one star on the F-stat for having funded a toilet through informal loans, we note that this variable has very little variation, which is likely the driving force behind this small imbalance.

Table 4: Sanitation, Hygiene and Health - Summary

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
HH owns toilet	30.5 (1.737)	28.0 (2.788)	35.1* (2.783)	28.5 (3.313)	1.95	0.15	3595
HH's toilet in use	95.5 (0.835)	95.4 (1.699)	96.6 (1.102)	94.3 (1.572)	0.78	0.46	1097
Cost of toilet (Rs.)	26527 (712.4)	26527 (1433.0)	26684 (1044.2)	26610 (1224.2)	0.028	0.97	922
Source of funding toilet: Savings	87.1 (1.791)	86.5 (3.068)	86.3 (3.557)	88.9 (2.076)	0.32	0.73	1097
Source of funding toilet: Loan (formal)	0.27 (0.157)	0.29 (0.283)	0.48 (0.338)	0 (0)	1.52	0.22	1097
Source of funding toilet: Loan (informal)	2.01 (0.484)	2.88 (1.029)	0.72* (0.529)	2.70 (0.915)	2.85	0.062*	1097
HH no toilet: cannot afford it	83.1 (1.100)	83.1 (1.795)	83.6 (1.896)	82.6 (2.019)	0.077	0.93	2498
HH no toilet: find it too expensive	12.6 (1.000)	14.4 (1.791)	12.5 (1.809)	10.9 (1.524)	1.10	0.33	2498
HH would borrow to build toilet	54.6 (1.386)	56.7 (2.205)	52.1 (2.679)	54.7 (2.299)	0.88	0.42	2476
HH purifies water	78.4 (1.561)	76.1 (2.784)	82.0* (2.151)	77.4 (3.018)	1.64	0.20	3595
HH stores water	63.2 (1.544)	62.6 (2.803)	66.0 (2.304)	61.0 (2.812)	1.03	0.36	3594

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Household Questionnaire.

The last few variables in Table 4 tell us a bit more about hygiene practices of study households: 78% report to purify their drinking water, although it is worth noting that this is done in a rudimentary manner by filtering it with a cloth, a method that does not eliminate parasites and bacteria from the water. Further, on average 63% of households store water, which is reflective of the water access situation and time spent on collection water, which we discuss in detail in Section 5.3.9.

on toilet construction costs in three districts of Maharashtra (Chandrapur, Kolhapur and Nashik). The costs reported for a typical toilet of a household classified as APL is reported at on average USD332, which is the equivalent of USD373 in December 2014 USD value (the time of our baseline survey). This average hides some variation. The average reported costs in Kolhapur for example is as high as USD 434 for an average APL toilet, slightly above the average reported by our study households.

The next two Tables we discuss show some of the individual sanitation behaviour we discuss in detail in Section 5.4 of this report.⁷ Both tables report the same set of variables, the difference being that responses presented in Table 5 were provided by the female and in Table 6 by the male respondent.⁸ We already discussed above that about 30% of households own a private toilet and that the great majority of these are used. Individual sanitation habits reported for those that own a toilet are in line with these statistics and so we do not reproduce them here. We show though that the majority of male and female respondents report to go for open defecation at a minimum walking distance of 5 minutes. The percentage is slightly lower for males (65% versus 70%), who more often defecate closer to their home (unreported in this table). It comes at no surprise that water and soap is rarely available at the site of open defecation. For households with toilets, we find that more than one third of men (38%) and one quarter of women (27%) report not having access to water at their dwelling. This is in line with findings reported above on access to water for sanitation, which only about a third of households reported to be piped water into the house and considerable average walking times to collect water (see Section 4.4.2). Of those who do have water available, the majority also seem to have soap (88% for woman and 75% for men), not reported in this Table.

About one fifth of those going for OD (20% of males and 15% of females) report to be satisfied with their OD and the place they frequent.

Table 5: Individual Sanitation, women - Summary

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
OD (>5 mins home)	69.5 (1.654)	71.1 (2.824)	65.6 (2.546)	71.7 (3.111)	1.52	0.22	3434
Satisfied with OD (>5 mins home)	15.2 (1.661)	16.6 (3.096)	17.3 (2.882)	11.8 (2.510)	1.26	0.29	2386
Hand-washing facility at site - toilet	72.9 (3.353)	80.5 (4.459)	73.7 (4.720)	63.9* (7.152)	1.99	0.14	987
Hand-washing facility at site - OD (>5min)	1.45 (0.504)	1.82 (1.029)	0.96 (0.577)	1.51 (0.890)	0.33	0.72	2351
Even if HH has toilet, HH members don't use it	22.5 (2.080)	23.9 (3.960)	22.8 (3.550)	20.7 (3.200)	0.22	0.80	3434
If people OD, nobody minds as is common habit	26.4 (2.232)	27.7 (4.211)	30.1 (3.770)	21.2 (3.399)	1.67	0.19	3434

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Household Questionnaire.

⁷We also discuss there reasons for changes in sample sizes that can be observed across variables in these presented summary tables.

⁸See Section 4.3.4 for the description of how this selection was done.

Table 6: Individual Sanitation, men - Summary

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
OD (>5 mins home)	64.8 (1.877)	66.6 (3.136)	60.9 (2.729)	66.8 (3.762)	1.27	0.29	2923
Satisfied with OD (>5 mins home)	20 (2.693)	20.9 (4.291)	21.1 (5.166)	18.0 (4.598)	0.14	0.87	410
Hand-washing facility at site - toilet	62.1 (2.940)	64.1 (4.506)	62.3 (4.867)	59.6 (5.768)	0.19	0.83	817
Hand-washing facility at site - OD (>5min)	5.07 (0.698)	5.27 (1.166)	5.68 (1.376)	4.27 (1.082)	0.37	0.69	1757
Even if HH has toilet, HH members don't use it	20.0 (2.348)	22.5 (4.498)	15.6 (3.543)	21.8 (3.982)	1.01	0.37	3434
If people OD, nobody minds as is common habit	29.2 (1.464)	28.1 (2.359)	26.4 (2.359)	33.2 (2.758)	1.85	0.16	3434

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Household Questionnaire.

We finally present two variables expressing beliefs about sanitation practices: For one, we report the interesting finding that 20-23% of those households that do not own a toilet are of the opinion that even if households have a toilet, household members do not necessarily use it. Women are slightly more likely to hold this view. Male and female individual respondents are in similar agreement about their belief that nobody in their community minds about open defecation as this is a common habit (26% of women ascribe to this opinion and 29% of males do).

All of these variables at the individual level are balanced across treatment arms.

We finally show a set of variables related to self-reported measures of health. A common outcome considered in sanitation studies is diarrhea incidence. We ask our individual respondents whether they suffered from diarrhea, using a recall period of 7 days. We also ask the female respondents, if they are mothers of children under the age of 2 (slightly more than half of the sample), whether their child had any diarrhea within that same time period.

As can be seen in Table 7, about 5% of women and children suffered from diarrhea, and 1% of men.

A considerably larger percentage of household members received any type of medical care in the last month (17%) and almost a quarter of households report that a household member was hospitalised for at least one night in the last year. Reasons and other related statistics are presented in Section 5.4.7. As with variables describing the sanitation situation and practices in our study villages, also indicators related to health do not display any significant imbalances across the study arms.

Table 7: Health indicators - Summary

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Female - diarrhoea in last 7 days	5.24 (0.646)	5.47 (1.202)	4.31 (1.151)	5.96 (0.964)	0.61	0.54	3432
Male - diarrhoea in last 7 days	1.14 (0.392)	1.26 (0.701)	0.44 (0.437)	1.69 (0.813)	1.13	0.33	703
Child (<2yrs) - diarrhoea in last 7 days	5.85 (0.631)	5.34 (0.987)	6.60 (1.212)	5.67 (1.069)	0.33	0.72	1897
Medical care received	16.8 (1.095)	16.0 (1.930)	16.5 (1.946)	18.0 (1.791)	0.33	0.72	3414
HH member hospitalized, last year	24.4 (1.039)	24.3 (1.715)	22.4 (1.587)	26.5 (2.010)	1.33	0.27	3416

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Household Questionnaire.

Our results from the treatment control comparison are generally very encouraging. The number of imbalances found fall well within the expected proportion of ten percent and those imbalances observed are for the most part significant only at the 5 or mostly at the 10% significance level. We further find no noteworthy significant differences in key variables across treatment and control units, particularly in the sanitation situation and health outcomes. This is important since it implies that our treatment and control households were not significantly different in terms of their sanitation practices and starting level of their health status before the start of the intervention.

Despite these encouraging overall results, it is worth pointing out that we observe more imbalances between the GK treatment arm and the control group than we do between the GK+ND treatment arm and the control group. Findings suggest that GK communities are somewhat better off with respect to assets they own. Households in GK communities are for example more likely to own their dwelling, the dwelling type owned is more likely to be of strong (pucca) structure, and they own more land than households in the control communities. It is likely that we observe a number of imbalances due to correlations between some of the variables. Likely related to the three variables just mentioned is for example the length of time households lived in a village. In line we find that one of the other significant differences observed is that households in the GK treatment arm are more likely to have been born in the village they now live in. We also note though that, while household in GK communities seem to be richer in terms of some assets, they do report to earn the same levels of income and have similar yearly consumption expenditures. These results suggest that households residing in the GK treatment arm are not necessarily systematically richer.

In any event, it will be important to account for characteristics for which imbalances are observed when we conduct the impact analysis and check for robustness of findings to the in- and ex-clusion of these characteristics.

In the final section of this executive summary we detail any other potential threats to the IE design and how we plan to monitor and, if necessary, mitigate them .

Significant risks to IE design

At this stage of the project we perceive that the risks to the IE design are low on the data collection side: We were able to collect our key indicators satisfactorily, and we have a high response rate to our surveys. We describe data collection problems we faced in Section 4.5, but believe that we were able to solve these in the process in a way that would not significantly affect our design.

Higher (albeit not 'high') risks are identified with respect to take-up of treatment and contamination. We present these in Section 4.2 and reproduce here a slightly shortened version of our five potential sources of concern:

1. There could be **non-compliance with the randomisation** by field and/or branch level staff of the implementing institutions. To prevent this, the research team has visited each of the implementing partner's field offices to give an in-depth training about the research to all stakeholders involved. Moreover, our implementing partner has programmed into their management information system that sanitation loan applications from control areas cannot be processed. If this is tried, head office receives a warning. That means that it is technically not possible for anyone in the control GPs to receive any sanitation loans from our implementing partner. This will also be monitored using GK loan monitoring data (see Section 4.4.6).
2. **Contamination** in our design might arise through sanitation information spilling over to neighbouring GPs, which may be control GPs. To be able to account for this, we will collect information on possible interactions between neighbouring GPs, e.g. common markets, and common branch level GK meetings, characteristics of GK branch officers, and distances between GPs.
3. **Other programs** (particularly by the GoI) could intensify their efforts at improving sanitation infrastructure across study areas and thereby increase sanitation density in our control GPs. To some extent, this is not a problem for our evaluation design, as long as the government's efforts are similar across both treatment and control GPs, and as long as the government does not manage to get toilets to all of the households in the control group. As described in Section 3.4, the objective of the intervention is to complement government's efforts rather than substituting for it. In our context - and in fact in most other evaluations of development impact - control communities represent cases of "business-as-usual" government activity, rather than "doing nothing" (Ravallion 2008). The impact we are interested in relates to the contributions of GK and ND's activities *in addition* to the government's program. However, if the government scales up its efforts to the extent that most households in the control group acquire a toilet, the potential for impact from ND and GK is reduced. Even in that case, however, we may still expect there to be an impact on toilet quality rather than toilet ownership. Section 3.4 above lists various channels through which we think GK and ND can make a difference despite the accelerated efforts that are being made by the government. Moreover, the results of the two rapid assessment surveys

which will be carried out between baseline and midline will allow us to monitor this potential risk.

4. In terms of **attrition**, we anticipate low attrition among GK clients given their long-standing relationship with GK. To reduce attrition particularly among non-clients for all surveyed households, we will collect mobile phone numbers of household members and contact details of relatives and friends who are likely to know of their whereabouts if they move. As will be discussed below, our baseline findings show that only around 4% of the households in our sample reported to have migrated in the year prior to the survey. This is consistent with other studies in areas near our study area, which have encountered relatively low attrition rates, driven predominantly by migration (at 3-4%, see for example Mahal et al 2012). Moreover, our baseline results show that more than 97% of all households report it to be unlikely that they will have moved in a year from now.
5. We might have a **threat to our power** given unanticipated policy changes by our implementing partner which affect loan eligibility. Right after the baseline data collection had taken place, a change was introduced which allows only clients that have been with the organisation for at least one year to receive a sanitation loan. As we describe in the body of this document, we will monitor this risk using loan uptake data (see Section 4.4.6 for more information). Once the programme has been running for some time, we will be able to judge whether this change is a serious risk to power or not. The planned rapid assessment will additionally help in this process.
6. There is a possibility that some of the **kendras** that are part of our study group **close down** between baseline and endline, which would imply a reduction in our effective GK sample size. This risk will be followed up very closely by examining the GK monitoring data.
7. Finally, we note that there are currently some undefined parameters around the implementation of the (primary) **health insurance component**, which is crucial for the second part of the study. The coming months will focus on narrowing down the necessary steps and putting things for this component in place.

2 Introduction

Poor sanitation has obvious implications for public health and provision of safe sanitation has thus been recognized to be an indispensable element of disease prevention and primary health care programs (e.g. the Declaration of Alma-Ata, 1978). Moreover, lack of safe sanitation is acknowledged to affect broader outcomes such as productivity and investment, which ultimately constrain economic growth (WSP, 2010). Nonetheless, there has been weak overall investment in improving sanitation infrastructure in poor countries, evident from the fact that the Millennium Development Goal (MDG) on sanitation fell short of its target to halve the 1990 level of proportion of people without sustainable access to sanitation by 2015 (WHO-UNICEF, 2014).

Poor sanitation is a particularly important policy issue facing India, which accounts for over half of the 1.1 billion people worldwide that defecate in the open (WHO-UNICEF, 2014). The Indian Government (GoI) has shown strong commitment to improving sanitation, establishing the Total Sanitation Campaign in 1999, which was revamped in 2012 as the Nirmal Bharat Abhiyan (NBA) policy and most recently in October 2014 as the Swachh Bharat Mission (SBM). This policy aims at attaining 100% Open Defecation Free India by 2019 (GoI, 2014). However, despite these efforts, safe sanitation uptake and usage remains low. For instance, the 2011 Indian census reports that almost 50% of Indian households do not have access to a private or public latrine. This highlights the need for novel approaches to foster the uptake and sustained usage of safe sanitation in this context.

In light of these challenges, researchers from the Institute for Fiscal Studies, with support from the FINISH Society⁹, teamed up with the Water and Sanitation Program (WSP), Grameen Financial Services Pvt. Ltd. (henceforth referred to by its popular name 'Grameen Koota' of just 'GK') and Navya Disha (ND), to shed light on (i) innovative ways of increasing the uptake and usage of safe sanitation practices and (ii) provide evidence on the links and interactions between improved sanitation and health insurance. It does so by studying two distinct but topically-linked projects in rural Maharashtra, India: The smaller of these two projects is designed to explore the potential of providing primary community health insurance for free to communities that reduced open defecation conditional on sustaining this tendency. This component of the project is still in the development phase and will hence not be covered in this report. The second project, which includes a full randomised controlled trial impact evaluation, analyses two variants of an intervention, which in achieving sustainable improvements in household and community sanitation, aims to improve the health and reduce health expenditures of the poor in rural India – potentially reflected in lower health care claims volumes.

More specifically, the interventions offer on the one hand sanitation loans to members of women groups which can be used to construct toilets or water connections and on the other hand a package of awareness creating activities - e.g. street plays, wall banners,

⁹FINISH stands for Financial Inclusion Improves Sanitation and Health and the programme is a response to the preventable threats posed by poor sanitation and hygiene. It was launched in 2009 as a new approach to improve the health and welfare outcomes of poor households. This approach focuses on financial tools to improve the sanitation situation in both rural and urban areas in India.

sanitation and hygiene workshops, etc. The rationale of these interventions is to improve toilet uptake and usage by i) lifting credit and liquidity constraints faced by poorer households and ii) improving awareness about the usefulness and cost-effectiveness of low cost safe sanitation. The longer term objective is to improve a wider set of outcomes including health, income, etc.

A rigorous impact evaluation, based on a randomised control trial, will evaluate the relative impact of these two interventions over a two-year period. The evaluation design includes a baseline survey before the intervention, as well as a midline survey after one year of the program and an endline survey after two years of the program. The study will concentrate on 120 GPs in two districts (Latur and Nanded) in rural Maharashtra, India.

This report provides details of the sampling methodology and the baseline data collection process. It also describes our study population with regard to a wide spectrum of community, household and individual household member characteristics. Whilst doing so, we test for any systematic differences between treatment and control groups which could undermine the validity of our randomised evaluation design.

3 Project description

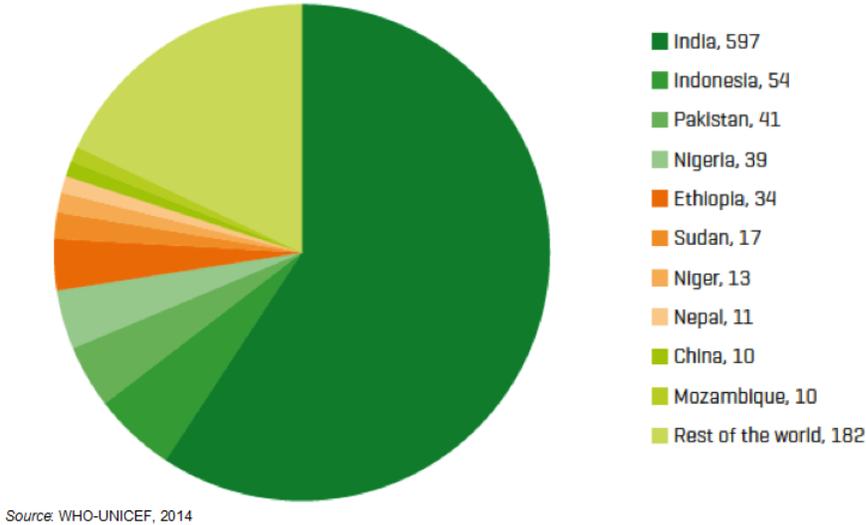
3.1 Project background

According to the most recent Joint Monitoring Program report for Water and Sanitation ([WHO-UNICEF, 2014](#)), sanitation has significantly improved globally: Sanitation coverage has increased from 49% in 1990 to 64% in 2012 and open defecation has fallen from 24% to 14%, respectively. Behind these average numbers, however, lie staggering disparities across countries. Whereas countries like Ethiopia, Cambodia and Nepal experienced noteworthy successes, some countries are lacking far behind. As shown in Figure 2, India currently tops the world ranking of the number of people still practicing open defecation in 2012 ([WHO-UNICEF, 2014](#)). More than half of the 1.1 billion people worldwide who practice open defecation live in India, which is more than ten times the number of any other country. Proportionately, this number amounts to 48% of the total Indian population.

Lack of improved sanitation can have disastrous consequences. Recent studies by [Spears \[2012\]](#) and [Kumar and Vollmer \[2012\]](#) suggest that improved sanitation decreases the risk of contracting diarrhea and associated infant mortality. Open defecation has also been associated with stunting among children and impaired cognitive development ([Pruss-Unstun et al., 2008](#); [Humphrey, 2009](#); [Dangour et al., 2013](#); [Spears, 2013](#)). Moreover, lack of safe sanitation is acknowledged to affect broader outcomes such as productivity and investment, which ultimately constrain economic growth. The Water and Sanitation program (WSP, 2010) of the World Bank (WB) estimates that poor sanitation costs India US\$48 per person per year, the equivalent of 6.4% of the country's gross domestic product (GDP) annually.

To address these challenges, in 1999 the Government of India (GoI) launched its ambitious Total Sanitation Campaign (TSC), which has been described as the largest

Figure 2: Top 10 countries with the highest numbers of people (in millions) practicing open defecation in 2012



sanitation initiative in the world (Hueso and Bell, 2013). The TSC included community awareness campaigns, provision of sanitation funds to communities to build sanitation infrastructure in public places such as schools and hospitals, and provision of small subsidies to individual below-poverty-line (BPL) households after they could demonstrate having constructed their own toilets.

In the year 2003, the TSC introduced the Nirmal Gram Puraskar (NGP) or 'clean village' award scheme. This scheme offers rewards (ranging from US\$1,000 to US\$10,000, depending on the population size) to local governments that achieve 100 percent open defecation free status and ensure total sanitation. Since its introduction in 1999, the TSC program has been redesigned twice: once in 2015 to become the Nirmal Bharat Abhiyan (NBA) policy and most recently in October 2014 to become the Swachh Bharat Mission (SBM). In essence, each reform implied an expansion of range rather than a paradigm change, including for example increases in the subsidy amounts.

Studies analysing the effectiveness of financial incentives for the continued usage of preventive health behaviors find that though successful in promoting simple behaviors such as one-time visits for preventive health checks, there is little supportive evidence of such incentives generating long-term changes. In the context of sanitation, Spears [2012] shows that the NGP prize in India has been successful in driving sanitation uptake, but little is known about the effectiveness of financial incentives in promoting sustained sanitation usage. Indeed, available evidence suggests that the NGP prize is not effective in sustaining long term sanitation usage (UNICEF, 2008). Recent figures from the Indian Ministry of Drinking Water and Sanitation 2012 baseline survey show that 39% of households reported to own toilets, but that two out of 10 of those toilets are reported to be out of use. Critics argue that the government programs yield poorly constructed toilets and that they do not sufficiently address the population's insufficient desire to construct, maintain and use toilets. Moreover, in other contexts some have argued that government subsidies do not always reach the poor (Jenkins and Scott,

2007; Jenkins and Sugden, 2006).

Possible constraints to health investments in general and toilet adoption and usage in particular include a lack of information, and credit and liquidity constraints. For instance, Pattanayak et al. [2009] provides evidence that the former is an important constraint hampering sanitation uptake: Latrine ownership in a rural Indian setting increased by 30 per cent on the provision of information, while consistent with the latter constraint, Tarozzi et al. [2011] find in a study in rural Orissa, that uptake of mosquito nets jumped from 2% to 52% when micro-loans were made available for their purchase. The intervention subject to evaluation in this study will relax both of these constraints by providing information on the benefits of safe sanitation and available low cost sanitation technologies and making available micro-loans for sanitation construction.

The first component of our intervention, i.e. the provision of sanitation loans, will be delivered by Grameen Financial Services Pvt. Ltd., popularly known as Grameen Koota (GK). GK was founded in 1999 as a project of nongovernmental organisation T Muniswamappa Trust and has since become an independent non-banking finance company (NBFC). GK is actively engaged in the microfinance sector in Karnataka, Maharashtra and Tamil Nadu states and provides a range of financial (primarily microcredit and micro insurance) and non-financial services to groups of women of age 19 to 55 years in rural and semi-urban low income households.¹⁰

Grameen Koota provides a wide range of loans, including emergency loans, festival loans, medical loans, income generating activity loans, etc. Since 2009, GK started providing microcredit for the construction of sanitation systems to its clients.

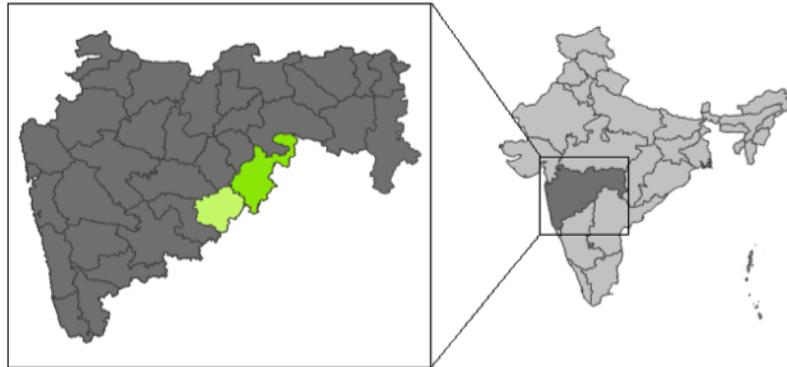
Understanding that providing finance only is not sufficient to reach high sanitation density, and that hygiene promotion also plays an essential role, GK further created an NGO, the Navya Disha (ND) Trust, which helps its clients understand the benefits of safe sanitation and available toilet technology and infrastructure, the materials needed, and procedures for procuring parts, labour, and government approvals. In the intervention subject to evaluation in our study, ND will be in charge of delivering the sanitation awareness creation component.

3.2 Project geographical focus

Our study will concentrate on the south-eastern area of the Maharashtra state, districts Latur and Nanded (see Figure 3), where GK has branch offices. Maharashtra, with its capital Mumbai, is one of the largest Indian states, counting approximately 100 million people living in almost 44,000 villages (Census, 2011). While this is the second richest state in the country in terms of per capita income, incidence of poverty remains close to the national average, implying severe inequalities within the state. According to the last human development report of Maharashtra (Government of Maharashtra, 2002) “Maharashtra’s economy has demonstrated a strong track record of growth and is a visible success story for the rest of the country but its weakness is its uneven distribution of the gains of the growth.”

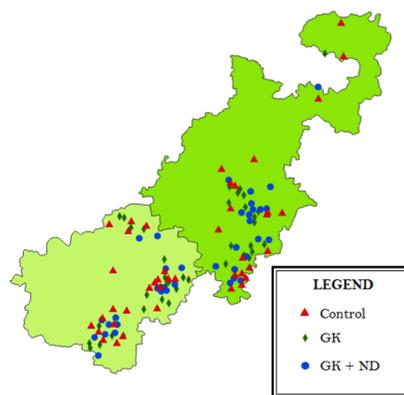
¹⁰As of 2014, GK reported total assets of USD 161 million and a gross loan portfolio of USD 141 million outstanding to 543,000 active borrowers (MIX). By February 2015, the number of active clients had further increased to 863,198 active borrowers.

Figure 3: Geographical focus of the study



The study districts Latur and Nanded belong to those areas that have only moderately benefitted from the growth. According to the latest District Level Household and Facility Survey (DLHS-4) collected in 2012-13, these districts face incidences of poverty of on average 35 per cent, only 17% of the households use some type of toilet facilities and 40% (77%) of (female) household heads have not attended any school. Access to health services is also very poor, with only 13 per cent of villages having a primary health center available in their village and the nearest government hospital being on average 63km away.

Figure 4: Location of the 120 study GPs



Within Maharashtra, our study will cover 120 GPs located in Latur and Nanded districts, a complete list of which is provided in Appendix D. Figure 4 shows the location of each of the study GPs, with an indication of their 'treatment' status in the study (this will be discussed in detail in Section 4 below).

These GPs were randomly selected from the total list of rural GPs that are serviced by five branches of Grameen Koota in Maharashtra (i.e. where they have a client base) but where no sanitation loans or health insurance had been provided and no ND activities had taken place to date. Section 4.3.1 describes the selection of these GPs in more detail. The selected five branches lie in administrative blocks Degloor, Udgir, Ahmadpur, Naigaon and Nilanga.

According to the 2012 baseline survey conducted by the Indian Ministry of Drinking Water and Sanitation (Baseline survey, 2012), on average 27.4% of households in our study GPs reported to have a toilet in 2012. Note that this number hides information on actual usage of the toilet, which is an additional indicator we will consider in our study.

Note also that the set of GPs in which GK is operational is a result of a careful selection exercise by GK head quarters (e.g. has to be politically stable, certain number of women, etc) and can therefore *not* be considered representative for the state of Maharashtra.

3.3 Description of the intervention

The first year of our study will focus on the evaluation of the relative impact of two variants of a sanitation intervention:

1. The first intervention is the provision of sanitation loans up to Rs. 15,000 to Grameen Koota clients, against an interest rate of 22% per annum over a 2-year repayment period. These loans can only be used for the construction of a new toilet or for repair of an old toilet. In addition, GK offers water connections loans of amounts up to Rs. 5,000. Sub-section 3.3.1 discusses further details of these loans (e.g. loan caps, eligibility criteria, etc).
2. The second intervention, which is a variant of intervention 1, combines the provision of sanitation loans to GK clients with a package of sanitation awareness creation activities run by Navya Disha. Education and awareness creation on sanitation issues are targeted to GK staff, GK members *and* the broader community. This is done through community-level activities such as theatre plays, wall banners and information sessions for GK clients at their weekly joint liability group meetings (see below) and GK branch level workshops. Furthermore, ND and GK engage with the sarpanch to gain support for their activities and to strengthen sanitation and hygiene awareness. This is important because the GP plays a pivotal role in the implementation of the GoI sanitation program, the SBM. Lastly, ND organises mason trainings. Sub-section 3.3.2 provides more details about the specific package of awareness creation activities that will be run by ND in our study area.

3.3.1 Sanitation loans

Sanitation loan eligibility criteria

Basically any GK clients is eligible to apply for a sanitation loan. There are no collateral requirement and the only constraint recently (in March 2015) introduced is that a client needs to have been with GK for more than 1 year to be eligible for this loan product. To become a GK client, women must be between the ages of 19 and 55 years and form into groups of 5-10 members. Multiple such women groups in a GP are then grouped together to form a so-called GK kendra. The purpose of the kendra is mainly for the management of weekly loan repayments which take place at the GP level (see below). Each kendra has a maximum of 30 members.

In order to track loan repayments and loan utilisation, each client is required to hold a GK passbook, in which all loan repayments are registered and the results of loan utilisation checks are kept. Sanitation loans are provided for the construction of a new toilet or for repair of an old one. GK does not put any restrictions on the type of toilet the beneficiary decides to build, except that they advise against single pit technologies. GK and ND staff (see below) are trained to provide advice on different models, but the ultimate choice is left to the client.

GK kendra managers, who are in charge of providing funds and collecting repayments on a weekly basis, are instructed to provide a sanitation loan only after the client has clearly demonstrated her intention to build a toilet, e.g. by having provided space and having digged a pit. The kendra managers are in charge of conducting a series of loan utilisation checks and to note the results of these checks in the kendra member's passbook. However, at present no sanctions are imposed in case loans are used for any other purposes. Lack of reinforcement can have implications for actual loan usage and needs to be taken into consideration in the analysis of our study.

Sanitation loan caps and costs

Today, GK sanitation loans cover a maximum amount of Rs. 15,000 (which is considered sufficient to build a quality toilet), charging a 22% interest rate per annum at a declining balance over a 2-year repayment period. This makes sanitation loans relatively attractive, as the amount received is higher than any other type of loan provided by GK and the interest rate is lower than those charged for any other loans (typically 25%).

In addition to the interest, loan costs include a processing fee of 1.1% of total amount and Rs. 306 life insurance premium. Each client can obtain one sanitation loan only, but clients can take an additional water connection loan of Rs. 5,000. The overall loan cap on total loans taken from GK is Rs. 35,000 for new clients and Rs. 40,000 for women who have been client for longer than 3 years. The overall loan cap on total loans from any microfinance institution in India currently stands at Rs. 100,000.

Sanitation loan disbursement and repayment process

The process from sanitation loan application to money disbursement can take any time up to 4 weeks. Like most of GK's loans, the GK client is required to physically go to the GK branch office for sanitation loan disbursement¹¹. Clients repay the loan on a weekly (Rs. 179) or bi-weekly (Rs. 358) basis during their weekly meetings in the village with a GK kendra manager. Loan groups are held jointly liable for repayment of a loan. That means that if one group member defaults on any loan, no one else in the loan group can take out a new loan.

3.3.2 Sanitation awareness creation activities

Whereas sanitation loans are targeted at and available only for GK clients, Navya Disha targets education and awareness creation more widely, i.e. to GK members, GK staff, GP officials and any other residents of the GP community.

First of all, ND organises one-off GK branch office trainings (one for each ranche), through which GK branch managers and kendra managers receive (i) information about ND's activities, (ii) get trained about the details of the sanitation loan procedures, (iii) get an awareness training on the importance of hygiene and sanitation (by use of IEC material), (iv) obtain information about the government's SBM scheme, (v) receive a brief introduction to different available sanitation technologies and (vi) receive awareness creation handouts for distribution in the GP (see Figure 5 for an example of an IEC handout).

Figure 5: IEC handout



ND sets up separate one-off trainings but of similar content for GK clients at their weekly kendra meetings in the GP (one for each kendra). Similarly, they organise mason trainings (one for all masons of 2-3 GPs combined) where again similar material is covered, but in addition the masons receive an in-door technical training on the different sanitation technologies (using demonstration material such as a pan and a pin trap and display of demonstration videos). Through an interactive toilet costing session, ND also creates awareness about the actual costs of different sanitation technologies

¹¹Exceptions where loans can be disbursed in the village include emergency loans (Rs. 1,000), festival loans (Rs. 1,000) and medical loans (Rs. 2,000).

with the aim of convincing the audience that Rs. 15,000 should be more than sufficient to construct a good quality and sustainable toilet.

Figure 6: Street play



Furthermore, ND organises street plays of about 30-45 minutes (one per GP) performed by a group of 6 Maharathi actors. The objective of the play is to convey serious sanitation and hygiene messages in a fun and creative way. Street plays are organised in central, open public places, are announced through speakers and attended by various demographic groups in the community (men, women and children). Figure 6 shows the actors in action.

In addition to the various trainings and street plays, Navya Disha hangs awareness banners in public places (usually at the wall of the GP office - see Figure 7 for an example)¹² and conducts water quality tests of which the results are presented and discussed with the GP officials as a means of creating awareness about the hygiene issues in their communities.

Finally, for each GK branch in our study area (five branches in total) ND runs a half-day awareness workshop together for all GK members and GP officials who are living in GPs in which the branch provides sanitation loans. The purpose of those workshops is to re-emphasise the messages covered in the individual sessions described above by going through similar IEC materials and procedures (including another performance of the street play theatre group). In addition, ND invites external speakers to the workshop, such as government extension officers and local doctors, to come and share their views about the importance of sanitation. Figure 8 gives an idea of how a typical audience of such a branch level awareness meeting looks like.

¹²The more traditional approach is to do actual wall paintings which are possibly more sustainable. However, ND wanted to experiment with banners, inspired by practices of politicians during election periods in the study areas, who make extensive - and apparently highly effective use of banners. They perceive that banners have the following two main advantages over paintings: (i) they are significantly cheaper and hence one can have more than one such banner in a village, and (ii) they are movable and can hence, in subsequent visits, be placed elsewhere in the village, reaching potentially a larger audience than a wall painting.

Figure 7: Wall banner



Figure 8: Branch level awareness workshop



3.4 Intervention objectives

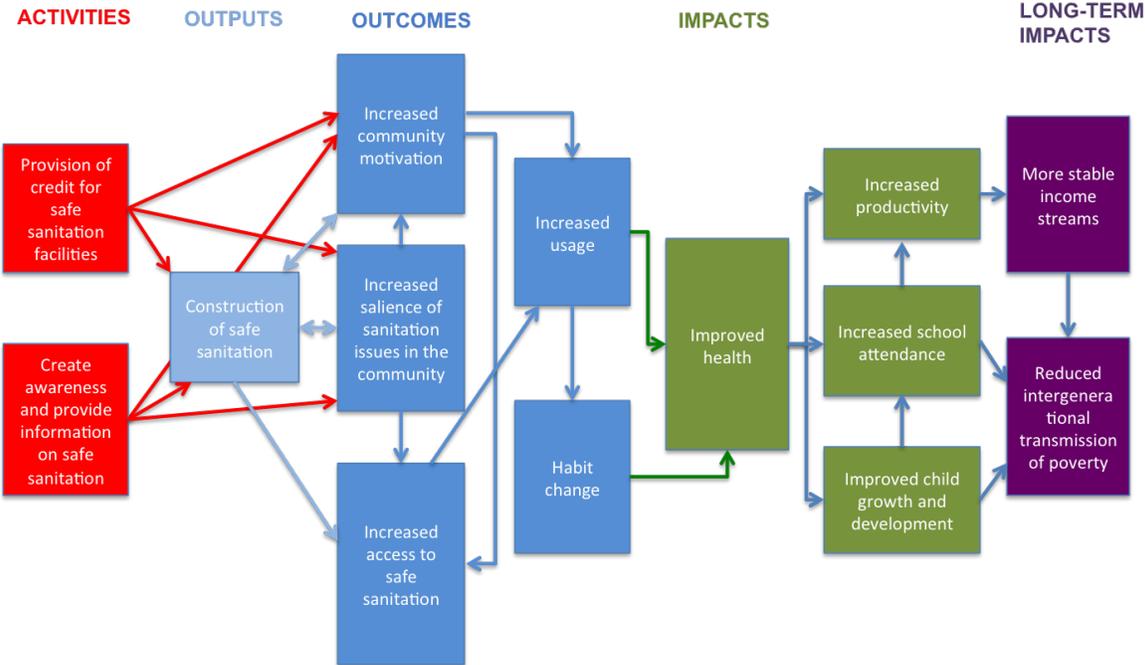
As described in Section 3.1, the GoI has recently strengthened its sanitation efforts with the aim of attaining a 100% Open Defecation Free India by 2019 (GoI, 2015). The government's large-scale sanitation programs are active throughout India, including in our study area. The purpose of the intervention subject to evaluation in this report is to complement rather than substitute government's efforts and should be considered in that light.

Just like the GoI, the key objectives of GK and ND are, in the first instance, to improve awareness about safe sanitation and to boost toilet construction and usage. Increased access to safe sanitation, in turn, is expected to positively impact on health through reduced parasitic/gut infections and reduced illness symptoms such as diar-

rhoea. Some studies suggest that health impacts can only be observed if the entire village becomes open defecation free (Pearson, 2013). However, so far this evidence is suggestive rather than conclusive. This will be examined more closely as part of this study.

In the longer term, better health of young infants is expected to yield improved child growth and development, increased school attendance and increased productivity. Ultimately, improved sanitation practice is expected to yield more stable income streams and to reduce intergenerational transmission of poverty. Figure 9 summarises the project’s short-term and longer-term objectives.

Figure 9: Project objectives



In spite of the government’s accelerating efforts to boost access to sanitation throughout India, we expect there to be a role for GK and ND in assisting the government in the achievement of their targets, for the following reasons:

- ND activities include the provision of information about the existence of the SBM program to villagers and what actions eligible households can take to get access to the government subsidies
- The government subsidies target specific types of households, primarily those living below the poverty line and other marginalised groups. Other demographic groups in need of financial support may reach out to sanitation loans to pay for toilet construction.
- Government subsidies are disbursed only *after* the household can demonstrate having built a toilet. Households in need of money up-front may not be able to construct a toilet without first acquiring a sanitation loan.

- The government subsidy currently stands at Rs. 12,000. Any additional expenses need to be covered by the households themselves.
- Moreover, an eligible household can receive a government subsidy only once. In case a household received a subsidy in the past, when the magnitude of the subsidies were much lower, it may still require a loan, for instance to upgrade its old toilet.
- Given the magnitude of the scale of India, and the speed at which the SBM program is being rolled out, this naturally poses a challenge to the efficiency of the program and the accuracy of the monitoring process. WSP (2011) points out the various challenges involved with an exponential increase in the number of applicants for NGP awards and makes a set of valuable recommendations. In general, we expect there to be some delays in the distribution of government sanitation subsidies to some eligible households. Although the government subsidies may be free whereas loans come at a cost in the form of interest payments, the subsidies may be relatively more difficult to come by whereas the loans are more readily available.
- WSP (2011) notes that experience has shown that there is an inverse relationship between government program scaling up and quality. ND awareness creating activities include an emphasis on the importance of high quality and safe toilets and hence may complement the government's program in terms of improving the quality of the toilets that are being built either through government subsidies or through sanitation loans.

4 Evaluation design

This evaluation will be based on a 'randomised control trial'. In what follows, we will provide details on the evaluation design, and highlight how it will enable us to evaluate the relative impact of each of the two project components described above. We will also consider potential sources of contamination of the design and discuss the actions we will take to mitigate those risks.

4.1 Randomised evaluation approach

While our implementing partner GK was already operating in our study area prior to the start of the project, neither the sanitation loans, nor the ND awareness creation activities were available to their beneficiaries nor to the communities their clients are living in. The interventions analysed in this study have been introduced during the course of this project in a manner that both facilitates implementation and allows us to rigorously assess the impact of the interventions. Specifically, the interventions have been randomly introduced in some of the 120 GPs and not in others.

Randomly providing sanitation services to some ('treatment group') and not to others ('control group') helps the impact evaluation process as it ensures that prior to the intervention treatment and control individuals are, on average, statistically the same in terms of observable and unobservable characteristics. In other words, randomisation removes 'selection bias' (i.e. pre-existing differences between the treatment and control groups, such as different levels of education, that might make one household more likely to follow hygiene practices than another). In theory, this should ensure that when we compare the outcomes of treatment and control groups at the end of the intervention, the only difference is due to exposure to intervention activities and not due to any unobserved differences between them. It allows one to obtain unbiased effects of the treatment on outcomes of interest.

While the need for randomisation is clear from a methodological point of view, one should also take its ethical implications into account. In particular, during the period of the experiment (approximately two years), some areas will be excluded from the GK and/or ND sanitation services although they would qualify to be covered in principle. Here it should be noted that implementing agencies would not be able to roll-out the programme across all areas of operation within the time of the evaluation. In practice, implementing agencies work in phases – covering one area, and then extending to another and so forth. We simply exploit the existing capacity constraints during the expansion phase of the programme to define the control groups.¹³

The interventions were randomised at the GP level (rather than at the household or individual level), for three key reasons: First, GPs play a key role in implementing

¹³While GK is theoretically able to roll-out a new loan product in all branches at the same time, experience shows that these are typically not marketed until loan officers were trained on the product which, in the case of sanitation loans, is done by ND staff. ND's resource capacity hence basically implies a phased-in approach. However, in the context of this study, the phase-in of loan products was made more explicit and was, in fact, held back more than originally envisioned, due to delays in evaluation activities.

government and state policies related to sanitation; second, because of intra-community spillovers and feedback effects associated with sanitation adoption; and finally because the intervention involves activities open to all GP members (e.g. street plays, engagement with the sarpanch, etc). Randomising at the individual or household level would therefore lead to the contamination of the control group, thereby leading to biased estimates.

Specifically, the 120 GPs in our study area (choice of which is described in Section 4.3.1 below and the list of GPs included is shown in Appendix ??) were more or less equally allocated to three evaluation groups:

1. Those that will receive both ND awareness activities and GK sanitation loans (ND + GK);
2. those that will only receive GK sanitation loans (GK); and
3. a control group that will receive no GK or ND services (Control).

The experimental design is summarised in Table 8. The reason why we do not have an equal distribution lies in the fact that randomisation was stratified at the level of the GK branch and by GP size (see Section 4.3.1).

Control	GK	GK + ND	Total
41	40	39	120

All 120 GPs (including those in the control group) will continue to receive a standard package of services from GK, which includes access to income generating micro-loans, and other loans offered by GK. This is important as it implies that our evaluation assesses the impact of providing sanitation activities (through GK and ND) *over and above* microfinance loans for other purposes. This first reason for this choice is operational: GK (as other MFIs providing sanitation loans) see sanitation loans to bare greater risks than non-income generating activities and hence want to know their clients better before they provide this new loan product (their recent policy change to only provide loans to clients that were with them for at least a year reflects this concern). The second reason is more important from our evaluation perspective and is that we would not want sanitation loans to crowd out other investment opportunities (which could also affect outcomes of interest).

This experimental design will allow us to assess 1) the impact of the provision of sanitation micro-loans, by comparing the outcomes of the ‘GK’ treatment arm with the outcomes of the ‘Control group’; and 2) the additional impact of ND activities by comparing the outcomes of the households in the ‘ND + GK’ treatment arm with the outcomes of the households living in the ‘GK’ treatment arm.

Finally, note that a large part of our analysis will be conducted at the individual or household level, and so all estimates will be identified through an intention-to-treat analysis, based on the initial community-level randomisation, regardless of whether or not an individual or household actually takes up the intervention (sanitation). Though

conservative, this will ensure that estimates are not subject to bias arising from selection, i.e. those that choose to build safe sanitation may be different from other households in the community. Moreover, all households in treated GPs will receive some aspect of the intervention (information for all, and credit for some), making the intention-to-treat analysis best suited for our purposes.

We conducted a power analysis to determine our sample size and understand the minimum detectable effects we can achieve. For this power analysis, which was done at the time of writing the proposal, we had to make a number of assumptions, some of which we are now able to refine and update with baseline data on our study population at hand. The key outcome discussed in the power analysis in the submitted proposal was the number of health insurance claims. We will not be able to update this analysis until the health insurance component is implemented. We will therefore update only the power analysis related to sanitation uptake. Details are provided in Appendix A. To summarize, the actual baseline mean is higher than was anticipated. The raw ICC is 0.143, which is reduced to 0.105 by taking household characteristics into account. Given this mean and ICC and keeping all other assumptions constant, we will be able to detect increases of about 10.9-12.3 percentage points. This assumption compares well with our preferred scenario discussed in the proposal and we conclude that for the time being, we anticipate that our design will allow us to detect expected minimum impacts.

4.2 Potential risk factors

There are four potential sources of contamination to the evaluation design that are of concern:

1. There could be **non-compliance** with the randomisation by field and/or branch level staff of the implementing institutions. To prevent this, the research team has visited each of the GK branches and the ND office to give an in-depth training to all stakeholders involved about the research design and the importance of not doing any sanitation related activities in the control area. Post this training, branches prepared posters listing all communities in which no loans should be provided, for everyone to see and be reminded of when coming to the office. Picture 10 shows an example. Moreover, GK head office has provided additional monitoring of branch staff and has disabled sanitation loan applications from control areas in their central management information system. That means that it is technically not possible for anyone in the control GPs to receive any sanitation loans. This will also be monitored using GK loan monitoring data (see Section 4.4.6).
2. **Other programs** (particularly by the GoI) could intensify their efforts at improving sanitation infrastructure across study areas and thereby increase sanitation density in our control GPs. To some extent, this is not a problem for our evaluation design, as long as the government's efforts are similar across both treatment and control GPs, and as long as the government does not manage to get toilets

Figure 10: Reminder of control GPs in GK branch office

Sanitation lone not paerndrogy leandry

hush

Branch	Gram Panchayat	Kendra ID	Kendra name	Remark
Degloor	BORGAON	0173-000624106	Bergao	Not permission
Degloor	DHOSNI	0173-000743503	Dhosani	Not permission
Degloor	HANEGAON	0173-000679040	Hanegaon	Not permission
Degloor		0173-000678697	Hanegaon-1	Not permission
Degloor	HOTTAL	0173-001022169	Hostal 01	Not permission
Degloor	MALEGAON (M)	0173-000948824	Malegaon	Not permission
Degloor		0173-000948922	Malegaon 01	Not permission
Degloor	MARKHEL	0173-000955808	Markhel-3	Not permission
Degloor		0173-000898141	Markhal	Not permission
Degloor		0173-000908440	Markhal-2	Not permission
Degloor	MARTOLI	0173-000982574	Martoli	Not permission
Degloor		0173-001419203	Martoli B	Not permission
Degloor		0173-001502664	Marattoli 03	Not permission
Degloor	NARANGAL BK	0173-000638324	Narangal 2	Not permission
Degloor		0173-000642970	Narangal 1	Not permission
Degloor	ZARI	0173-000729470	Zari	Not permission

to all of the households in the control group. As described in Section 3.4, the objective of the intervention is to complement government’s efforts rather than substituting for it. In our context - and in fact in most other evaluations of development impact - control communities represent cases of “business-as-usual” government activity, rather than “doing nothing” (Ravallion 2008). The impact we are interested in relates to the contributions of GK and ND’s activities *in addition* to the government’s program. However, if the government scales up its efforts to the extent that most households in the control group acquire a toilet, the potential for impact from ND and GK is reduced. Even in that case, however, we may still expect there to be an impact on toilet quality rather than toilet ownership. Section 3.4 above lists various channels through which we think GK and ND can make a difference despite the accelerated efforts that are being made by the government. Moreover, we integrated in our design two rapid assessments, approximately 4 and 8 months after the baseline survey. These rapid assessments will collect information on sanitation uptake in both control and treatment communities, providing us a picture of uptake in both study groups and hence allowing us to monitor this potential risk.

3. **Contamination** in our design might arise from sanitation information spilling over to neighbouring GPs, which may be control GPs. To be able to account for this, we will collect information on possible interactions between neighbouring GPs, e.g. common markets, and common branch level GK meetings, characteristics of GK branch officers, and distances between GPs. However, in general this risk is assessed as low given significant distances between GPs.
4. In terms of **attrition**, we anticipate low attrition among GK clients given their long-standing relationship with GK. To reduce attrition particularly among non-clients for all surveyed households, we will collect mobile phone numbers of household members and contact details of relatives and friends who are likely to know of their whereabouts if they move. As will be discussed below, our baseline find-

ings show that only around 4% of the households in our sample reported to have migrated in the year prior to the survey. This is consistent with other studies in areas near our study area, which have encountered relatively low attrition rates, driven predominantly by migration (at 3-4%, see for example Mahal et al 2012). Moreover, our baseline results show that more than 97% of all households report it to be unlikely that they will have moved in a year from now.

5. Finally, we note that there are currently some undefined parameters around the implementation of the (primary) **health insurance component**, which is crucial for the second part of the study. The coming months will focus on narrowing down the necessary steps and putting things for this component in place.

In addition to this list of potential contamination concerns, the change in GK’s policy in terms of eligibility for loans (allowing only clients who stayed with GK for a year) might reduce our power to detect impacts. We purposely sampled GK client households (details are provided in the following section) but did not take into account how long they have been with GK. Ideally, we would have liked to sample only amongst clients who were for at least a year with GK. However, since the policy change happened only when our baseline data collection was already completed, we were not able to take this into account. The extent to which this policy change affects the study sample is unclear at the moment. Based on loans provided in study areas in February and March of this year, we know that roughly 20-30% of sampled clients are with GK for less than one year. However, for one, this data does not cover all sampled clients and second, the policy is not yet followed strictly (i.e. some of the new clients have already received a sanitation loan). We will use GK’s loan data, which we expect to receive roughly every two months, to monitor this issue closely.

Lastly, there is a possibility that some of the kendras that are part of our study group close down between baseline and endline, which would imply a reduction in our effective GK sample size. This risk will be followed up very closely by examining the GK monitoring data.

4.3 Sample selection strategy

4.3.1 Selection and randomisation of study GPs

The study took off by determining the set of 120 GPs to consider as part of the evaluation and then to randomly allocate each of these to the three different evaluation arms. This Section provides details on this selection and randomisation process.

The first step was to identify the set of feasible study GPs. A feasible study GP was defined as one where (i) GK had operations (i.e. active GK kendra groups) and (ii) where neither sanitation nor health insurance activities had been undertaken by GK in the past. Through interactions with GK, we identified areas where they had neither sanitation nor health insurance activities ongoing (and did not have them in the past), but where they were planned (and/or considered feasible) in the near future. During the discussion six blocks (corresponding to branches) within three districts in Maharashtra, namely Solapur (the block of Akkalkot), Nanded (Degloor, Naigaon) and

Latur (Ahmedpur, Nilanga, Udgir) were identified. Below it will become clear why later it was decided to drop district Solapur from the study.

GK does not keep information on the gram panchayat (our unit of randomisation) that the kendra members live in (which corresponds to where the Kendra is registered). This information was however crucial in identifying feasible study GPs (since not all GPs in the operating area of the branches are necessarily covered by GK). The next step therefore was to get a list of GK kendra groups that GK had in these branches. With the list of kendras at hand, kendras were matched to the GPs they were located in. During the process, kendras located in urban areas were excluded (as they would by definition not be in a GP and hence not be eligible to become part of the study). Approximately 80% of all kendras were matched to GPs based on their names using 2001 census information, and Google (i.e. searching for the Kendra name to see if it was associated with a village with a different name to the GP name) where possible. The remaining kendras needed to be matched to their GPs with the input from GK branch staff.

The first list was compiled in November 2012. Due to project delays the list had to be updated at multiple times before the baseline survey to account for closing of kendras, opening of new kendras and also some matching mistakes were detected (even after the listing exercise had taken place, which has some implications, as will be explained in Section 4.3.2). Once the list was available, it turned out that there was more than a sufficient number of GPs in just five of the six branches. After consultation with all stakeholders in this study, it was decided to drop one branch. This was done for two main reasons: (i) to ease logistics of data collection and with that reduce data collection costs and (ii) to allow GK to start with sanitation activities in the dropped branch¹⁴. The selected branch to drop was Akkalkot in Solapur. The final list of study branches was the following: Degloor, Naigaon, Ahmedpur, Nilanga, and Udgir.

At that stage, the resulting list of feasible GPs included 142 GPs. Since the study design only required 120 GPs, this allowed us to additionally drop the largest 5% of GPs, to reduce the list to 133 GPs. This was useful for two main reasons: (i) it allowed us to reduce the variation of GP size which was quite substantial, and (ii) it again facilitated data collection logistics, reducing costs. Finally, we randomly selected 120 GPs out of the list of 133 GPs, conditional on retaining a number of GPs per branch that was divisible by three (to ensure randomisation across three treatment arms in the next step). We refer to Table 117 in Appendix D for a list of the final set of study GPs, the administrative blocks in which each of these are located and the GK branch by which there are covered.

The second step was the randomisation of the 120 GPs to one of three evaluation arms (see Section 4.1): Control, GK, GK + ND. Randomisation was stratified by branch and by size of the GP (size in terms of number of households). The reason for stratifying by GP size lies in the data collection process which will be described in Section 4.3.2 below. That Section will also explain which GPs were classified as ‘small’

¹⁴The significant project delays implied that GK was only able to start the interventions in study areas much later than initially envisioned. Dropping one branch was therefore a very welcomed decision by GK.

and which GPs were classified as ‘large’.¹⁵

Table 9 summarises the results of the randomisation to the three evaluation arms, stratified by GK branch office and GP size. Table 117 in Appendix D shows for each GP the evaluation arm that they have been allocated to and provides baseline information on the number of GK kendras and the number of GK clients living in each of the study GPs.

Table 9: Random allocation of study GPs to evaluation arms, by branch and size

		Control	GK	GK + ND	Total
Ahmadpur	Large	2	2	1	5
	Small	3	2	2	7
Degloor	Large	4	4	5	13
	Small	5	4	5	14
Naigaon	Large	5	5	5	15
	Small	5	5	5	15
Nilanga	Large	5	5	4	14
	Small	4	4	4	12
Udgir	Large	5	5	5	15
	Small	3	4	3	10
Total		41	40	39	120

4.3.2 GP segmentation and listing

An important component of the study design was the implementation of an extensive listing exercise, the purpose of which was threefold:

- To serve as a sampling frame for the selection of respondents for the baseline survey (see Section 4.3.4) - since no up-to-date household level information was otherwise available. It was important to have a comprehensive and up-to-date sampling frame to ensure we would have a random sample of the whole community - including households that might have moved in recently, and covering all

¹⁵The randomisation process, stratified by branch and by GP size, worked as follows: First the GPs were stratified into 10 strata: five branches and within each branch one stratum for large GPs (more than 480 households) and one stratum for small GPs (less or equal to 480 households). Randomisation was done in 2 steps (after drawing a random number for each GP from a uniform distribution): Step 1 sorted the data by branch and magnitude of the random number from the highest value to the lowest. Based on this ordering, the first N GPs, where N is the largest number of GPs in the stratum that is fully divisible by 3 (e.g. a stratum could have 14 GPs, in which case N would be 12), were assigned in sequence to one of the three evaluation arms. For strata where the number of GPs was not divisible by 3, the ‘remainder’ GPs were randomly allocated in a second step. The idea here was to treat the unallocated units as an additional stratum, and assign them randomly to the evaluation group, but ensuring that: (i) no treatment or control group got allocated more than 1 of these units within a stratum, and (ii) we randomly chose which treatment group got the extra units. The process is based on a blog post by McKenzie and Bruhn: <http://blogs.worldbank.org/impacetevaluations/tools-of-the-trade-doing-stratified-randomisation-with-uneven-numbers-in-some-strata>

marginalised groups (both examples of groups that could be less likely to invest into sanitation);

- To provide us with some important GP level insights of our main outcome variable particularly (sanitation ownership) at baseline;
- To collect GPS data that allow us, among other: (i) to check that all households interviewed are indeed situated within the selected GPs, (ii) to make it easier to identify panel households during the follow-up survey, (iii) to calculate distances between neighbourhoods, (iv) to assess for potential spillovers across treatment arms, (v) to construct social networks measures to assess the role of neighbours in the sanitation uptake, etc.

Based on the assumption that there would be on average 400 households per GP (as per the 2001 census data and taking population growth into account), the original idea was to administer a short questionnaire of about five minutes and to collect GPS coordinates of every household in all of the study GPs. However, after we had identified the list of study GPs and updated the population estimates using the latest 2011 census data, it turned out that the actual number of households in the study GPs was much higher, on average, around 590, i.e. exceeding the assumption by a considerable amount and with that exceeding the available budget. Therefore, the following decisions were made:

- Complete listing was done of GPs of size 480 households or smaller;
- In cases where GPs consisted of more than 480 households that were at the same time spread over more than one village, one village was selected at random for complete listing and the other one(s) were excluded from listing;
- GPs consisting of only one village of size greater than 480 households were segmented into segments of 120 households and 4 segments were subsequently randomly selected for listing purpose, conditional on having at least one segment with GK clients (see Appendix B for more details on how segmentation was done in practice);
- The randomisation of GPs to treatment arms took this distinction between ‘small’ and ‘large’ GPs into account (both for fully and only partially listed GPs, see discussion in Section 4.3.1 for details). Specifically, GPs (or sampled village within the GP in case there was more than one village) that had more than 480 households were classified as being ‘large’ in the stratified randomisation process and those with less or equal to 480 households were denoted as ‘small’.

Table 117 in the Appendix shows for each GP the total number of villages it contains, the name of the village within the GP that was sampled for the study, the number of households living in that sampled village (i.e. the indicator on the basis of which GPs were stratified in the randomisation process) and an indicator for whether or not the GP was segmented during listing.

After completion of the listing exercise, it was discovered that among the 120 GPs that had originally been sampled for the study - and therefore had been listed - there

were 13 GPs that did not have any GK clients residing in them. This was caused by errors made in the complicated process of matching GK kendras to GPs and the time the exercise took (during which some kendras became inactive), as described in Section 4.3.1. These 13 GPs were subsequently dropped from the study sample and replaced by 13 new GPs to form the final set of 120 study GPs that were visited for baseline survey and that will be re-visited in the next survey rounds. Note, however, that this means that for 13 GPs in our sample, we do not have any listing data available. Table 117 in the Appendix shows for which GPs we have listing data available and for which these data are missing.

Table 10: Sample composition listing survey (107 GPs)

	Control	GK	GK + ND	Total
# of HHs visited during listing (total)	14,277	12,830	13,960	41,067
# of HHs visited during listing (GP average)	397	377	377	384
# of HHs available for interview (total)	13,482	12,337	13,049	38,868
# of HHs available for interview (GP average)	375	363	353	363
% of households available for interview (total)	94%	96%	94%	95%

A total of 41,067 households (living in residential dwellings) were visited during listing in those 107 GPs where GK has active operations. Out of these, a total of 38,868 households (95%) were available for the short listing interview. The remaining ones were unavailable, either because the dwelling was locked (4%) or the respondent refused to be interviewed or for any other reason (1%). Table 10 breaks down the number of available and unavailable households by evaluation treatment arm.

Section 4.4.1 provides more details on the information that was collected as part of the listing survey.

4.3.3 Matching listing dataset to GK clients database

As will be described in the next section, we do not only stratify at the GP-level randomisation, but we also adopt a stratified sampling approach for respondents. Study households were stratified by (i) whether a GK client lives in the household and (ii) whether a child under the age of two years lives in the household. In order to stratify the population on the basis of GK membership, households of GK clients first had to be identified in the listing dataset. To avoid the survey contaminating the study design, the listing survey did not explicitly ask the respondent whether or not any household member was a GK client. Instead, the identification of GK households was to be achieved ex-post listing, by matching the names of the (only female) clients in the GK clients database to the names of household members provided during the listing survey. Three challenges occurred related to this matching exercise:

- In large GPs not all households were listed (see discussion in Section 4.3.2), which meant that GK clients living in areas not covered by listing could not be matched and we do not know the percentage that was not listed;

- In the interest of time and costs, the listing survey did not collect information on the names of all of the household members. It only asked about the name of the household head and the name of the spouse of the household head or, in case there was no spouse, the name of the eldest female member. This meant that in case the GK client was not either the household head or the head’s spouse/eldest female member in the household, the GK client could not be identified in the listing dataset (even after using client/spouse’s middle names and last names);
- As explained in the previous Section, for 13 out of 120 study GPs we do not have listing data available. GK clients living in those areas could not be identified ahead of the start of the baseline survey.

As a result, only half of the GK clients in the study area could be reliably matched to the listing dataset and the original sampling strategy had to be adjusted. Appendix C discusses the complementary sampling strategies that were adopted in order to circumvent these problems.

4.3.4 Sample selection and sample size

Household survey

By design, the evaluation included a baseline household survey of a sample of 30 households in each of the 120 study GPs (3600 households in total), stratified by (i) whether a GK client lives in the household and (ii) whether a child under the age of two years old lives in the household. Originally, based on the information that was available at the study design phase, the distribution of the sample was planned to be as depicted in Figure 11, i.e. half of the sample being GK client households, the other half non-client households and within these interview 8 households with children under the age of 2 years and 7 without.

Figure 11: Original sampling strategy per GP



However, as it turned out, in some GPs it was not possible to reach the target number of available GK households in the stratum of kids less than two years old (i.e. eight). Whereas the sampling strategy for the non-GK group remained the same, the sampling strategy for the GK group was revised as follows: In case there were not sufficient (i.e. 8) client households with kids less than two years of age, households with kids less than six years were sampled instead. If still not enough client households falling in this category could be identified, then households without kids were over-sampled to obtain a total of 30 sampled households in each GP. To compensate for the lower

number of sampled clients with kids in some GPs, in neighbouring GPs a respectively higher number of GK clients with kids was sampled, although always ensuring that a minimum of five non-clients with kids was sampled in each of the study GPs.

Further, a total of 14 completed household surveys got lost as a result of a technical failure with the electronic survey system (see Section 4.5).

The overall result is that in some GPs there is a shortfall relative to the target sample size of 30 surveys, whereas in others there is a surplus. Table 118 in Appendix D shows the shortfall/surplus by GP. Given that the survey teams had surveyed a higher number of households than was required, the net result is a shortfall of 5 households in total, which is insignificant relative to the target sample size of 3600, i.e. 0.001 %. For none of the three treatment arms the shortfall/surplus exceeds 1%.

This yields a total sample size of 3,595 in the household survey dataset. Table 11 summarises the sample size by stratum (by client/non-client and by kids/no kids) and by treatment arm.

Table 11: Sample composition baseline household survey

		Control	GK	GK + ND	Total
Non-GK clients	Kids < 2 years old	330 (53%)	318 (53%)	314 (54%)	962 (53%)
	Kids 2-5 years old	48 (8%)	34 (6%)	57 (10%)	139 (8%)
	No kids < 6 years old	239 (39%)	248 (41%)	213 (36%)	700 (39%)
	Total	617	600	584	1801
GK clients	Kids < 2 years old	230 (37%)	182 (31%)	203 (35%)	615 (34%)
	Kids 2-5 years old	120 (19%)	120 (20%)	97 (17%)	337 (19%)
	No kids < 6 years old	271 (44%)	285 (48%)	286 (49%)	842 (47%)
	Total	621	587	586	1794
Total	1238	1187	1170	3595	

The distribution of the total number of surveys across the different treatment arms is as expected given the number of GPs included in each of the treatment groups. The control group sample has slightly more households than the ‘GK’ and ‘GK + ND’ treatment groups, given that it covers 41 GPs, whereas the other two groups include 40 and 39 GPs, respectively. As can be seen from the numbers in brackets in the table, which show the share of each client/non-client stratum within its treatment category, the distribution across treatment arms remains similar when breaking the composition down by stratum.

As expected, we observe differences (albeit small) in sample composition between the GK and non-GK sub-groups. In the non-GK client group, about 53 % of the sample of the non-GK group contains households with kids less than two years old and 47% of the sample has households without kids less than two years old. This is very much in line with the sampling strategy summarised in Figure 11 which targeted for each

GP eight households with kids less than two years old (i.e. 53%) and seven households without kids of two years old (i.e. 47%).

The sample of GK households, on the other hand, contains a relatively smaller share of households with kids less than two years old, i.e. on average 34% of the total. This is consistent with the revised sampling strategy for the GK group, discussed above, which allowed for an increase in the threshold of the age of kids to six years of age in case not a sufficient number of GK households with kids less than two years old were available. Note that the share of households without kids (here defined as kids less than 6 years old) is exactly the same as in the non-GK group, i.e. 47%.

Individual man and woman survey

In addition to the household survey, an individual man and woman survey was conducted for most of the households. Figure 12 describes the conditions under which the administration of these individual surveys was required and who would be the respondents in this case. In sum, an individual man/woman survey was required if there was either a father/mother of a child less than 6 years old available in the household or, if not applicable, a male/female household head/spouse of head should be interviewed. Section 4.4.5 and Section 4.4.4 describe the purpose and contents of these questionnaires in more detail.

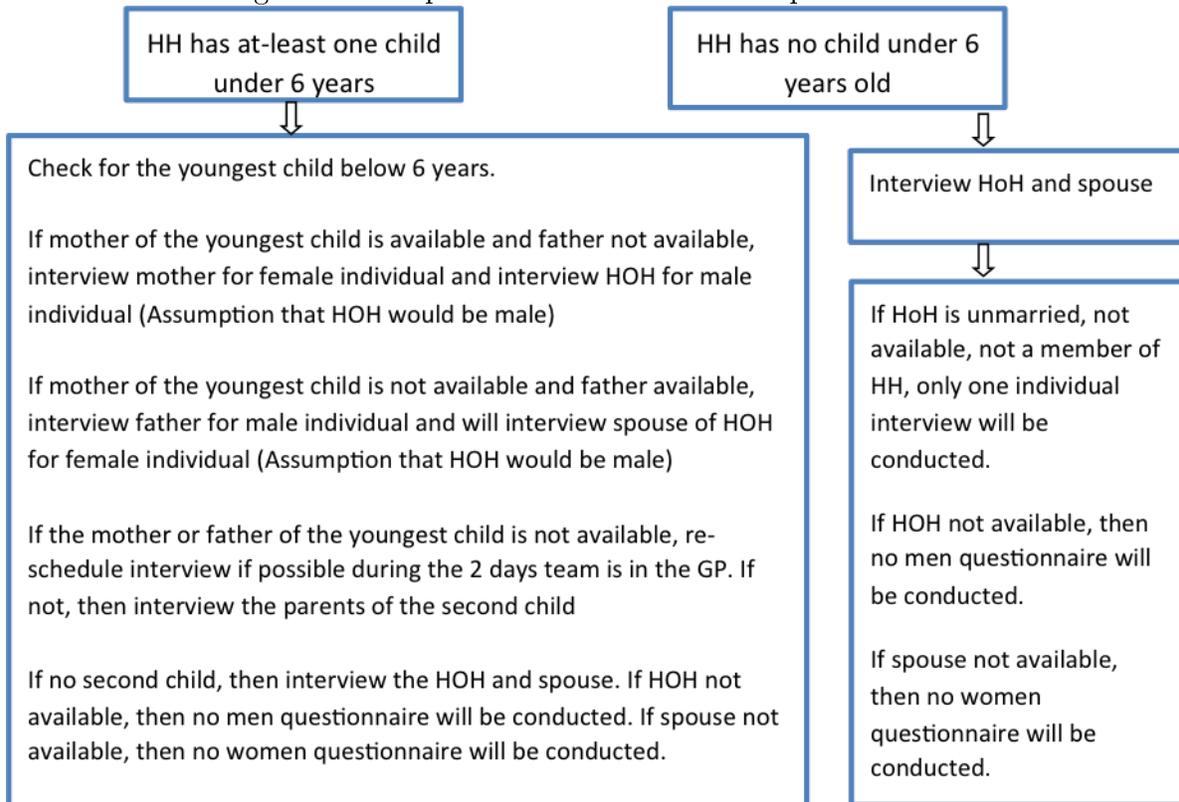
Table 12 summarises the sample composition of the individual man survey dataset. It should be noted that for 13 records in the man and woman datasets, there is no parent data record available in the household survey dataset, as a result of the fact that 14 household surveys went missing because of a technical failure (see Section 4.5). These 13 records are not included in this summary table.

In total, 96% of the households that were covered by the baseline household survey received an individual man questionnaire. This means that the response rate among male respondents was very high. 1% of the households in the sample received a man survey even though this was not required according to the guidelines described in Figure 12. On the other hand, for 1% of the households for which a male survey was required there is no individual man survey available, because of unavailability of the respondent, or because of loss of surveys as result of the technical problems encountered (see Section 4.5).

The total man survey sample size is proportionately similar across treatment arms. The difference in sample size composition between GK and non-GK households is consistent with the differences observed in the composition of the household survey sample size which was discussed before. That is, the man survey dataset in the GK group contains proportionally more respondents from households with kids less than six years old relative to the non-GK group.

Table 13 presents similar findings for the composition of the individual woman survey dataset (excluding 14 observations for which no parent household record is available, see discussion above). In total, 99% of all households in the baseline household survey got administered an individual woman survey. Similarly as for the man survey, 1 % of the households in the baseline survey completed a woman survey despite this survey not being required according to the guidelines. For another 1% of households there is

Figure 12: Respondent woman and man questionnaire



no woman survey data available even though this was required.

Anthropometrics

Finally, the baseline survey took anthropometrics (height and weight) from all children up to the age of 2. Table 14 summarises the number and proportion of households for which anthropometrics data for both height and weight is available, by GK clients/non-GK clients. Complete anthropometrics were taken in 95% of all households where this was required, i.e. those with kids less than to years old. For the remaining 5% of households child measurements could not be captured/are not available because of the child or his/her mother being out of station (4%) or because of other reasons (1%), such as refusals or data getting lost because of the technical issues faced in the field (see Section 4.5).

4.3.5 Randomisation of branch order

Finally, we randomised the order in which data collection in each branch was taken place. While not necessarily high risk, this was to avoid that data collection would be

Table 12: Sample composition individual man survey

		Control	GK	GK + ND	Total	% HHs with men survey
Non-GK clients	Kids < 6	371 (62%)	352 (60%)	369 (65%)	1,092 (63%)	99%
	No kids < 6	223 (38%)	233 (40%)	199 (35%)	655 (37%)	94%
	Total	594	585	568	1,747	97%
GK clients	Kids < 6	337 (58%)	298 (53%)	294 (52%)	929 (54%)	98%
	No kids < 6	245 (42%)	263 (47%)	277 (48%)	785 (46%)	93%
	Total	581	560	570	1,711	95%
Total		1,176	1,146	1,139	3,461	96%
% HHs with man survey		95%	96%	97%	96%	

Notes: The numbers in this table exclude 13 observations for which no parent household record is available.

Table 13: Sample composition individual woman survey

		Control	GK	GK + ND	Total	% HHs with woman survey
Non-GK clients	Kids < 6	377 (62%)	350 (59%)	364 (63%)	1091 (61%)	99%
	No kids < 6	236 (38%)	244 (41%)	210 (37%)	690 (39%)	99%
	Total	613	594	574	1781	99%
GK clients	Kids < 6	349 (56%)	300 (51%)	298 (51%)	947 (53%)	99%
	No kids < 6	270 (44%)	285 (49%)	284 (49%)	839 (47%)	100%
	Total	619	585	582	1786	99%
Total		1232	1179	1156	3567	99%
% HHs with woman survey		99%	99%	98%	99%	

Notes: The numbers in this table exclude 14 observations for which no parent household record is available.

Table 14: Sample composition anthropometrics

	Control	GK	GK + ND	Total	% HHs (total sample)	% sample (with kids)
Non-GK clients	319	305	282	906	50%	94%
GK clients	225	168	191	584	33%	95%
Total	544	473	473	1490	41%	95%
% HHs (total sample)	44%	40%	40%	41%		
% (HHs sample with kids)	97%	95%	92%	95%		

done in an order in which for example other programs could be rolled out. One could think of government programs that following a specific geographic or alphabetical ordering. The outcome is the following sequence in which baseline survey was conducted:

1. Naigaon
2. Nilanga
3. Udgir
4. Degloor
5. Ahamedpur

4.4 Instruments for data collection

The baseline survey covered five different types of surveys (listing questionnaire, community questionnaire, household questionnaire, women questionnaire - including anthropometrics - and man questionnaire), which were organised and implemented by an external survey firm. The instruments were developed by the research team and refined in conjunction with the survey firm. They were piloted in two stages: First, the paper-versions were tested in the field. This took place in two villages in rural Maharashtra, outside our study area¹⁶ in the period May 10-13 2014.¹⁷ The second piloting focused on CAPI pre-testing and took place in the period 22-25 September 2014. This testing was done by the survey firm exclusively. We refer to the Implementation Report, for details about the overall field work and quality control procedures, and to the various field training manuals for more information about the instructions provided to the enumerators.

Except for the community questionnaire, all surveys were conducted electronically using PDAs (for listing) or tablets (household, man and woman surveys). The listing survey started on 14 September 2014 and was completed within a month, on 12 October 2015. The baseline household surveys (community, household, woman and man questionnaires) took place between 24 November 2014 and 26 January 2015.

4.4.1 Listing questionnaire

The listing exercise was carried out by an external survey firm using PDAs with in-built GPS devices. The questionnaire was designed to be brief and cover only names of the household head and spouse, and the eldest female in case the household head was not married. These information were used to match GK clients to the listing data. It further collected information about children under the age of 2 years for stratification purposes and finally, it asked about the household's caste as well as some basic information on the households sanitation situation. A more detailed discussion of the purpose and procedures of this listing exercise is provided above in Section 4.3.2.

¹⁶Specifically, the PAPI pre-testing took place in two villages (Shindvane and Degaon) in Pune district.

¹⁷Participants in the training included researchers from the survey firm, one male and one female field supervisor, a representative from WSP India as well as from IFS London.

4.4.2 Community (village) questionnaire

The main purpose of this instrument is to collect information on factors that are expected to facilitate or potentially constrain the uptake and success of the interventions we study. Accounting for characteristics of the environment in which the intervention is implemented will be crucial to assess its success or understand failures. To that end, the questionnaire covers information on the village's population (size, caste structure, main activities in general and main farming activities in particular, as well as wages for the main activities), transportation links, information on access and distances to specific infrastructure (public services such as schools, police, health services etc, artisans and masons, and shops), information on the political economy, community activities, sources of water and sanitation and waste disposal and finally shocks experienced and prices of a list of goods.

This questionnaire focused on one village within the GP. As discussed above, in most cases this one village would be equivalent to the GP. In remaining ones, the village would be picked at random. The information in this questionnaire were collected from relevant, well-informed members of the community. The information could be collected in a group discussion, or on a one-on-one basis with the relevant informants. Community members that were approached include the sarpanch, gram sevak, teacher, talati, anganwadi worker, opinion leader, SHG federation leader, but should also include villagers that do not have such special posts in the community, and may be from poorer parts of the community.

Due to some logistical problems, one village questionnaire¹⁸ (of the GP Dhosni) was not yet completed at the time of writing up this baseline report. Completion of this survey is scheduled to take place in the next couple of weeks. This report will present the community findings for 119 villages in 119 GPs.

4.4.3 Household questionnaire

The household questionnaire was designed to provide us with information on the baseline levels of the outcomes of interest for the study: This is very important so as to be able to assess whether the intervention was effective and also how large any effect is. Moreover, the baseline level could in some cases be used as a control variable to increase power of the experiment. The outcomes of interest are as follows:

Primary Outcomes: Sanitation uptake, uptake of safe sanitation, usage of safe sanitation, health insurance claims;

Secondary Outcomes: Health (diarrhea, child nutritional status), awareness about sanitation, changes in perceptions of costs and benefits of safe sanitation (for the Navya Disha intervention); Uptake of credit, awareness about health insurance, uptake of health insurance, household income and consumption (affected by credit);

Tertiary outcomes: Productivity, schooling, among other.

¹⁸Dhosni village in Dhosni GP, Deglur block and Nanded district. This village was allocated to the control group.

It should be noted that the secondary and tertiary outcomes allow us to investigate how the intervention works (i.e. does it follow the channels outlined in the theory of change? Which of these channels is most important?), and also impacts on broader socio-economic outcomes (e.g. productivity and schooling), which are hypothesised to be affected by poor sanitation.

The second major purpose of the household questionnaire is to collect characteristics of the household that would provide a good description of the study population, poverty levels and wealth. We will also want to investigate heterogeneous impacts of the interventions by some of these variables (e.g. is it only the wealthiest households that take up the loans/build safe sanitation, etc). Finally, some of these variables will become covariates in our analysis and help improve the power of the study.

The household questionnaire was hence the most extensive module, covering socio-economic characteristics of the household (assets, income, savings, credit, consumption expenditures), household member information (age, gender, education, etc) and a detailed section on sanitation and hygiene infrastructure, practices and beliefs.

Where possible, the head of the household was interviewed, however contributions from other knowledgeable household members were also allowed. At the first instance though, if the household head was not available, the interview was re-scheduled to a convenient time for the household head. If he or she was not available in the next 2 days, another knowledgeable adult household member was interviewed. The interview took on average 66 minutes to complete.

4.4.4 Women questionnaire

There are four key reasons for the choice of having a questionnaire administered to a woman in households of our study population, partly in line with reasons described above for the household questionnaire:

1. To collect information on individual sanitation behaviour, reported by the woman herself, the former one being one of the key outcome measures in the first stage of this experiment;
2. To collect information on individual sanitation preferences and beliefs to understand women's perceptions of the costs and benefits of sanitation, and also to identify what they value about it. This set of variables will allow us to identify the components of the interventions that households are likely to value the most, and thereby are likely to be most effective;
3. To collect information on child health, child care practices and child nutrition for children aged < 6 years;
4. To collect information on women's status in the household.

The interview was conducted by a female interviewer with a female respondent of the household, selected as described in Section 4.3.4. The interview took on average 32 minutes.

4.4.5 Men questionnaire

The main purpose for the men questionnaire is in line with points 1 and 2 presented for women above. On the second, sanitation preferences and perception, this is driven by anecdotal evidence that suggests that men and women value sanitation differently, and that this explains the slow take-up rates of sanitation.

The interview conducted by a male enumerator with the men (selected as described in Section 4.3.4), took on average 19 minutes.

4.4.6 Monitoring data and rapid assessments

Finally, emphasis will be placed on monitoring on-going progress and processes of the sanitation intervention. This will be done through two means: First, we will receive information, on an on-going basis (roughly bi-monthly), from GK on the number of sanitation loans they have provided in study areas. This will inform us about the sanitation uptake among GK clients and will allow us to monitor some of the risks described in Section 4.2.

We will further conduct two rapid assessments (RAs), roughly 4 and 8 months after implementation start. These rapid assessments will be done in a similar manner as done by the GoI when checking OD(F) status of villages. Specifically, for the first RA we will select half of the 120 study GPs and will within each of these, interview about 30 respondents. These respondents will be different to the baseline respondents, to avoid survey fatigue. The second RA will then concentrate on the remaining 60 GPs. Further details, including the questionnaire to be fielded, are in development at the time of writing this report.

4.5 Problems in data collection

The previous sections have already touched upon the various problems encountered during data collection. This section sums up the main challenges that were faced and the subsequent decisions that were made to move forward.

1. After selection of the study GPs (see Section 4.3.1), it turned out that the number of households living in these GPs is significantly higher than what was anticipated at the project proposal stage (based on the latest census data). Instead of listing all households in all of the study GPs, for large GPs (more than 480 households) it was decided to list only a random sub-set of households. If the GP had more than one village, one village was randomly sampled. If the sampled village had more than 480 households, the village was segmented for listing according to a procedure described in Section 4.3.2.
2. After completion of the listing exercise, it was discovered that among the 120 GPs that had originally been sampled for the study - and therefore had been listed - there were 13 GPs that did not have any GK clients residing in them. This was caused by errors made in the complicated process of matching GK kendras to GPs and the time the exercise took (during which some kendras became inactive), as

described in Section 4.3.1. These 13 GPs were subsequently dropped from the study sample and replaced by 13 new GPs to form the final set of 120 study GPs that were visited for baseline survey and that will be re-visited in the next survey rounds. Note, however, that this means that for 13 GPs in our sample, we do not have any listing data available. Table 117 in the Appendix shows for which GPs we have listing data available and for which these data are missing. This had implications for the sampling strategy, which are discussed in Section 4.3.1.

3. Reviewing the listing data, IFS/WSP observed that the GPS readings collected during household listing were missing for 26.7% of the households and were duplicated for 20.2% of the households. This means that for almost half of the households in the listing dataset we currently do not have accurate GPS coordinates available. This issue has been discussed in length with the survey firm that was in charge of collecting these data, who blame poor satellite connectivity as the main cause for this. The survey firm has agreed to go back to the field to re-capture GPS readings for those households for which these data are missing. This is scheduled to take place in the next few weeks.
4. There was a long delay in the process of obtaining an official support letter from Yasahada for the study. This letter was needed by the survey field team as a means to introduce themselves to the Sarpanch office. The delay in this letter until long time after the start of the field work meant the field teams had a hard time trying to explain the purpose of their visit to the community officials.
5. In some GPs the target number of 8 GK households in the sample with kids less than 2 years old could not be met. In those incidences, it was decided to increase the child age threshold for a 'GK household with kid' from 2 years to 6 years and, if necessary, to over-sample GK households without kids to obtain the target number of 30 GK households per GP. To compensate for the lower number of sampled clients with kids less than 6 years old in some GPs, in neighbouring GPs a respectively higher number of clients with children less than 6 years old, or if possible less than 2 years old, was sampled, although always ensuring that a minimum of five non-clients with kids was sampled in all of the study GPs. Section 4.3.4 discusses this issue in more detail.
6. As a result of some technical problems with the synchronisation of data between the CAPI (Computer-Assisted-Personal-Interviewing) units employed by the survey firm during data collection and the online server, a (proportionally small) number of completed household/woman/man surveys were lost. This only concerns 0.04%, 0.22% and 0.8% of the household/woman/man surveys, respectively, which are too small in number to significantly affect the study. Also, the survey teams completed a higher number of surveys than actually required, which largely compensates not only for these technical issues but also for those incidences where either man and/or woman was unavailable for completion of the individual surveys (because of being out of station or refusals). The net shortfall in the number of surveys comes to 5 household records (0.1%), 23 woman surveys (0.6%) and 44

man surveys (1.3%). These missing records are more or less equally distributed across treatment arms.

5 Findings

The data analysis presented in the following sections of this report has two main purposes: On the one hand, the data provides a snapshot of our study population and environment, serving as a useful tool to understand the context in which the intervention is taking place.

On the other hand, the analysis allows us to formally test whether we see any systematic differences between the treatment and control group prior to the intervention starting.

As explained in Section 4.1, the evaluation methodology will be based on comparing the outcomes between the different treatment groups, i.e. 'C' (control), GK (sanitation loans only) and 'GK + ND' (sanitation loans + awareness creation). In order to be able to attribute any effects to the sanitation interventions, it is imperative that the three groups being compared are similar in all respects at the outset of the intervention. Section 4.1 explains that randomisation is the best tool at our disposal for achieving this; the key is to conduct it properly. To test whether randomisation was properly done, we will compare the observable (pre-treatment) characteristics and test that there are no significant differences in their distribution between the different treatment arms. If we pass the test, this can be taken as evidence that the samples are balanced in the unobservable dimension as well, given there has been randomisation in the first place.

At baseline we can compare variables such as consumption, enterprise, assets and savings, as well as background characteristics that cannot be changed by the program such as age, sex, adult education, and so on. We present tables showing the average values of different variables for each of the treatment groups. The standard way to check that the randomisation was successful and that the two sets of data are not significantly different is the Student's t-test. This t-test compares the means of the treatment and control populations and returns a value known as a *p-value*, which is a measure of the probability that an outcome as the one observed is the result of randomly generating samples under the assumption that both means are identical. A p-value corresponds to a significance level and we will present in our tables asterixes, indicating significant differences between the treatment groups and the control group at the 10, 5 and 1 percent significant level.

Our key balance test will however be based on the statistical joint 'F-test' to test whether overall there are any differences between any of the three evaluation groups. The advantage of this test is that it controls for the fact that we are making multiple comparisons, which a normal t-test fails to do. For that reason, if we find that a difference in characteristics passes that test, we conclude that the treatment arms are balanced. The disadvantage of this test is that if we reject the balance test, i.e. if the test suggests at least one difference, we do not know which treatments can be said to be significantly different from each other. This can be inferred from the results of two-way comparisons between control and each of the treatment groups just discussed.

Figure 13: Guide to technical jargon

Magnitude of a difference in outcome

In the first instance, we will be interested in the size or magnitude of the difference between people who receive support (treated group) and people who did not (control group). For instance, imagine that we measure at baseline that the average sampled beneficiary household earns \$50 more than the average control household in our study sample, we would conclude that the magnitude of the difference is \$50.

Statistical significance and p-value

However, when we talk about ‘significant’ or ‘insignificant’ differences in this report, we mean differences that are ‘statistically significant’ or ‘statistically insignificant’, respectively. In the example above, although a difference of \$50 seems quite large, we remember that this estimated average impact is derived from data on a sample of treatment and control households, rather than data on the true population of beneficiaries and their respective counterfactual values (see discussion in Section 4.1). Given that we only have one sample available, there is always a possibility that we have been unfortunate with our sample and that we happen to have drawn beneficiaries with relatively high values from a population pool with a true average difference of zero. Therefore, it is necessary to take into consideration the statistical probability of measuring differences of \$50 (or more) if there were truly no differences in earnings between the treatment and control groups, that is, if in reality the difference was zero. This statistical probability is usually referred to as the p-value. If the p-value is small, for instance one per cent, this means that the probability of obtaining an estimate of \$50 (or more) if the true difference was zero is very small (one per cent) so we can reject the hypothesis that there is no difference in earnings ‘at the one per cent significance level’. We would then say that the result is ‘significant’. This is different from saying that the magnitude of the difference is large, because magnitude refers to effect size, as explained above. Taken together, significance and magnitude help us to interpret the results. In the results tables below, statistical significance will be indicated with asterisks, with one asterisk (*) indicating a p-value of less than one per cent, two asterisks (**) indicating a p-value of less than 5 per cent and three asterisks (***) indicates a p-value of less than ten per cent. The higher the p-value, the less confident we are that the measured difference reflects a true difference. Values with a p-value of more than ten 10 per cent are considered statistically insignificant in this report.

Joint Hypothesis Testing and the F-Statistics

In our study, we deal with two treatment groups and one control group. This presents an additional concern when comparing the mean results by treatment group. It is customary in this setting to run a joint hypothesis test (F-test), which tests for equality of all means together. This test would detect any of the several possible differences between the groups, rather than just comparing one group to another. The disadvantage of this test is that if we reject the balance test, i.e. if the test suggests at least one difference, we do not know which treatments can be said to be significantly different from each other. The pairwise tests described above can instead be used for that purpose.

The box presented in Figure 7 explains some key statistical concepts that we use in order to accept or reject balance between different groups. In sum, typical balance tests rely on statistical probabilities of observing any differences between groups if in reality there is zero difference. If this probability is lower than 10%, then we conclude that there is a statistically significant difference between groups. However, as the number of comparisons increases (for instance, if you make comparisons over many variables, which is what we do in this report), it becomes more likely that the groups being compared will appear to differ in terms of at least one attribute. For example, if one test is performed at the 10% level, there is only a 10% chance of incorrectly concluding that there is a difference even if there is no true difference between two groups. However, for 100 comparisons, the expected number of incorrect rejections is 10. We therefore expect (and therefore accept) there to be around 1 statistically significant difference for every 10 comparisons we make.

Before proceeding, note that in all of the tables presented below, we use the following format: The first column gives information on the variable under consideration. We then show the mean for the whole sample (treatment groups and control combined). The following three columns show the mean of the control and the treatment groups separately. The fifth column shows the F-statistic of the test of statistical differences between any of the treatment groups and the second last column shows the associated p-value. The last column shows the total the number of observations over which the whole sample mean is calculated. We also run two-way comparisons between each of the treatment groups and the control group. Statistical differences with the control group, if any, are indicated with asterixes (*). Note that throughout, the tests account for clustering of the standard errors at the gram panchayat level.

We will refer to the sample average when presenting the results, unless the per treatment group averages are statistically different.

5.1 Study population (listing data)

We start off by presenting the key characteristics of our study population, based on the information that was collected during the listing survey in 107 of our study GPs. Given that the respondent to the (quick) listing survey was not necessarily the household head or another knowledgeable member of the household, some answers were set to ‘don’t know’ leading to missing values. This means that the total number of observations (presented in the last columns of the relevant tables) does not always add up to the total number of households in the listing dataset (i.e. 38,868). We note though that the drops are typically very small as in most cases a household member could be interviewed as we will see shortly.

We provide information on who responded to the listing survey in Table 15. We can see that 84 percent of the interviews were conducted by the household head, his/her spouse, the son/daughter or son-in-law/daughter-in-law of the head of household. The other surveys were conducted by other relatives of the household (14%) or non-related persons such as neighbours (3%). The composition in respondent type is very well balanced over the treatment arms.

Table 16 presents general demographic information by treatment status, which is

Table 15: Respondent to the listing survey (%)

	Treatment Status				P-value	N
	Whole Sample	Control	GK	GK + ND		
Head	33.85 (1.031)	35.40 (1.517)	34.90 (2.096)	31.27* (1.629)	0.154	38848
Wife or Husband	32.99 (1.249)	32.02 (1.842)	33.59 (2.491)	33.42 (2.158)	0.835	38848
Son or Daughter	12.37 (0.467)	12.99 (0.836)	11.09 (0.899)	12.95 (0.623)	0.193	38848
Son-in-Law or Daughter-in-Law	4.50 (0.250)	4.39 (0.396)	4.19 (0.477)	4.90 (0.418)	0.500	38848
Grandchild	0.80 (0.0669)	0.75 (0.112)	0.86 (0.131)	0.80 (0.102)	0.816	38848
Parent	5.35 (0.309)	5.05 (0.496)	5.01 (0.526)	6.00 (0.559)	0.352	38848
Parent-in-Law	0.31 (0.0541)	0.18 (0.0550)	0.34 (0.0827)	0.42* (0.126)	0.098*	38848
Brother or Sister	3.93 (0.255)	3.70 (0.408)	4.16 (0.509)	3.96 (0.400)	0.767	38848
Brother-in-Law or Sister-in-Law	1.27 (0.125)	1.04 (0.209)	1.30 (0.213)	1.47 (0.224)	0.363	38848
Niece or Nephew	0.83 (0.114)	0.87 (0.200)	0.62 (0.176)	0.99 (0.208)	0.366	38848
Other Relative	1.06 (0.126)	0.91 (0.186)	1.09 (0.190)	1.19 (0.266)	0.633	38848
Adopted/Foster/Stepchild	0.07 (0.0217)	0.03 (0.0139)	0.10 (0.0580)	0.08 (0.0307)	0.233	38848
Not Related	0.19 (0.0463)	0.22 (0.0722)	0.24 (0.113)	0.12 (0.0383)	0.338	38848
Neighbour	2.45 (0.269)	2.41 (0.495)	2.51 (0.486)	2.45 (0.409)	0.990	38848
Don't know	0.02 (0.0122)	0.05 (0.0327)	0.01 (0.00806)	0.01 (0.00762)	0.417	38848

Note: Standard Errors in parenthesis, clustered at the GP level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Listing Questionnaire.

balanced across groups. Only 9% of the households interviewed had a female household head. The average self-reported age of the household head was about 48 years and that of the head's spouse was 41 years. Furthermore, the average number of household members is close to 5 and the average number of children under 2 years per household is 0.18.

Table 16: General Demographics

	Treatment Status				P-value	N
	Whole Sample	Control	GK	GK + ND		
Female (HH head)	9.17 (0.301)	9.25 (0.520)	9.67 (0.512)	8.61 (0.509)	0.342	38868
Age of HH head	47.76 (0.267)	47.92 (0.467)	47.54 (0.474)	47.81 (0.444)	0.842	38589
Age of HH head's Spouse	40.98 (0.245)	41.14 (0.450)	40.86 (0.416)	40.94 (0.399)	0.898	34673
No. of HH members	4.88 (0.0591)	4.91 (0.100)	4.75 (0.123)	4.96 (0.0740)	0.341	38817
No. of children under 2 years	0.18 (0.00444)	0.19 (0.00735)	0.17 (0.00866)	0.18 (0.00652)	0.269	38868

Note: Standard Errors in parenthesis, clustered at the GP level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Listing Questionnaire.

Table 17: Caste (%)

	Treatment Status				P-value	N
	Whole Sample	Control	GK	GK + ND		
Other Backward Class (OBC)	23.53 (1.739)	22.43 (2.889)	25.76 (3.641)	22.55 (2.386)	0.727	38855
Special Backward Castes (SBCs)	1.43 (0.265)	1.53 (0.477)	1.61 (0.511)	1.16 (0.386)	0.732	38855
Scheduled Castes (SCs)	23.56 (1.091)	24.18 (1.749)	22.40 (2.339)	24.02 (1.533)	0.808	38855
Scheduled Tribes (STs)	4.65 (0.711)	4.35 (0.967)	5.43 (1.713)	4.23 (0.903)	0.820	38855
Denotified Tribes (DTs) / Vimukta Jati (VJ)	0.86 (0.254)	1.34 (0.580)	0.37 (0.103)	0.82 (0.432)	0.174	38855
Nomadic Tribes (NTs) / Bhatkya Jati	10.91 (1.070)	10.70 (1.958)	9.99 (1.782)	12.01 (1.776)	0.719	38855
General	34.11 (1.814)	34.63 (3.008)	33.80 (3.466)	33.87 (2.963)	0.978	38855
Do not want to reveal	0.27 (0.0913)	0.09 (0.0384)	0.21 (0.0931)	0.52* (0.246)	0.125	38855
Don't Know	0.67 (0.148)	0.76 (0.230)	0.43 (0.132)	0.81 (0.344)	0.325	38855

Note: Standard Errors in parenthesis, clustered at the GP level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Listing Questionnaire.

In terms of caste composition in our study areas, we can see from Table 17 that the most common castes are General (34%), Scheduled Caste (24%), Other Backward Caste (24%) and Nomadic Tribes (11%), with no significant differences between groups.

Moving on to latrine use, Table 18 shows that the majority of households (65%) practice open defecation, while only 29% use any type of toilet (27% at home and 2% at a neighbour's, school or community toilet). We find a slight imbalance in the percentage of households going for OD for the GK+ND study group. We note however that this imbalance is only significant at the 10% significance level and the F stat is not significant. While we are generally not too concerned about imbalances at this level for the pairwise t-test, given that this is one of our main outcomes, it deserves a bit more attention. We break down the information on open defecation practices, splitting the sample by households that report to go for OD close to their own and those reporting

Table 18: Latrine ownership/usage (%)

	Treatment Status				P-value	N
	Whole Sample	Control	GK	GK + ND		
Use HH toilet	27.15 (1.808)	26.17 (2.592)	30.45 (3.659)	25.05 (3.043)	0.502	38868
Use other toilet	1.56 (0.144)	1.50 (0.273)	1.72 (0.281)	1.47 (0.180)	0.753	38868
HH toilet not in use	0.65 (0.0815)	0.57 (0.102)	0.61 (0.166)	0.77 (0.151)	0.534	38868
Open Defecation	64.57 (2.278)	67.79 (3.165)	57.57* (4.644)	67.86 (3.686)	0.152	38868

Note: Standard Errors in parenthesis, clustered at the GP level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Listing Questionnaire.

to go further away (at least 5min walking). When doing so, the small imbalance we see otherwise disappears. It is interesting to note that, only 7% of households report that they go for OD within a 5min walking radius from their home (not shown).

We also find that few households report to have a toilet that is not in use: less than 1% of the households fall into this category.

Table 19: Age of owned toilet (%)

	Treatment Status				P-value	N
	Whole Sample	Control	GK	GK + ND		
Here when we moved in	1.21 (0.290)	1.08 (0.271)	1.64 (0.746)	0.86 (0.212)	0.537	10806
Less than 3 months old	6.16 (0.610)	6.55 (1.175)	6.24 (1.030)	5.67 (0.914)	0.825	10806
Between 3 and 6 months old	4.05 (0.361)	3.27 (0.406)	4.23 (0.666)	4.69* (0.742)	0.179	10806
Between 6 months and 1 year	8.72 (0.691)	7.05 (0.812)	9.66 (1.572)	9.44* (0.956)	0.108	10806
Between 1 and 2 years old	15.97 (0.958)	16.75 (1.492)	14.93 (1.969)	16.32 (1.267)	0.755	10806
Between 2 and 3 years old	14.47 (0.851)	14.98 (1.419)	14.88 (1.691)	13.47 (1.149)	0.651	10806
More than 3 years	48.13 (2.227)	49.43 (3.578)	46.46 (4.312)	48.64 (3.482)	0.866	10806
Don't know	1.28 (0.278)	0.89 (0.335)	1.96 (0.608)	0.92 (0.288)	0.270	10806

Note: Standard Errors in parenthesis, clustered at the GP level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Listing Questionnaire.

From Table 19 we learn that for those with a toilet at home, a large portion of households (48%) have a toilet that has been constructed at least 3 years ago, 15% of households have a toilet for between 2 and 3 years and 16% of the respondents with latrine report that their toilet is between 1 and 2 years of age. Only 19 % of households with a toilet (or 5% of the total population) constructed a toilet in the year preceding

the listing survey visit.

5.2 Village profile (community survey)

In this section, we discuss the findings from the community survey conducted with GP officials regarding the characteristics of the village that was sampled for our study. At the time of writing this baseline report, the village questionnaire data for one GP were still pending (see Section 4.4.2). Therefore, this section only presents the information available for 119 GPs.

4.1.1 Demographics

Table 20: General Demographics from Community Data

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
What is the current population of this village?	3860.78 (309.4)	3660.13 (189.1)	4150.00 (489.6)	3769.95 (607.8)	0.476	0.636	119
How many household are in this village?	594.05 (37.03)	618.38 (33.10)	598.35 (38.65)	564.69 (58.04)	0.367	0.703	119

Note: Standard Errors in parenthesis, clustered at the block level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Community Questionnaire.

Table 20 summarises information on the size of the average village in our study sample, which is relatively large: the GP officials report an average village population size of 594 households and an average of 3,861 residents in the village. This former estimate is close to the one we used for classification of GPs into small and large groups for stratification purpose. It was based on the 2011 census numbers, accounting for population growth, and resulted in an average village size of 605 households. Combined with the resident estimate, the community data implies an average household size of about 6.5 members. We will see later that this is a slight overestimate: Our listing and household data indicate that the average number of households is about 0.5-1 members lower. Important for the purpose of this report is that estimates provided by village officials are balanced across treatment arms as is indicated in the Table by large p-values in the second to last column.

4.1.2 Economic Activities

Village leaders were asked about the main three economic activities in their respective villages and Table 21 presents the findings. All respondents report farming and agricultural labour. Since there is no variation across villages, the first two rows do not provide any F-stat or P-value. We find more variation in the third activity reported: 79% of the villages have families involved in daily wage labour (other than agriculture), and 66% have households that are employed in shops. Other activities mentioned were animal husbandry (goats and sheep) and dairy, pottery and other handy crafts as well as other services.

Table 21: Community's engagement in economic activities (%)

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Farming	100.00 (0)	100.00 (0)	100.00 (0)	100.00 (0)	.	.	119
Agricultural labourer	100.00 (0)	100.00 (0)	100.00 (0)	100.00 (0)	.	.	119
Daily wage labour	78.99 (3.750)	75.00 (9.270)	75.00 (9.410)	87.18 (8.566)	0.947	0.423	119
Goat/Sheep raising	26.89 (4.082)	27.50 (5.898)	30.00 (9.972)	23.08 (9.270)	0.617	0.561	119
Dairy	10.92 (2.872)	15.00 (7.461)	12.50 (5.629)	5.13 (3.472)	1.074	0.382	119
Shops	65.55 (4.375)	62.50 (12.63)	70.00 (10.45)	64.10 (16.19)	0.386	0.691	119
Potters	21.01 (3.750)	22.50 (6.324)	17.50 (8.368)	23.08 (7.148)	0.324	0.731	119
Other handicrafts	21.01 (3.750)	17.50 (3.707)	25.00 (7.096)	20.51 (5.639)	0.587	0.576	119
Other services	15.97 (3.372)	15.00 (5.079)	10.00 (4.916)	23.08 (7.148)	3.216	0.088*	119

Note: Standard Errors in parenthesis, clustered at the block level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Community Questionnaire.

Respondents were then asked to report on average wages paid in their village for the main three activities. These are presented in Table 22. Average daily wages over the three main activities in the village are 221 Rs per day for males and 133 Rs per day for females, putting female earnings at 60% of those of males. 29 villages (~24%) report values for minors. For this sub-set of villages the average wage reported is Rs. 105 per day per minor. The table also presents the average wages for adult males and females for the two common activities in our study villages, farming and agricultural labour. Farming gives a slightly lower daily wage than the average just presented, namely Rs. 215 for males and Rs. 127 for females, and agricultural labour pays again a bit less at on average Rs. 208 for males and Rs.124 for females. None of these activities and corresponding wages differ statistically between treatment and control arms.

In Table 23 we show the major crops, grains and fruit grown by farmers in our study areas. We can see that our study villages are most homogeneous in terms of the types of grains they grow, with 92% of villages growing staple dhal and 85% growing Jowar. Also tomatoes are very common with 77% of study villages producing this vegetable. Crops and other grains and fruits show slightly more variation. Still quite common ones to be grown are cotton (56%) sugarcane (50%), groundnut (37%) and okra (59%). The GK treatment group reports to depend statistically more (at the 5% level) on sugarcane than the control group - 58 vs 40%, but otherwise the villages are well balanced in terms of the agricultural products.

Table 22: Daily Wages

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Av. Wage for Males (Top 3 activities)	221.19 (2.902)	220.71 (4.565)	223.42 (3.350)	219.40 (5.910)	0.186	0.833	119
Av. Wage for Females (Top 3 activities)	132.54 (2.530)	130.33 (4.514)	139.54 (5.719)	127.61 (4.662)	1.959	0.197	119
Av. Wage for Minors (Top 3 activities)	104.94 (4.187)	106.25 (6.847)	102.00 (3.405)	106.67 (9.802)	0.152	0.863	29
Avg farming wage for males	214.20 (4.686)	206.50 (5.726)	217.50 (6.624)	218.72 (10.26)	1.278	0.325	119
Avg farming wage for females	127.31 (3.077)	123.75 (3.879)	132.00 (2.460)	126.15 (4.603)	1.097	0.375	119
Avg wage for male agric. labourers	208.31 (3.219)	208.25 (5.937)	209.49 (2.972)	207.18 (7.504)	0.045	0.956	118
Avg wage for female agric. labourers	124.24 (2.737)	123.63 (4.087)	129.74 (6.261)	119.36 (5.080)	1.155	0.358	118

Note: Standard Errors in parenthesis, clustered at the block level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Community Questionnaire.

Table 23: Main cash crops, grains, fruits and vegetables grown in the community (%)

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Groundnut	36.97 (4.444)	45.00 (12.97)	35.00 (12.67)	30.77* (9.251)	1.814	0.218	119
Sugarcane	49.58 (4.603)	40.00 (15.60)	57.50** (16.95)	51.28 (20.38)	3.697	0.067*	119
Cotton	56.30 (4.566)	60.00 (15.60)	57.50 (11.96)	51.28 (21.71)	0.602	0.568	119
Turmeric	10.92 (2.872)	15.00 (8.043)	10.00 (5.055)	7.69 (2.460)	0.598	0.570	119
Sorghum	8.40 (2.554)	10.00 (6.530)	7.50 (2.804)	7.69 (5.127)	0.157	0.857	119
Staple dhal	91.60 (2.554)	95.00 (3.471)	92.50 (7.256)	87.18 (3.289)	1.229	0.337	119
Wheat	32.77 (4.321)	37.50 (9.398)	25.00 (5.103)	35.90 (4.489)	3.482	0.076*	119
Rice	2.52 (1.443)	2.50 (2.295)	0.00 (0)	5.13 (2.749)	1.710	0.235	119
Jowar	84.87 (3.298)	77.50 (9.189)	90.00 (4.773)	87.18 (4.492)	1.004	0.404	119
Okra	58.82 (4.531)	65.00 (5.599)	47.50 (6.091)	64.10 (5.258)	1.969	0.195	119
Cauliflower	32.77 (4.321)	37.50 (7.321)	37.50 (8.765)	23.08 (5.453)	1.898	0.205	119
Tomatoes	77.31 (3.856)	72.50 (9.828)	80.00 (6.831)	79.49 (5.835)	0.375	0.697	119

Note: Standard Errors in parenthesis, clustered at the block level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Community Questionnaire.

4.1.3 Service and Product Availability

We now concentrate on the types of services and products that villages have access to, starting with those relevant for the construction of toilets: Table 24 displays information on the percentage of villages with access to different artisans and shops in their community. It can be seen that access to some crucial artisans for toilet construction is provided: 93% report to have a mason and 87% a carpenter. Further 51% have a plumber in their village. Access to materials is on the other hand not provided for the majority of the villages: only one third report to have a concrete block producer, 20% a brick producer and 18% a sanitary hardware store. We see a highly significant imbalance between the GK+ND treatment arm and the control group in terms of access to carpenters, the treatment group having significantly lower access. As we discussed above, some imbalances are expected to be found at random, which, given the overall balancedness seen so far is what we expect to lie behind this imbalance.

Our data provides us with some more information on access, namely distances to these services in case not available in the village. For communities that do not have services discussed in Table 24, the travel distance is as follows: on average 9km to reach a mason, approximately the same to reach a plumber, carpenters, cement block producers and sanitary hardware stores are reached within roughly 10km and furthest away are brick producers with an average distance of 21kilometers. These statistics are not reported in any Table.

Table 24: Percentage of villages with Artisans available within the village

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Mason	93.28 (2.305)	95.00 (2.681)	92.50 (5.579)	92.31 (4.546)	0.233	0.796	119
Plumber	51.26 (4.601)	55.00 (6.527)	60.00 (11.45)	38.46* (9.752)	3.505	0.075*	119
Carpenter	86.55 (3.140)	90.00 (6.841)	92.50 (4.435)	76.92*** (7.303)	11.024	0.004***	119
Cement Block Producer	32.77 (4.321)	32.50 (4.698)	35.00 (5.969)	30.77 (11.86)	0.109	0.898	119
Sanitary Hardware Store	17.65 (3.509)	22.50 (4.971)	15.00 (3.755)	15.38 (5.357)	0.625	0.557	119
Brick Producer	19.33 (3.635)	22.50 (5.143)	12.50 (7.598)	23.08 (8.378)	2.664	0.123	119

Note: Standard Errors in parenthesis, clustered at the block level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Community Questionnaire.

Table 25 provides more information on access to construction materials in the villages. In line with the observation from Table 24 we see that availability is quite low: only 17% of villages have cement within their village available, 13% have bricks, and 7-7% concrete blocks and rings. We see again only one imbalance for the GK+ND study arm.

In terms of village access to health facilities we can see in Table 26 that a bit more than half of the study communities have access to some type of health facilities. The

Table 25: Construction materials available in the village (%)

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Cement	16.81 (3.442)	20.00 (5.869)	20.00 (7.032)	10.26 (4.000)	0.755	0.497	119
Brick	13.45 (3.140)	15.00 (5.346)	17.50 (7.607)	7.69 (3.577)	0.819	0.471	119
Concrete block	6.72 (2.305)	7.50 (4.021)	10.00 (7.012)	2.56 (2.653)	0.538	0.602	119
Concrete ring	8.40 (2.554)	12.50 (4.908)	10.00 (7.012)	2.56** (2.653)	2.654	0.124	119

Note: Standard Errors in parenthesis, clustered at the block level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Community Questionnaire.

different types are displayed in the same table and we can see that about one third of the sample has access to a sub-center or ANM (auxiliary nurse-midwife), and about the same proportion has access to a private clinic. Most unusual to have within the village is a government hospital (4%). As above we calculate, but don't report, the average distances to the service if not located in the village and we learn that, except for the government hospital, which is on average 16km away, all facilities can be reached within ten kilometers at a maximum. We also ask for one-way costs to reach these facilities. These costs, which are balanced across treatment arms, are roughly equivalent to one Rupee per kilometer travelled.

Table 26: Access to health facilities (%)

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Access to some health facility	52.94 (4.595)	55.00 (3.851)	62.50 (6.215)	41.03 (8.448)	3.738	0.066*	119
Sub Centre or ANM	35.29 (4.399)	40.00 (8.097)	37.50 (10.31)	28.21 (7.476)	0.631	0.554	119
Primary Health Centre (Phc)	23.53 (3.905)	20.00 (6.540)	25.00 (4.751)	25.64 (5.790)	0.702	0.521	119
Community Health Centre	10.92 (2.872)	17.50 (3.799)	7.50* (3.154)	7.69** (5.016)	4.447	0.045**	119
Government Hospital	4.20 (1.847)	2.50 (2.295)	5.00 (3.195)	5.13 (3.682)	0.241	0.791	119
Private Clinic	32.77 (4.321)	32.50 (6.682)	42.50 (2.866)	23.08 (7.531)	4.436	0.046**	119
Private Hospital/Doctor	27.73 (4.121)	30.00 (6.593)	32.50 (7.630)	20.51 (9.158)	1.252	0.331	119

Note: Standard Errors in parenthesis, clustered at the block level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Community Questionnaire.

We also present availability of medicines in the villages, which can be seen in Table 27. The two medicines we have information on are antibiotics as well as ORS solutions. The latter is available within only just over 40% of the study villages and antibiotics is available for less than 30%. We find this to be balanced across study arms.

Table 27: Medicines available within the village (%)

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
ORS solution	41.18 (4.531)	45.00 (10.68)	40.00 (5.578)	38.46 (7.467)	0.107	0.899	119
Antibiotics	28.57 (4.159)	22.50 (6.460)	35.00 (5.609)	28.21 (7.195)	1.541	0.266	119

Note: Standard Errors in parenthesis, clustered at the block level. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$
Source: Community Questionnaire.

The final set of services we consider are NGOs and GP level committees focused on those related to water, sanitation and education. These are presented in Table 28. It is relevant to highlight that while the presence of projects from NGOs focusing on water is about 26%, only 2.5% of villages are subject to NGOs offering sanitation related services. Most villages have on the other hand established a Health, Sanitation and Nutrition committee (ranging from 85% to 97% of communities across treatment arms). It is worth noting that the data reports statistically significant differences between the groups both from the pairwise t-tests and the F test, at the 1% level. Specifically, the GK + ND treatment group reports lower community organization in topics of Health, Sanitation and Nutrition, (69.23% vs 92.68% from the control group). To the extent that previous existence of Sanitation Committee or Health Centres relates with community views on sanitation or health, our estimates on sanitation and health insurance take-up will (likely) be biased downwards. We can however account for this information in our analysis to avoid biased impact findings.

Table 28: NGO and Village committees

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
NGO undertaking water projects	26.05 (4.040)	27.50 (12.22)	25.00 (10.29)	25.64 (9.935)	0.037	0.964	119
NGO undertaking sanitation projects	2.52 (1.443)	2.50 (2.295)	2.50 (2.271)	2.56 (2.653)	0.000	1.000	119
Village Health, Sanitation and Nutrition Committee	84.87 (3.298)	92.50 (4.772)	92.50 (4.882)	69.23*** (4.876)	49.405	0.000***	119
Village Water committee/association	92.44 (2.434)	92.50 (3.753)	95.00 (3.405)	89.74 (5.533)	0.555	0.592	119

Note: Standard Errors in parenthesis, clustered at the block level. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$
Source: Community Questionnaire.

4.1.4 Finance and Insurance Services

We next review the financial services available in the villages (Table 29). According to our data, (micro)finance services are not widely available in our study villages: only about 7% have a microfinance institution offering credit. Interestingly, almost all of the villages (8 out of 9 villages that report to have access to an MFI in the village) report that the MFI offers loans for sanitation and/or home improvement purposes.

Table 29: Availability of Financial Services Availability

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Bank	21.85 (3.804)	25.00 (7.156)	30.00 (5.477)	10.26 (4.847)	5.306	0.030**	119
Microfinance Institution	6.72 (2.305)	5.00 (2.681)	12.50 (5.929)	2.56 (2.354)	1.215	0.341	119
Life Insurance - government	60.50 (4.500)	55.00 (7.329)	62.50 (8.961)	64.10 (8.991)	0.585	0.577	119
Health Insurance - government	20.17 (3.694)	15.00 (8.000)	22.50 (7.261)	23.08 (10.73)	0.915	0.435	119
Crop Insurance - government	17.65 (3.509)	17.50 (7.326)	17.50 (5.702)	17.95 (7.741)	0.007	0.993	119
Life Insurance - private	34.45 (4.375)	37.50 (13.94)	32.50 (5.823)	33.33 (9.703)	0.106	0.900	119
Health Insurance - private	14.29 (3.221)	17.50 (7.741)	15.00 (7.514)	10.26 (7.130)	1.031	0.395	119
Crop Insurance - private	14.29 (3.221)	20.00 (8.563)	12.50 (7.366)	10.26 (5.533)	1.444	0.286	119

Note: Standard Errors in parenthesis, clustered at the block level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Community Questionnaire.

Unfortunately, we did not ask this separately, so cannot say whether it is the home improvements and/or sanitation loans that are offered.

Of further interest in the context of this research study (particularly the longer-term objective), is the availability and access to insurance products. The most common product available is life insurance, with 60% of villages reporting access to this product from the government and about 35% from private insurance companies. Also available is health insurance, with 20% of villages reporting coverage from government providers and 14% from private ones. Similar access proportions are reported for crop insurance. All these are balanced across treatment arms. We find one imbalance in the access to banks, for which the access proportion is on average 20%, but is observed to be significantly lower for our GK+ND treatment arm.

4.1.5 Community Life

Before embarking into our final and main community variables of interest (sanitation, water and waste disposal), let's review briefly the general community life of our study GPs. This is of interest since some of the activities conducted by ND target communities and one might believe that these would be more efficient in communities which are already active in their communal life. We show in Table 30 that our study GPs seem to have a strong community life, with 91.67% of GPs holding religious activities, 90.83% political activities and 89.17% social activities happening. We find this to be balanced across treatment arms.

Table 30: Community-wide Activities from Community Data

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Religious activities	92.44 (2.434)	92.50 (4.432)	92.50 (3.469)	92.31 (4.546)	0.001	0.999	119
Political activities	91.60 (2.554)	92.50 (4.432)	92.50 (3.757)	89.74 (3.164)	0.478	0.635	119
Social activities	89.08 (2.872)	90.00 (3.051)	90.00 (3.979)	87.18 (5.079)	0.129	0.880	119

Note: Standard Errors in parenthesis, clustered at the block level. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$

Source: Community Questionnaire.

4.1.6 Sanitation

We now embark on describing the sanitation situation in the villages, as inferred from the community questionnaire.

Table 31 shows that, congruent with findings we will present below on the listing and household data, open defecation is widespread practice with 70% of GPs stating that it is either very common or common for their community members to openly defecate. The most commonly reported places where people practice open defecation are: in the field (68%), along the roadside (58%), in woods/forests (28%), and along the river/lake (19%). Regretfully, as increasingly highlighted in the media, OD is not perceived to be safe for women. In line, we find that, on average, 24% of villages report that the OD places frequented are not safe for women. In addition, on average 30% of these OD places for women are reported to be exposed. These practices are statistically the same across our treatment and control groups.

Table 31: Community Sanitation Practices

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Open defecation (very) common	70.59 (4.195)	67.50 (9.196)	72.50 (5.098)	71.79 (6.483)	0.237	0.794	119
OD - field	68.07 (4.292)	75.00 (3.608)	70.00 (7.322)	58.97 (9.150)	1.319	0.314	119
OD - roadside	57.98 (4.544)	57.50 (15.70)	60.00 (14.36)	56.41 (12.07)	0.091	0.914	119
OD - woods/forest	28.57 (4.159)	25.00 (9.270)	35.00 (14.15)	25.64 (7.681)	0.627	0.556	119
OD - river/lake	19.33 (3.635)	17.50 (7.420)	20.00 (10.90)	20.51 (9.480)	0.160	0.854	119
OD - railway line	3.36 (1.659)	7.50 (4.985)	0.00 (0)	2.56 (2.354)	1.154	0.358	119
OD - other	10.08 (2.772)	12.50 (5.571)	7.50 (3.757)	10.26 (4.962)	0.819	0.471	119
Are the places safe for women?	24.37 (3.952)	22.50 (5.956)	20.00 (6.412)	30.77 (6.286)	0.744	0.502	119
Are the places for women exposed?	30.25 (4.229)	25.00 (13.27)	30.00 (5.349)	35.90 (2.986)	0.623	0.558	119

Note: Standard Errors in parenthesis, clustered at the block level. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$

Source: Community Questionnaire.

We also collected information on the sanitation facilities available at schools in the GPs. Table 32 shows our findings for government and private primary schools. Our data shows that GPs have on average one government-run primary school, and that these are generally well equipped with respect to sanitation infrastructure: 96% of the schools have sanitation facilities, 87% have sanitation facilities (supposedly working and accessible) specific for children and 83% have such sanitation facilities specific for girls. Privately run primary schools are much less prominent in our study vilages with on average one school per ten GPs. The sanitation infrastructure is reported to be considerably worse with on average ten percentage points lower availability than in government run schools. A similar overall pattern holds for middle schools, albeit with much lower average number of government middle schools available (0.36 per GP). We find no imbalances across study arms.

Table 32: Community Schools and Sanitation Infrastructure

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
No. Primary- government	0.99 (0.00840)	1.00 (0)	1.00 (0)	0.97 (0.0265)	0.918	0.363	119
San. Facilities: Primary- government	95.76 (1.862)	97.50 (2.512)	95.00 (4.838)	94.74 (4.083)	1.833	0.215	118
San. Facilities for children: Primary- government	87.29 (3.080)	92.50 (5.089)	82.50 (8.512)	86.84 (6.045)	1.437	0.287	118
San. Facilities for girls: Primary- government	83.90 (3.398)	87.50 (5.571)	82.50 (8.512)	81.58 (7.953)	0.338	0.722	118
No. Primary- private/aided	0.13 (0.0314)	0.20 (0.0511)	0.10 (0.0506)	0.10 (0.0316)	1.225	0.338	119
San. Facilities: Primary- private/aided	87.50 (8.539)	87.50 (9.682)	75.00 (25.62)	100.00 (0)	0.929	0.454	16
San. Facilities for children: Primary- private/aided	68.75 (11.97)	75.00 (13.69)	75.00 (25.62)	50.00 (27.39)	0.588	0.590	16
San. Facilities for girls: Primary- private/aided	62.50 (12.50)	62.50 (11.86)	75.00 (25.62)	50.00 (27.39)	0.223	0.808	16

Note: Standard Errors in parenthesis, clustered at the block level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Community Questionnaire.

The final community level sanitation related variables we discuss relate to sanitation campaigns and activities that took place the year previous to the baseline survey. As can be seen in Table 33, 28% of villages report for some activities to have taken place and this number is balanced across treatment groups. Most common are street plays, which 75% of those villages that had any activities report. Village meetings are also quite common (44%). Less frequently reported are film showings (17%), wall paintings (11%) and flyer distributions (2%).

4.1.7 Water Sources & Garbage Disposal

We finally turn to present information on village access to water source and practices related to garbage disposal.

Table 34 shows the results on water sources. We find that only 57% of GPs report to provide access to piped water to all or most of their households, making it the most commonly used source of water for drinking and cooking purpose. In line with this, 88% of villages report that their households get drinking water piped into their home or yard. Other reported sources include public handpumps (52%), tubewell or borewell

Table 33: Sanitation-related Activities from Community Data

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Sanitation-related activities	37.82 (4.464)	37.50 (8.215)	37.50 (11.20)	38.46 (7.977)	0.014	0.987	119
Street Plays	75.56 (6.479)	80.00 (7.621)	80.00 (11.88)	66.67 (9.601)	1.164	0.374	45
Film Showings	17.78 (5.764)	20.00 (11.88)	26.67 (10.50)	6.67 (6.405)	0.928	0.445	45
Village Meetings	44.44 (7.491)	40.00 (11.34)	46.67 (11.68)	46.67 (11.97)	0.677	0.543	45
Flyer Distribution	2.22 (2.222)	6.67 (7.406)	0.00 (0)	0.00 (0)	0.773	0.413	45
Wall Paintings	11.11 (4.738)	20.00 (13.51)	13.33 (8.614)	0.00 (0)	2.139	0.199	45

Note: Standard Errors in parenthesis, clustered at the block level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
 Source: Community Questionnaire.

(47%), public taps (42%), and protected dug wells (37%). We find statistical differences between groups in usage of handpumps and public taps. In particular, we find that the GK group has lower access on both relative to the control only for the former significant at the 5% significance level though.

Table 34: Water Sources from Community Data

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Do most HHs have access to piped water?	57.14 (4.556)	52.50 (8.724)	62.50 (10.81)	56.41 (11.00)	0.331	0.726	119
HHs drink from Piped water into dwelling yard/plot	88.24 (2.966)	85.00 (4.651)	87.50 (7.003)	92.31 (3.577)	0.858	0.456	119
HHs drink from Public tap/stand pipe	41.18 (4.531)	45.00 (7.877)	27.50* (7.702)	51.28 (9.077)	2.091	0.180	119
HHs drink from Tubewell or borewell	46.22 (4.590)	45.00 (10.75)	55.00 (9.268)	38.46 (3.765)	2.441	0.142	119
HHs drink from Handpump	51.26 (4.601)	57.50 (6.733)	40.00** (6.055)	56.41 (7.808)	3.138	0.093*	119
HHs drink from Dug well - protected	36.97 (4.444)	32.50 (7.605)	42.50 (8.450)	35.90 (7.185)	0.443	0.656	119

Note: Standard Errors in parenthesis, clustered at the block level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
 Source: Community Questionnaire.

Our survey also collected information on general garbage disposal habits. While we have information on both short-term (daily) and longer-term basis, we present here only the former, noting that the habits do not differ greatly by this distinction. We find that the most common means of disposing of garbage in the study area is to simply leave (55%) or bury (45%) it on the own land. Another quite common practice is to burn the garbage on own land (36%) or somewhere else (22%), or people simply throw it in the street (35%). This practice shows the only imbalance significant at the 5% significance level, indicating that communities allocated to the GK treatment arm are less likely to throw their daily rubbish into the street.

Table 35: Garbage Disposal from Community Data

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Daily: Throw it away in waste baskets - trucks pick it up	8.40 (2.554)	10.00 (5.737)	7.50 (3.469)	7.69 (4.420)	0.153	0.860	119
Daily: Burn on their own land	36.13 (4.422)	42.50 (7.786)	32.50 (8.471)	33.33 (7.429)	1.446	0.285	119
Daily: Burn somewhere else	21.85 (3.804)	15.00 (5.960)	35.00* (8.831)	15.38 (4.311)	2.310	0.155	119
Daily: Leave it on their own land	55.46 (4.575)	55.00 (5.024)	57.50 (4.701)	53.85 (2.563)	0.185	0.834	119
Daily: Throw it into the street / drainage	35.29 (4.399)	37.50 (5.082)	27.50** (5.491)	41.03 (13.35)	4.568	0.043**	119
Daily: Bury it on own land	45.38 (4.583)	50.00 (7.217)	47.50 (6.968)	38.46 (9.279)	0.892	0.443	119
Daily: Bury it somewhere else	20.17 (3.694)	20.00 (6.952)	27.50 (6.342)	12.82 (4.696)	1.710	0.235	119
Daily: Village garbage pit	21.01 (3.750)	25.00 (6.387)	15.00 (8.006)	23.08 (6.743)	1.277	0.325	119

Note: Standard Errors in parenthesis, clustered at the block level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Community Questionnaire.

5.3 Households in study sample (household survey)

In this section, we present summary descriptive statistics for key demographic, socio-economic, hygiene, health, and child development variables.

5.3.1 General household characteristics

Table 36 provides information on household's religion, primary activity and living status. The majority of households in our sample profess Hinduism (76%), the dominant religion in India, followed by Islam (14%) and Buddhism (10.5%). We can see from the last column, that there is no significant difference between these proportions between the treatment and control areas. For example in the control area, 76% of the sample declares to be Hindu, whereas this number stands at 72% for the GK treatment group and 78% for the 'GK + ND' group.

The overwhelming majority of households are engaged primarily in agricultural activities (77%). Of these, almost half are cultivators and the rest derive their livelihoods from doing labour for others (29%) or allied agricultural activities (11%).¹⁹ The remaining households rely on other kind of waged employment (17%) or self-employment (4%) to sustain their households.

Barely 4% of the households report any members having migrated in the last year.

Even though almost everyone in the sample holds a government benefit card, only 14% own one of the Below Poverty Line (BPL) kind, given only to the poor belonging to a vulnerable section of the society. Provided that the sanitation subsidy scheme of the SBM government program primarily targets BPL households (see Section 3.1), most households in our study sample may not be eligible for government support. This gives space for sanitation loans to potentially play an important role in tackling financial barriers to sanitation uptake.

¹⁹ Activities like dairy farming, poultry farming, bee-keeping, fisheries etc. are defined as belonging to the allied agriculture category.

Table 36: General household characteristics

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
HH religion: Hinduism	75.8 (1.634)	76.7 (2.414)	72.6 (3.428)	77.9 (2.481)	0.81	0.45	3595
HH religion: Islam	13.7 (1.500)	12.4 (2.218)	15.9 (3.248)	12.7 (2.162)	0.43	0.65	3595
HH religion: Jainism	0.056 (0.0392)	0 (0)	0.17 (0.117)	0 (0)	2.09	0.15	3595
HH religion: Buddhism	10.5 (0.934)	10.8 (1.557)	11.3 (1.794)	9.32 (1.465)	0.43	0.65	3595
Primary activity HH: Cultivator	37.1 (1.248)	35.9 (2.135)	36.2 (1.979)	39.1 (2.322)	0.59	0.56	3595
Primary activity HH: Agriculture wage labour	28.9 (1.383)	29.2 (2.242)	27.1 (2.463)	30.3 (2.465)	0.42	0.66	3595
Primary activity HH: Allied agriculture	10.9 (0.839)	11.2 (1.450)	12.6 (1.503)	8.89 (1.337)	1.74	0.18	3595
Primary activity HH: Waged employment	16.8 (0.952)	17.9 (1.599)	16.3 (1.527)	16.2 (1.804)	0.34	0.71	3595
Primary activity HH: Self-employment	4.09 (0.396)	3.47 (0.519)	4.63 (0.642)	4.19 (0.857)	1.02	0.36	3595
Primary activity HH: Other	1.72 (0.282)	1.62 (0.429)	2.44 (0.654)	1.11 (0.280)	1.93	0.15	3595
HH members have migrated last year	4.28 (0.516)	3.39 (0.687)	3.71 (0.832)	5.81* (1.080)	1.86	0.16	3595
HH owns BPL card	13.9 (0.937)	14.3 (1.537)	12.4 (1.653)	15.1 (1.657)	0.73	0.49	3595
HH owns government card	98.9 (0.176)	98.8 (0.276)	98.7 (0.370)	99.3 (0.248)	1.35	0.26	3594

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Household Questionnaire.

We next describe some general characteristics about the dwellings where the households in our sample reside (Table 37).

Almost all the sampled households live in a dwelling they own (97%). Statistically significant differences in ownership status are found among treatment groups due to the slightly larger mean recorded for the GK sub-sample. In magnitude, however, the difference is relatively small (1.6 percentage points). Similarly, the GK group presents a larger proportion of households with a pucca house²⁰ than the other groups, i.e. 21% compared to an overall sample average of 16%. Most households (60%) possess a semi-pucca house, which has either the walls or the roof but not both, made of pucca materials. The ownership of this kind of house is balanced across treatment status. This is also the case for the materials of the floor, with the sole exception of tiled floors, which are more commonly found in the dwellings of the GK group (a result directly related to the higher prevalence of pucca houses within this group). Given that semi-pucca (60%) and kutchha (23%) houses are the most common ones in our sample, it comes as no surprise that earth or mud (42%) and cement (27%) floors are the most usual ones. Firewood is used by 88% of the households as fuel for cooking, whilst most houses (92%) have electricity as their source of lighting.

²⁰These are houses whose major portion of walls and roof are made of *pucca* (“strong”) materials, like cement, concrete, oven burnt bricks, stone, timber etc.

Table 37: Characteristics HH dwelling

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Owned by HH member	97.1 (0.353)	96.4 (0.654)	98.0** (0.413)	96.8 (0.699)	2.47	0.089*	3595
Dwelling structure: Pucca House	16.3 (0.976)	14.0 (1.330)	21.2*** (1.927)	13.8 (1.490)	5.69	0.0044***	3595
Dwelling structure: Semi-pucca house	60.3 (1.425)	61.8 (2.327)	57.4 (2.264)	61.6 (2.738)	1.14	0.32	3595
Dwelling structure: Kutchra House	23.4 (1.317)	24.2 (2.202)	21.4 (2.161)	24.5 (2.441)	0.60	0.55	3595
Material floor: Mud/clay/earth	42.1 (1.538)	43.0 (2.474)	38.7 (2.545)	44.6 (2.881)	1.34	0.26	3595
Material floor: Cement	27.3 (1.323)	25.6 (2.297)	27.9 (2.127)	28.4 (2.413)	0.41	0.67	3595
Material floor: Tiles	5.34 (0.492)	5.25 (0.870)	6.82 (0.903)	3.93 (0.702)	3.22	0.043**	3595
Material floor: Other	30.7 (1.668)	31.4 (2.909)	33.4 (2.969)	27.0 (2.671)	1.39	0.25	3595
Cooking fuel: Lpg, biogas, kerosene or other	12.1 (0.923)	11.1 (1.294)	15.0* (1.947)	10.3 (1.395)	2.01	0.14	3595
Cooking fuel: Firewood	87.9 (0.923)	88.9 (1.294)	85.0* (1.947)	89.7 (1.395)	2.01	0.14	3595
Lighting fuel: Kerosene lamps, candles...	7.48 (0.615)	6.70 (1.055)	6.66 (0.857)	9.15 (1.209)	1.60	0.21	3595
Lighting fuel: Electricity	92.5 (0.615)	93.3 (1.055)	93.3 (0.857)	90.9 (1.209)	1.60	0.21	3595

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Household Questionnaire.

5.3.2 Household member characteristics

We now turn to the sampled households and their members. Table 38 shows the general composition of these household units. The average household in our sample has more than five members, of which almost three are male. The sampling strategy of the study targeted a disproportionately high number of households with kids less than 2 years, a feature corroborated by the actual data. The estimated proportion of households with very young children is similar for all groups, with an overall sample average of 44%. The number of two year olds is slightly above one for those households who have them, whereas the number of children below five is closer to two. Our average study household has close to one young child. Conditional on having a young child, households have typically almost two.

In Table 39 and Table 40 some general information about the household head and spouse is displayed. The great majority of heads of households in the sample are male (93%). They are on average 44 years of age, are married and have received close to six years of formal education, yet 30% of them are illiterate. By contrast, their (predominantly female) spouses are some five years younger, less educated and have acquired a lower level of literacy (about 50% report being illiterate). Here, we observe significant differences in baseline education and literacy levels : Both the household head and spouse in the ‘GK + ND’ group seem to have lower educational status compared to their counterparts in the control group and the GK group. Given the consistency of this pattern across the various educational outcomes, we will need to control for this during the endline analysis. We at the same time note that only one of the imbalances

Table 38: Household composition

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Nr of HH members	5.43 (0.0465)	5.45 (0.0830)	5.36 (0.0738)	5.48 (0.0830)	0.67	0.51	3595
Nr of male HH members	2.77 (0.0252)	2.75 (0.0450)	2.77 (0.0351)	2.78 (0.0492)	0.15	0.86	3595
Nr of female HH members	2.66 (0.0289)	2.70 (0.0460)	2.59* (0.0517)	2.69 (0.0502)	1.62	0.20	3595
HHs with children <2 years	46.7 (0.848)	48.5 (1.382)	45.3* (1.290)	46.3 (1.677)	1.41	0.25	3595
Nr of children <2 years	0.56 (0.0129)	0.59 (0.0211)	0.54 (0.0215)	0.55 (0.0239)	1.27	0.28	3595
HHs with children <5 years	57.1 (0.865)	58.8 (1.423)	55.1* (1.299)	57.4 (1.700)	1.89	0.16	3595
Nr of children <5 years	0.95 (0.0173)	0.98 (0.0259)	0.91* (0.0309)	0.95 (0.0323)	1.65	0.20	3595
Nr of children 6-14 years	0.78 (0.0182)	0.78 (0.0288)	0.78 (0.0317)	0.79 (0.0339)	0.035	0.97	3595
Nr of elderly (>64 years)	0.28 (0.00933)	0.28 (0.0182)	0.28 (0.0120)	0.27 (0.0173)	0.12	0.89	3595

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Household Questionnaire.

observed is significant at the 5% level.

Table 39: Information on the household head

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Gender HH head (fraction male)	92.5 (0.566)	92.2 (0.826)	92.5 (0.992)	92.9 (1.115)	0.11	0.89	3595
Age HH head	44.5 (0.314)	44.1 (0.631)	44.2 (0.485)	45.2 (0.475)	1.37	0.26	3595
Marital status HH head	92.4 (0.511)	91.6 (0.902)	92.8 (0.798)	92.6 (0.936)	0.58	0.56	3595
Years of education HH head	6.02 (0.115)	6.09 (0.161)	6.34 (0.199)	5.63* (0.219)	2.97	0.055*	3433
Head able to read	67.0 (0.983)	67.4 (1.638)	69.6 (1.562)	63.8 (1.770)	3.07	0.050*	3595
Head able to write	70.7 (1.045)	71.6 (1.431)	73.4 (1.766)	66.8* (2.044)	3.08	0.050**	3595

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Household Questionnaire.

Table 40: Information on the head's spouse

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Gender head's spouse (fraction male)	0.67 (0.221)	0.80 (0.298)	0.46 (0.193)	0.75 (0.570)	0.51	0.60	3291
Age HH head's spouse	38.1 (0.299)	37.7 (0.577)	37.9 (0.467)	38.7 (0.487)	0.98	0.38	3291
Years of education HH head's spouse	4.09 (0.113)	4.18 (0.173)	4.41 (0.197)	3.67* (0.199)	3.64	0.029**	3081
Head's spouse able to read	49.9 (1.232)	50.6 (2.148)	53.7 (1.931)	45.1* (2.072)	4.63	0.012**	3291
Head's spouse able to write	53.9 (1.312)	54.3 (2.074)	57.7 (2.214)	49.7 (2.336)	3.11	0.048**	3291

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Household Questionnaire.

5.3.3 Income generating activities

The focus of this subsection is to present a clear picture of the income earned by households to sustain themselves and what they do to obtain it.

Household income

We start by presenting overall descriptives of the total household income, the members who work to earn it and sources where it comes from in Table 41. On average over two household members contribute to income. Given that households usually consists of five members, this yields an average ratio of two working members for three dependents. In almost 80% of the cases one of these working members is the household head. This proportion is a little higher for those in the GK treatment group (81%) with respect to the control (75%). The income data in our sample is balanced in any other respect.

Table 41: Household income

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Total HH income	60365 (3460.6)	60365 (4398.7)	56257 (5016.8)	63158 (8049.4)	0.43	0.65	3595
Typical HH yearly income	64746 (4002.1)	64746 (4463.5)	64590 (8111.5)	65833 (7820.8)	0.024	0.98	3595
Nr HH members contributing income	1.69 (0.0278)	1.66 (0.0487)	1.72 (0.0541)	1.70 (0.0395)	0.46	0.63	3595
Head contributing income	0.78 (0.0111)	0.75 (0.0215)	0.81** (0.0174)	0.77 (0.0168)	2.76	0.067*	3595
Income source: Wages agriculture	48.1 (1.459)	49.3 (2.563)	46.3 (2.329)	48.7 (2.653)	0.42	0.66	3595
Income source: Wages non-agriculture	30.2 (1.353)	28.9 (2.240)	28.9 (2.337)	32.9 (2.403)	0.96	0.39	3593
Income source: Business/Farm	34.0 (1.316)	35.5 (2.483)	33.0 (2.123)	33.4 (2.170)	0.31	0.74	3595
Income source: Others	3.01 (0.321)	3.07 (0.632)	3.21 (0.496)	2.74 (0.518)	0.22	0.81	3590

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Household Questionnaire.

Households in our sample earned in the last year on average Rs. 60,365 (~USD

970)²¹. This figure includes two hundred households who report not having received any income (in-kind or in cash) over the last year. Employing the sub-sample who reported positive yearly earnings, back of the envelope calculations indicate that the average household in the sampled area lives on \$2.8 per day. Recalling that each household has on average 5.43 members, it translates into barely \$0.52 per person per day. Even excluding children, i.e. members of less than 15 years of age, from this estimation the daily income per adult amounts to only \$0.76 per person per day, far below the internationally recognised poverty line of US\$ 1.25 per person per day. We can make this comparison more meaningful by adjusting for the purchasing power parity conversion factor. This factor is the number of Rupees required to buy the same amount of goods and services in the local market as a US\$ would buy in the United States. Adjusting for this, the average household member has access to US\$ 1.69 per day, placing our study population slightly above the international poverty line.²²

The *typical* yearly income reported by the respondents was slightly higher, reaching Rs. 64,746 (~USD 1,039) for the whole sample.

Our study households' earnings come to a very large extent from agriculture-related activities, with 48.1% of the sample reporting receiving wages for agricultural labour and 34% deriving income from the farm.²³ Another important source of income are wages from employment outside the agriculture and allied sector, benefitting 30% of the households sampled.

Table 42: Household business

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
HH owns business	31.1 (1.308)	32.4 (2.437)	30.0 (2.108)	30.9 (2.196)	0.28	0.76	3595
Primary activity business: Farm	94.3 (0.937)	95.0 (1.461)	94.1 (1.656)	93.6 (1.774)	0.19	0.83	1119
Primary activity business: Shop/restaurant	4.92 (0.681)	4.99 (1.210)	3.93 (1.017)	5.80 (1.254)	0.69	0.50	1119
Primary activity business: Other	3.49 (0.603)	3.49 (1.006)	3.37 (1.052)	3.59 (1.078)	0.011	0.99	1119
Nr of workers in HH business	5.79 (0.232)	6.20 (0.342)	5.74 (0.435)	5.40 (0.409)	1.16	0.32	1119
Enterprise expenses last year	34717 (2654.6)	34717 (6249.8)	35858 (3128.1)	33344 (3169.6)	0.16	0.85	1118
Enterprise revenue last year	63860 (5172.7)	63860 (11559.6)	60160 (6308.5)	52579* (6487.1)	1.76	0.18	1119

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Household Questionnaire.

Household business

From Table 42 we learn that 31% of households own a business, which for the majority of them is a farm, and for the rest is generally a shop or restaurant. These businesses

²¹Exchange rate are derived from http://www.imf.org/external/np/fin/data/param_rms_mth.aspx

²²The PPP conversion is based on PPP conversion factors available on <http://data.worldbank.org/indicator/PA.NUS.PPPC.RF>.

²³This category also includes other type of businesses. However, for 94% of the households this business is a farm (see Table 42).

employ an average of almost six people under normal circumstances, the majority of them household members, which outnumber non-members two to one. These enterprises provide a good profit margin, averaging almost Rs. 30,000 annually (\sim USD 482), about 47% over the revenue.

5.3.4 Assets

This section provides some insight into the wealth of households in terms of their assets. First we provide an overview of the ownership of assets that can be found in rural households (Table 43). Control and treatment groups do not display any significant difference with respect to asset ownership, with the sole exception of four-wheeled motor vehicles which seem to be more common among households in the GK group. However, given the clear and consistent balance in terms of income and in terms of all of the other assets in the list we are not concerned that this one particular difference will influence the study.

We see in Table 44 that close to half of the households hold some land (other than the land the dwelling is built on, if applicable), have some furniture (typically chairs and tables, or beds), television and some jewellery. Phones and electric appliances are common possessions among households in our sample. Other important assets held by households are cattle (26%) and bicycles (29%).

Those who own land, possess on average 3 and 4.6 acres of non- and agricultural land, respectively, the value of the latter being close to half a million rupees (\sim USD 8030). Land tenure is significantly higher for the GK group than for the control group, albeit only at the 10 percentage point significance level. The value of the dwelling amounts to almost two hundred thousand rupees (\sim USD 3212) with no significant differences across groups²⁴.

5.3.5 Consumption

We next turn to consumption expenditures of the households. The statistics presented in Table 45 correspond to three aggregated consumption categories: food consumption, other non-durable and durable goods. These measures are generated with information collected from a wide range of items. With regard to the household expenditure on food, consumption for 21 different items was employed to come up with this aggregate, which represents the value of annual consumption of all of these food items, including home grown food or items obtained as gifts in addition to goods purchased on the market. The mean value of this aggregate measure in a year is Rs. 63,797 (\sim USD 1,025), excluding alcohol and tobacco, where we make the assumption that the reported expenditures from the last week are representative for the whole year.²⁵

In order to obtain an estimate of total consumption one has to add other spending categories like services and non-food consumables, which represent Rs.1,666 (\sim USD 27)

²⁴Two outliers for the value of the dwelling and eight values for the total value of land owned were replaced by their respective GP median values.

²⁵Outliers in consumption values have been replaced by the sample median .

Table 43: Household asset ownership

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
HH assets: Land	44.5 (1.449)	44.4 (2.498)	43.1 (2.364)	46.1 (2.637)	0.36	0.70	3590
HH assets: Bicycle	28.5 (1.793)	28.0 (3.001)	30.5 (2.925)	26.9 (3.359)	0.36	0.70	3593
HH assets: Motorcycle	16.6 (0.756)	16.2 (1.214)	18.5 (1.433)	15.3 (1.227)	1.47	0.23	3593
HH assets: Four wheeler	1.14 (0.184)	0.73 (0.244)	1.86** (0.382)	0.85 (0.290)	3.26	0.042**	3593
HH assets: Chairs/tables	49.9 (1.830)	49.0 (3.197)	52.2 (3.241)	48.4 (3.020)	0.42	0.66	3593
HH assets: Beds	37.2 (1.797)	35.5 (2.854)	41.4 (3.087)	34.9 (3.303)	1.37	0.26	3593
HH assets: Cupboard	32.2 (1.360)	31.7 (2.116)	33.8 (2.647)	31.0 (2.267)	0.34	0.71	3593
HH assets: Other furniture	7.32 (0.846)	5.98 (1.330)	9.11 (1.629)	6.92 (1.380)	1.13	0.33	3593
HH assets: Refrigerator	6.04 (0.477)	5.74 (0.852)	7.00 (0.877)	5.38 (0.710)	1.07	0.35	3593
HH assets: TV	54.5 (1.473)	56.9 (2.442)	55.9 (2.216)	50.4* (2.845)	1.70	0.19	3593
HH assets: Radio	4.15 (0.462)	4.60 (0.637)	4.22 (0.675)	3.59 (1.036)	0.36	0.70	3593
HH assets: Phone	84.1 (0.767)	85.1 (1.199)	83.2 (1.406)	84.1 (1.363)	0.51	0.60	3593
HH assets: Electric appliances	73.4 (1.292)	73.3 (2.391)	75.8 (1.975)	71.3 (2.252)	1.15	0.32	3593
HH assets: Steel/glass plates	85.5 (1.387)	85.9 (2.421)	86.0 (2.202)	84.6 (2.565)	0.100	0.91	3593
HH assets: Jewellery	57.3 (2.108)	56.8 (3.612)	58.8 (3.646)	56.4 (3.685)	0.13	0.88	3593
HH assets: Cattle	26.5 (1.172)	27.1 (1.927)	26.1 (2.238)	26.3 (1.908)	0.064	0.94	3593
HH assets: Goats/Sheep	4.90 (0.448)	4.12 (0.694)	5.32 (0.768)	5.30 (0.851)	0.88	0.42	3593
HH assets: Chickens	5.26 (0.616)	4.52 (0.853)	5.06 (0.970)	6.24 (1.324)	0.59	0.55	3593
HH assets: Agricultural equipment	14.0 (0.975)	12.2 (1.655)	14.9 (1.712)	15.1 (1.662)	0.95	0.39	3593
HH assets: Other	33.4 (1.227)	31.7 (1.982)	35.2 (2.241)	33.2 (2.117)	0.66	0.52	3593

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Household Questionnaire.

Table 44: Household assets value

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Agricultural land owned by HH - Acres	4.61 (0.142)	4.44 (0.218)	5.06* (0.259)	4.34 (0.249)	2.37	0.098*	1570
Non-agricultural land owned by HH - Acres	3.01 (0.357)	2.37 (0.239)	3.92* (0.876)	3.20 (0.805)	1.79	0.18	107
Value land owned (Rs.)	471665 (7925.6)	471665 (13997.0)	465354 (10532.9)	469825 (16013.9)	0.32	0.72	3595
Value main dwelling (Rs.)	184458 (8138.1)	184458 (12163.3)	197518 (16389.1)	181426 (13296.6)	0.62	0.54	3595
Value representative assets (Rs.)	672351 (11442.0)	672351 (20136.8)	676745 (18188.2)	668455 (20994.5)	0.046	0.95	3595

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Household Questionnaire.

Table 45: Household consumption

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Food consumption (last year)	63797 (1010.2)	63797 (1750.0)	61286* (1494.9)	64956 (1907.4)	1.81	0.17	3595
Nondurable consumption expenditure (past month)	1666 (64.73)	1666 (106.6)	1691 (106.1)	1649 (123.4)	0.039	0.96	3595
Durable consumption expenditure (last year)	26008 (937.6)	26008 (1804.5)	24621 (1470.3)	25889 (1531.1)	0.74	0.48	3595
Total consumption expenditure (last year)	110128 (1775.9)	110128 (3196.1)	106506 (2724.2)	110985 (3186.5)	1.23	0.29	3595

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Household Questionnaire.

monthly for the typical household. Durable goods consumption, in turn, is reported to be in the region of twenty-six thousand rupees annually (\sim USD 418).

Overall, the grand total for the value of household consumption is an estimated Rs. 110,128 (\sim USD 1,769), almost one third of which are food expenditures. Note that the total estimated value of these consumables is higher than the reported annual income of Rs.60,365 that we discussed above. The average household in our dataset, however, produces at home or receives as gifts almost 20% of all food items it consumes.

Just like income and assets, consumption expenditures are nicely balanced across treatment arms, suggesting that the experimental design was successful in randomly allocating households of varying income groups to the different samples.

5.3.6 Savings, credit and insurance

In this section we look at the financial transactions of the household, namely, at the savings at their disposal, their credit history and the insurance policies that they may hold.

Savings

Table 46 shows that 25% of households have savings, averaging thirty-five thousand rupees (\sim USD 562) for those who have them. The vast majority of these households employ formal financial institutions like banks as their vehicle for saving.

Table 46: Household savings

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
HH has savings	24.9 (1.633)	25.0 (2.680)	26.2 (2.741)	23.2 (3.044)	0.27	0.77	3593
Amount savings (Rs.)	35095 (2725.7)	35095 (4398.9)	40461 (5466.9)	29571 (3475.0)	1.47	0.23	893

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Household Questionnaire.

Credit

Households are asked about their credit access and borrowing activities for both small amounts (less than Rs 500) and larger amounts (Rs 500 and above). This distinction was made as small amounts of borrowing are typically non-negligible in rural India, but rarely reported when asking about household debts. We start by discussing responses to small amounts of credit, followed by the larger ones.

From Table 47 we can see that 61% of households state to know a source where they can turn to in case they need to borrow less than Rs. 500.. The main source for this type of credit are friends (72%) and relatives (53%), followed at great distance by moneylenders (15%) and more formal institutions like banks (12%). It is worth noting that almost nobody in our sample mentions microfinance institutions as a viable source of such small amounts. However as we will see shortly, MFIs are also not a prominent source for larger amounts in our study area. Borrowing from friends appears to be less common in the control group than in our treatment groups by a small margin. In any other respect, credit statistics are shown to be balanced across treatments in the sample.

The above table describes just potential borrowing and sources of credit for households, actual borrowing of small amounts was much less common among households than their knowledge of sources may suggest. As we can see in Table 48 only 21% of them took any loan of less than Rs. 500 (\sim USD 8) in the last month. As expected there is a high correlation between known sources and actual lenders, and most of these small loans were taken from friends (68%) and relatives (23%), with any other sources

Table 47: Household credit sources (<Rs.500)

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
HH knows credit source	61.0 (1.341)	61.3 (2.046)	62.6 (2.461)	59.1 (2.427)	0.52	0.60	3593
Sources of borrowing: Bank	11.8 (1.015)	12.3 (2.041)	11.2 (1.717)	11.8 (1.394)	0.087	0.92	2193
Sources of borrowing: Microfinance inst.	0.41 (0.132)	0.13 (0.131)	0.67* (0.287)	0.43 (0.240)	1.78	0.17	2193
Sources of borrowing: SHG	6.29 (1.033)	7.64 (1.930)	5.12 (1.696)	6.07 (1.690)	0.48	0.62	2193
Sources of borrowing: Moneylender	15.3 (1.361)	15.4 (2.292)	16.6 (2.469)	13.9 (2.268)	0.33	0.72	2193
Sources of borrowing: Relative	52.7 (1.880)	52.0 (3.363)	53.1 (3.262)	53.0 (3.116)	0.032	0.97	2193
Sources of borrowing: Friend	71.8 (1.619)	66.7 (2.714)	75.1** (2.846)	73.8* (2.643)	2.74	0.068*	2193
Sources of borrowing: Work	11.4 (1.179)	10.9 (2.144)	11.3 (1.800)	11.8 (2.164)	0.045	0.96	2193

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Household Questionnaire.

playing a very marginal role. The total amount borrowed adding up all loans by those who took any is less than a thousand rupees (\sim USD 16), of which 65% is still owed on average. 80% of the household who borrowed in the last month are still indebted.

Table 48: Household borrowing (<Rs.500)

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
HH taken loan last month	20.8 (1.380)	21.3 (2.597)	17.0 (1.820)	24.3 (2.567)	2.89	0.059*	2193
HH borrowed from: Bank	2.19 (0.659)	1.85 (1.024)	3.97 (1.570)	1.19 (0.824)	1.22	0.30	456
HH borrowed from: Moneylender	7.24 (1.258)	7.41 (2.103)	10.3 (2.910)	4.76 (1.637)	1.51	0.23	456
HH borrowed from: Relative	23.0 (2.139)	22.2 (3.933)	23.8 (3.820)	23.2 (3.348)	0.043	0.96	456
HH borrowed from: Friend	67.8 (2.395)	68.5 (3.898)	61.9 (4.739)	71.4 (3.776)	1.24	0.29	456
Amount borrowed last month	729.8 (67.78)	670.1 (82.72)	669.1 (114.6)	832.7 (140.9)	0.54	0.58	455
HH has debt outstanding	79.6 (3.058)	79.4 (5.492)	80 (5.728)	79.5 (4.734)	0.0034	1.00	451
Amount outstanding debt	470.1 (52.56)	407.1 (58.46)	496.3 (119.5)	511.1 (94.44)	0.55	0.58	451

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Household Questionnaire.

Given that the GK intervention in this study aims at promoting the construction of sanitation systems through the provision of credit, households were asked about their knowledge of credit sources for the specific purposes addressed by the intervention prior to its implementation. About half of the sample was asked about whether they know of any institution that provides loans for housing improvements, whereas the other half of the sample was asked about whether they know of any institution that provides

loans for educational purposes (neither of which are loans that were provided by GK prior to the intervention). The results reveal a general lack of awareness of any sources available for getting a loan with the purpose of making house improvements (only 7% know of any institution for this purpose, which includes toilet construction), or for funding education (less than 5%). Even among those who are aware of any source of funding, microfinance institutions are relatively unheard of, just 10% and 6% of these households know of a microfinance institution that disburses loans for improving the house or furthering education, respectively, within these tiny sub-samples.

Table 49: Knowledge of sources specific purpose

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
HH knows credit source for housing improvements	7.13 (0.705)	6.75 (1.138)	7.13 (1.339)	7.53 (1.181)	0.11	0.89	1780
HH knows MFI for housing loan	10.2 (3.019)	12.2 (6.555)	11.9 (5.022)	6.82 (3.799)	0.44	0.65	127
HH knows credit source for building toilet	43.3 (4.176)	43.9 (5.363)	42.9 (6.203)	43.2 (9.227)	0.0084	0.99	127
HH knows MFI for toilet loan	3.64 (2.485)	5.56 (5.518)	5.56 (4.880)	0 (0)	1.11	0.34	55
HH knows credit source for education	4.47 (0.628)	3.17 (0.991)	4.87 (1.210)	5.46 (1.015)	1.39	0.25	1812
HH knows MFI for education loan	6.17 (2.636)	10 (5.718)	3.45 (3.517)	6.25 (4.316)	0.48	0.62	81

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Household Questionnaire.

With regard to credit sources for large amounts (over Rs. 500), they are reportedly known by only half of the sample. Nevertheless, households tend to turn to similar credit sources as for smaller amounts with only two exceptions of note, both banks and moneylenders are considerably more often mentioned for larger amounts of credit: 23% and 28% of credit aware households would seek funding from these two sources respectively. Microfinance institutions continue to be a negligible source of borrowing in our sample area.

Information was also collected about the credit history during the last year from households whose borrowing exceeded five hundred rupees. According to our findings reported Table 51, 22% of households got indebted at some point during the year, 93% of whom have still outstanding debts. Both the amount borrowed (sum of the three most recent loans) and still owed are close to fifty thousand rupees, suggesting that probably the wealthiest households in the sample are responsible for most of the borrowing recorded. Only a meagre 2% of households took a loan for building a toilet, health expenses and purchase of agricultural inputs are the most frequent reasons for households to get into debt.

Table 50: Household credit sources (>Rs.500)

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
HH knows source to borrow from	52.2 (1.475)	53.7 (2.146)	52.4 (2.606)	50.3 (2.871)	0.44	0.64	3593
Sources of borrowing: Bank	23.1 (1.386)	22.0 (2.394)	23.7 (2.671)	23.9 (2.024)	0.22	0.81	1875
Sources of borrowing: Microfinance inst.	1.60 (0.266)	1.35 (0.408)	1.77 (0.489)	1.70 (0.480)	0.26	0.77	1875
Sources of borrowing: SHG	7.95 (1.189)	10.4 (2.167)	7.09 (1.947)	6.11 (1.946)	1.14	0.32	1875
Sources of borrowing: Moneylender	28.2 (1.703)	28.6 (2.895)	30.3 (2.804)	25.7 (3.094)	0.61	0.55	1874
Sources of borrowing: Relative	57.2 (2.318)	55.8 (4.162)	56.4 (4.117)	59.8 (3.671)	0.31	0.73	1875
Sources of borrowing: Friend	66.0 (1.892)	62.0 (3.111)	68.1 (3.484)	68.4 (3.078)	1.34	0.27	1875
Sources of borrowing: Work	12.7 (1.227)	12.9 (2.145)	12.7 (2.074)	12.6 (2.148)	0.0074	0.99	1875

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Household Questionnaire.

Table 51: Household borrowing (>Rs.500)

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
HH taken loan last year	22.0 (1.491)	22.2 (2.565)	18.8 (2.179)	25.1 (2.866)	1.59	0.21	3592
Amount borrowed last year	45998 (3821.1)	45998 (8191.3)	39561 (4473.8)	47498 (5870.3)	0.90	0.41	792
HH has debt outstanding	93.1 (1.264)	93.8 (1.917)	93.3 (2.085)	92.2 (2.405)	0.14	0.87	792
Amount outstanding debt	49668 (5536.5)	49668 (8291.3)	41653 (6090.2)	53820 (11681.4)	0.71	0.49	792
HH took loan for building toilet	2.02 (0.509)	2.18 (0.873)	2.24 (0.948)	1.70 (0.839)	0.12	0.89	792

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Household Questionnaire.

Insurance

Finally, we look at whether households have insurance and if so, of which type. As we can see in Table 52, 20% of households have taken an insurance policy in the sampled area. Among them, the most common insurance type is life insurance, held by almost 80% of those with a policy. Other types of insurance bought are crop insurance (19%) and health insurance (12%). Of this latter group almost 80% are participants of the Rajiv Gandhi Jeevandayee Arogya Yojana (RGJAY) health insurance scheme sponsored by the state government of Maharashtra; the program's main objective is to improve medical access for both Below Poverty Line (BPL) and Above Poverty Line (APL) families.

Table 52: Household insurance

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
HH has insurance	19.8 (0.968)	17.2 (1.519)	22.0** (1.651)	20.3 (1.779)	2.35	0.100*	3592
HH insurance type: Crop	19.0 (1.686)	19.7 (2.762)	18.5 (2.925)	18.9 (2.991)	0.050	0.95	711
HH insurance type: Weather	1.97 (0.816)	0.47 (0.461)	1.54 (0.893)	3.78 (2.156)	1.55	0.22	711
HH insurance type: Life	79.3 (1.963)	74.2 (4.522)	85** (2.775)	77.7 (2.768)	2.77	0.067*	711
HH insurance type: Health	12.0 (1.404)	14.1 (2.740)	8.08* (2.008)	14.3 (2.457)	2.53	0.084*	711
HH insurance type: Vehicle	7.17 (1.006)	4.69 (1.417)	9.23* (2.015)	7.14 (1.456)	1.82	0.17	711
HH insurance type: Livestock	0.56 (0.272)	0.94 (0.634)	0.38 (0.373)	0.42 (0.407)	0.30	0.74	711
HH insurance type: Other	1.97 (0.610)	2.82 (1.229)	1.93 (1.130)	1.26 (0.724)	0.62	0.54	710
HH has RGJAY	8.02 (1.211)	8.45 (1.966)	6.54 (2.027)	9.24 (2.229)	0.44	0.64	711

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Household Questionnaire.

5.3.7 Shocks

We now turn to negative shocks experienced by the households over the last year. Results are displayed in Table 53.

The most common shock suffered by households was bad harvest, which reportedly affected 35% of the households. It is interesting to put this finding in context with our income figures above, where we show that households had on average a lower income than their typical earnings. It is possible that these bad harvests played a role, given the large proportion of households who earn their livelihoods primarily in the agricultural sector. The occurrence of all other shocks was substantially less frequent, for example, the death of a household member only hit 3% of the sample. We see imbalances across our treatment arms in the occurrence of shocks but note that these imbalances are likely driven by the low occurrence of relevant shocks. Bad harvests on the other hand, which is a shock more likely to occur for households of our sample, is on the other hand balanced across arms.

Table 53: Shocks and risks

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Unexpected job loss	2.15 (0.357)	2.75 (0.772)	1.10* (0.349)	2.57 (0.608)	3.40	0.037**	3586
Robbery/ Theft	0.39 (0.0985)	0.49 (0.185)	0.084** (0.0838)	0.60 (0.205)	4.04	0.020**	3583
Bad harvest	35.1 (1.388)	35.1 (2.525)	34.5 (2.202)	35.6 (2.456)	0.056	0.95	3589
Death HH member	3.01 (0.362)	3.08 (0.689)	2.03 (0.414)	3.93 (0.691)	3.06	0.050*	3586

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Household Questionnaire.

5.3.8 Sanitation

In this section, we provide an overview of the sanitation facilities of the household, the motivation for constructing them or not and the way they are financed.

Table 54: Toilet ownership

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
HH owns toilet	30.5 (1.737)	28.0 (2.788)	35.1* (2.783)	28.5 (3.313)	1.95	0.15	3595
HH's toilet in use	95.5 (0.835)	95.4 (1.699)	96.6 (1.102)	94.3 (1.572)	0.78	0.46	1097

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Household Questionnaire.

Table 54 gives a broad picture of the sanitation situation in our sample area. Less than a third of households (31%) own a toilet, out of which 96% are reported to be currently in use. Toilet ownership at baseline (November 2014 - January 2015) is only slightly higher than the 28% coverage that was reported during the listing survey that took place two months prior to the baseline survey (see Section 5.1). These figures about existing sanitation facilities were validated through direct observation by the data collection team, who observed and confirmed that 29% of households had, indeed, a toilet. This observed measure is highly correlated (90%) with the self-reported measure.

Table 55: Toilet type

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Flush/pour flush to - piped sewer system	2.37 (0.575)	2.88 (1.119)	2.16 (0.967)	2.10 (0.867)	0.17	0.84	1097
Flush/pour flush to -septic tank	24.2 (1.890)	22.5 (3.466)	23.3 (3.381)	27.0 (2.692)	0.67	0.51	1097
Flush/pour flush to - pit latrine	26.0 (2.064)	31.1 (3.447)	23.0 (3.542)	24.3 (3.354)	1.58	0.21	1097
Ventilated improved pit latrine (vip)	16.0 (1.587)	15.0 (2.701)	16.8 (2.650)	16.2 (2.884)	0.12	0.89	1097
Pit latrine with slab	24.6 (2.142)	19.6 (3.922)	27.6 (3.605)	26.1 (3.351)	1.25	0.29	1097
Composting toilet	5.01 (1.220)	7.49 (3.069)	5.04 (1.561)	2.40 (1.172)	1.74	0.18	1097

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Household Questionnaire.

As seen in Table 55 flush or pour flush toilets are the kind most commonly owned by households, accounting for more than half of the existing ones. These toilets are characterised by having a water seal that prevents the passage of flies and odours. A comparatively smaller but still numerous group of households (42%) possess dry-type toilets, namely, improved pit latrines either ventilated (16%) or with a slab (25%).

The location of the toilet (Table 56) is for all of those households reporting to have one within or in the immediate vicinity of the dwelling, with the latter being the most common occurrence (85%). These results are backed by the interviewer's direct

Table 56: Toilet location

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Location toilet: In dwelling	14.1 (1.322)	14.4 (2.530)	13.9 (1.971)	14.1 (2.437)	0.012	0.99	1097
Location toilet: Attached dwelling	42.1 (1.993)	44.7 (3.666)	41.5 (3.058)	40.2 (3.696)	0.39	0.68	1097
Location toilet: In compound	42.8 (1.826)	40.9 (3.252)	42.9 (2.552)	44.4 (3.808)	0.26	0.77	1097
Location toilet: 10 mins walking	0.73 (0.275)	0 (0)	1.44** (0.605)	0.60 (0.419)	3.85	0.024**	1097

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Household Questionnaire.

Table 57: Pit/septic tank characteristics

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Material pit/tank: Brick	72.2 (2.103)	71.2 (3.706)	72.9 (3.069)	72.4 (4.281)	0.064	0.94	1097
Material pit/tank: Stone	2.19 (0.605)	1.44 (0.593)	3.84 (1.408)	0.90 (0.499)	1.97	0.14	1097
Material pit/tank: Cement rings	17.3 (1.816)	18.7 (3.633)	15.6 (2.601)	18.0 (3.327)	0.31	0.74	1097
Material pit/tank: Other permanent materials	7.84 (1.487)	8.65 (2.883)	7.67 (2.471)	7.21 (2.290)	0.077	0.93	1097
Material covering pit/tank: Cement	98.0 (0.550)	98.3 (0.618)	98.3 (1.088)	97.3 (0.970)	0.40	0.67	1097
Material covering pit/tank: Other	1.82 (0.542)	1.44 (0.579)	1.44 (1.067)	2.70 (0.970)	0.66	0.52	1097

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Household Questionnaire.

observations, which yield the same conclusions. The GK group is 1 percentage point more likely than the control group to have to walk for 10 minutes to reach its own toilet. This difference is statistically significant at the 5% level but given the small proportion of households this response option is relevant to we do not consider this of great concern.

The toilet's pit or septic tank is usually made of brick (72% of the cases) with cement rings as a distant second. In contrast, the super structure covering the pit or tank is almost invariably constructed with cement (Table 57).

For an overwhelming majority of households (96%) the pit has never reached its maximum capacity, this finding seems indicative of the relative novelty of the sanitation facilities in the area. For 15 of the 39 households whose pit or tank filled up, this took more than 5 years to occur (see Table 58), however, for a third of the household the pit needed to be emptied after just one to two years. These factual results come in marked contrast with the time that households who have never had a filled pit expect this to happen, the tendency here is to consider that the pit would take longer to become full than what the actual results suggest.

Most toilet owners (79%) in the sample plan to empty the pit when it fills up and continue to use it, yet only a minority of them (16%) would rather opt/be able to use a second pit instead of emptying the full one. Very few others would consider other

Table 58: Pit/tank filled

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Pit/septic tank filled up	3.62 (0.733)	3.77 (1.657)	3.90 (1.001)	3.10 (1.095)	0.16	0.86	1078
Pit/tank filled up in: One year or less	7.69 (4.382)	0 (0)	12.5 (8.113)	10 (9.873)	1.61	0.22	39
Pit/tank filled up in: 1-2 years	33.3 (9.924)	53.8 (14.23)	18.8* (12.68)	30 (18.39)	1.63	0.22	39
Pit/tank filled up in: 2-3 years	7.69 (4.261)	15.4 (10.81)	6.25 (5.846)	0 (0)	1.50	0.24	39
Pit/tank filled up in: 4-5 years	12.8 (5.791)	7.69 (8.039)	18.8 (10.49)	10 (9.873)	0.34	0.71	39
Pit/tank filled up in: More than 5 years	38.5 (9.257)	23.1 (10.01)	43.8 (15.55)	50 (19.05)	1.07	0.36	39
Pit/tank expected to fill up in: One year or less	4.54 (0.763)	5.69 (1.360)	3.49 (1.201)	4.64 (1.398)	0.74	0.48	1058
Pit/tank expected to fill up in: 2-3 years	11.1 (1.225)	14.1 (2.382)	8.73* (1.624)	10.8 (2.331)	1.73	0.18	1058
Pit/tank expected to fill up in: 4-5 years	13.2 (1.333)	14.7 (2.769)	13.2 (2.069)	11.8 (2.000)	0.38	0.69	1058
Pit/tank expected to fill up in: More than 5 years	57.6 (2.325)	49.4 (4.631)	62.6** (3.206)	59.8* (3.517)	2.79	0.065*	1058
Pit/tank expected to fill up in: Don't know	13.6 (1.208)	16.2 (2.265)	12.0 (1.887)	13.0 (1.943)	1.05	0.35	1058
When pit/tank full: Use second pit	16.1 (1.692)	17.6 (3.879)	14.1 (2.184)	17.1 (2.554)	0.52	0.59	1097
When pit/tank full: Empty it	78.5 (1.743)	76.9 (3.752)	81.8 (2.372)	76.0 (2.733)	1.44	0.24	1097
When pit/tank full: Use someone else's toilet	0.36 (0.221)	0.29 (0.283)	0 (0)	0.90 (0.659)	1.45	0.24	1097
When pit/tank full: OD	1.73 (0.500)	2.59 (0.896)	1.68 (1.011)	0.90* (0.484)	1.44	0.24	1097
When pit/tank full: Other	0.18 (0.128)	0 (0)	0.24 (0.236)	0.30 (0.298)	1.02	0.36	1097

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Household Questionnaire.

alternatives, like turning to open defecation.

Table 59: Construction and funding toilet

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Who constructed toilet-Self/Other family member	11.3 (1.134)	13.1 (2.172)	6.97** (1.356)	14.8 (2.050)	6.22	0.0027***	1091
Who constructed toilet-Self/Other family member with help	0.37 (0.181)	0.29 (0.287)	0.24 (0.239)	0.60 (0.421)	0.29	0.75	1091
Who constructed toilet-Hired labour	11.6 (1.453)	7.56 (2.133)	13.2* (2.290)	13.9* (2.801)	2.30	0.10	1091
Who constructed toilet-Mason	92.4 (0.865)	91.6 (1.706)	93.8 (1.116)	91.5 (1.674)	0.90	0.41	1091
Who constructed toilet-Arranged dwelling owner	0.37 (0.181)	0.29 (0.285)	0.48 (0.336)	0.30 (0.299)	0.11	0.90	1091
Who constructed toilet-Government official	1.19 (0.594)	1.16 (0.714)	1.92 (1.381)	0.30 (0.299)	1.18	0.31	1091
Who constructed toilet-NGO workers	0.46 (0.370)	0 (0)	1.20 (0.926)	0 (0)	1.68	0.20	1091
Cost of toilet (Rs.)	26527 (712.4)	26527 (1433.0)	26684 (1044.2)	26610 (1224.2)	0.028	0.97	922
Source of funding toilet: Savings	87.1 (1.791)	86.5 (3.068)	86.3 (3.557)	88.9 (2.076)	0.32	0.73	1097
Source of funding toilet: Loan (formal)	0.27 (0.157)	0.29 (0.283)	0.48 (0.338)	0 (0)	1.52	0.22	1097
Source of funding toilet: Loan (informal)	2.01 (0.484)	2.88 (1.029)	0.72* (0.529)	2.70 (0.915)	2.85	0.062*	1097
Source of funding toilet: Subsidy	11.8 (1.981)	11.5 (2.562)	12.9 (4.329)	10.5 (2.373)	0.13	0.88	1097
Source of funding toilet: Transfers	4.74 (0.693)	5.19 (1.230)	3.84 (1.129)	5.41 (1.225)	0.53	0.59	1097
Source of funding toilet: Other	0.27 (0.156)	0.58 (0.409)	0 (0)	0.30 (0.287)	1.54	0.22	1097

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Household Questionnaire.

In the vast majority of cases (92%) the toilet was built by masons hired for this specific purpose, only 11% of households took the task upon themselves or hired un-specialised labour to complete the job as displayed in Table 59. The average cost of the toilet was reportedly of approximately Rs. 26,000 (~USD 417), which came predominantly from the household's own savings. This amount is significantly higher than either the sanitation loan offered by GK (Rs. 15,000) or the subsidy provided by the government (Rs. 12,000). Few households had access or made use of external funding to cover the costs of the toilet and only 12% benefitted from any type of subsidy. These results are suggestive that only those with already the means to have a toilet built undertook its construction. Supporting this hypothesis are statistics presented in [Raman and Tremolet \[2010\]](#), which reports on toilet construction costs in three districts of Maharashtra (Chandrapur, Kolhapur and Nashik). The costs reported for a typical toilet of a household classified as APL is reported at on average USD332, which is the equivalent of USD373 in December 2014 USD value (the time of our baseline survey). This average hides some variation. The average reported costs in Kolhapur for example is as high as USD 434 for an average APL toilet, slightly above the average reported by our study households.

Table 60: Motivation building toilet

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
More convenience	65.0 (1.798)	65.1 (2.388)	63.3 (3.208)	67.0 (3.472)	0.30	0.74	1097
Female members in the household wanted one	78.0 (1.728)	76.1 (3.536)	79.6 (2.684)	78.1 (2.671)	0.32	0.73	1097
Health improvement	51.5 (2.051)	47.8 (3.476)	55.4 (3.578)	50.5 (3.346)	1.18	0.31	1097
People in village had one	9.75 (1.045)	9.80 (1.800)	10.1 (1.683)	9.31 (1.973)	0.043	0.96	1097
Imposed by someone in village	3.46 (0.706)	4.03 (1.127)	3.60 (1.349)	2.70 (1.071)	0.38	0.68	1097
Usage of toilets elsewhere	2.10 (0.560)	3.17 (1.238)	1.44 (0.801)	1.80 (0.828)	0.70	0.50	1097
Physical problems with OD	13.1 (1.505)	12.1 (2.455)	14.1 (2.476)	12.9 (2.861)	0.17	0.84	1097
Difficulties in finding alternatives	2.73 (0.651)	2.31 (0.918)	2.16 (0.726)	3.90 (1.614)	0.49	0.61	1097
So not embarrassed with guests	3.92 (0.712)	4.32 (1.035)	4.32 (1.381)	3.00 (1.133)	0.44	0.65	1097
Status in the village	7.29 (1.123)	6.05 (1.724)	6.71 (1.716)	9.31 (2.411)	0.62	0.54	1097
Better hygiene	17.3 (1.730)	16.1 (2.837)	17.3 (2.570)	18.6 (3.671)	0.14	0.87	1097
Greater safety	12.9 (1.472)	12.4 (2.585)	12.7 (2.417)	13.5 (2.678)	0.048	0.95	1097
Financial support from government	4.83 (1.001)	2.88 (0.989)	5.76 (1.777)	5.71 (2.087)	1.45	0.24	1097
Other financial support	0.46 (0.200)	0.58 (0.401)	0.24 (0.241)	0.60 (0.404)	0.44	0.65	1097
Owner/Landlord decided	1.00 (0.312)	0.86 (0.495)	0.72 (0.524)	1.50 (0.578)	0.56	0.57	1097
Government official decided	0.64 (0.386)	0.29 (0.284)	0.96 (0.912)	0.60 (0.419)	0.37	0.69	1097
NGO employee or other decided	0.36 (0.282)	0 (0)	0.96 (0.706)	0 (0)	1.84	0.18	1097
Other	0.27 (0.157)	0.29 (0.286)	0.24 (0.238)	0.30 (0.298)	0.015	0.98	1097

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Household Questionnaire.

From Table 60 we can see that 78% of respondents cite the influence of female household members as one of the main motivations leading to the decision of constructing a toilet. This is closely followed by the convenience of having access readily to a sanitation facility when it belongs to the household, mentioned by 65% of the sub-sample with a toilet. A third reason brought forward by half of the respondents is the perceived improvement in health that a toilet could bring. Other reasons stated are general issues (e.g. safety or physical problems) with other defecation alternatives (25%), the wish of better hygiene (17%), or social repercussions (11%) appear to be important as well but are less often cited.

Table 61 provides information on households that used to have a toilet but to not anymore, which is less the 1% of our study households. The main reason why they do not have it anymore is that is became unusable.

Table 61: Household does not own toilet anymore

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
HH had toilet	0.80 (0.176)	1.01 (0.304)	0.91 (0.362)	0.48 (0.233)	1.13	0.33	2498
Reason no longer toilet: Became unusable	60 (10.68)	88.9 (10.76)	42.9** (16.24)	25** (22.24)	4.43	0.027**	20
Reason no longer toilet: Never used it	15 (8.275)	0 (0)	14.3 (13.91)	50* (25.69)	2.17	0.14	20
Reason no longer toilet: Was in previous home	5 (5.020)	0 (0)	14.3 (13.91)	0 (0)	0.94	0.34	20
Reason no longer toilet: Other	20 (8.718)	11.1 (10.76)	28.6 (14.83)	25 (22.24)	0.45	0.64	20

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Household Questionnaire.

When asking households without a toilet about the reasons for the same we find that the main obstacle for building a toilet stated by those households is lack of financial resources (83%). This can be seen in Table 62. Limited space availability comes second but it is only argued by 17% of the households, and other reasons seem not to have much weight in preventing households from building a sanitation facility. These findings are supportive to the rationale behind the GK intervention, which aims to relax the households financial constraints and to lead to a sizeable increase in toilet adoption.

Table 62: Reasons for not owning a toilet

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
No need	4.36 (0.541)	3.37 (0.764)	4.68 (0.850)	5.14 (1.126)	1.10	0.34	2498
I cannot afford it	83.1 (1.100)	83.1 (1.795)	83.6 (1.896)	82.6 (2.019)	0.077	0.93	2498
I find it too expensive	12.6 (1.000)	14.4 (1.791)	12.5 (1.809)	10.9 (1.524)	1.10	0.33	2498
No space	17.1 (1.133)	17.4 (1.728)	17.3 (2.041)	16.7 (2.120)	0.033	0.97	2497
Had not the time to construct one	11.6 (0.952)	11.1 (1.530)	12.3 (1.604)	11.6 (1.793)	0.15	0.86	2498
Toilet shouldn't be close to house	0.76 (0.199)	0.67 (0.377)	1.04 (0.334)	0.60 (0.307)	0.52	0.60	2498
Never thought about it	8.89 (0.872)	9.54 (1.646)	8.57 (1.371)	8.48 (1.444)	0.14	0.87	2498
Religious and cultural reasons	0.080 (0.0562)	0 (0)	0.26 (0.178)	0 (0)	2.14	0.15	2498
Not everybody in HH wants it	0.44 (0.127)	0.22 (0.157)	0.39 (0.217)	0.72 (0.264)	1.29	0.28	2498
No one else I know has one	1.04 (0.231)	1.12 (0.439)	0.78 (0.340)	1.19 (0.398)	0.37	0.69	2498
Open defecation is seen as better	0.68 (0.165)	0.45 (0.214)	0.78 (0.351)	0.84 (0.294)	0.69	0.50	2498
Lack of info what toilet to build	1.68 (0.343)	1.01 (0.370)	2.08 (0.634)	2.03 (0.727)	1.50	0.23	2498
I dont know how to get a toilet	2.64 (0.480)	2.02 (0.813)	2.34 (0.792)	3.58 (0.866)	0.95	0.39	2498
Stay in rented house	0.64 (0.219)	0.45 (0.269)	0.26 (0.178)	1.19 (0.552)	1.35	0.26	2498
Can use neighbour's	0.20 (0.0886)	0.34 (0.191)	0.13 (0.129)	0.12 (0.119)	0.51	0.60	2498
Other	2.00 (0.392)	1.23 (0.436)	1.95 (0.748)	2.87* (0.804)	1.67	0.19	2498

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Household Questionnaire.

Conforming with the above results, 62% claim that they would not be able to cover the cost of building a toilet even if they wanted to. Table 63 provides information on the remaining 38% of households. These households state a relatively wide set of potential sources for funds for construction of a toilet: 40% state they would take a formal loan, 43% an informal one, 48% report they could get subsidy, and 41% report they could use transfers from family or friends. Lastly, the majority of the households without a toilet (55%) would be willing to borrow for covering the expenditures of constructing one, another important result justifying the approach of the present intervention.

Table 63: Potential for building toilet

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
HH able to fund toilet	37.7 (1.486)	36.1 (2.127)	36.7 (2.942)	40.1 (2.657)	0.73	0.48	2486
Formal loan for building toilet	39.4 (1.783)	39.7 (2.863)	39.4 (3.127)	38.9 (3.271)	0.016	0.98	2170
Informal loan for building toilet	43.2 (1.991)	45.7 (3.105)	42.1 (3.062)	41.6 (3.994)	0.47	0.63	2191
Subsidy for building toilet	48.3 (2.321)	51.7 (4.182)	48.9 (3.356)	44.2 (4.225)	0.81	0.45	2178
Transfers for building toilet	41.4 (2.267)	40.7 (3.777)	42.8 (4.198)	40.7 (3.828)	0.092	0.91	2350
HH would borrow to build toilet	54.6 (1.386)	56.7 (2.205)	52.1 (2.679)	54.7 (2.299)	0.88	0.42	2476

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Household Questionnaire.

5.3.9 Water

The survey also collected information on the household access to water for drinking, cooking and sanitation purposes. A summary of it is presented below. As it can be inferred from the hypothesis tests carried out, treatment and control groups do not differ in any respect in terms of their access to water.

From the figures in Table 64 we learn that the main sources of water for cooking and drinking are piped water into the dwelling (31%), public tap or pipe (25%), tubewell or borewell (23%), handpump (21%) and dug well (15%). This multiplicity of sources is due to the fact that not all houses have water all year round from the same source (for example, 21% of household rely on a different source for the dry season) or at least cannot get enough water from it. Consequently, people in the area need to turn to several sources to ensure a continuous supply of water in the house.

Only 38% of the household have access to water in their own dwelling. For the rest, the source is located an average of 142 metres from the household residence, taking household members about 20 minutes to reach the source, collect water and return home. This action needs to be performed on average eleven times a day to maintain the supply of water that the household requires. Very likely for this reason and other problems related to the difficulties in accessing water, a large proportion of households (63%) store it at home as shown in Table 65.

An equally large number of households pay for drinking and cooking water both in the wet and dry seasons (about 65% in both cases). In spite of almost 80% of the household purifying the water for own consumption, this is done in a rudimentary manner by filtering it with a cloth, a method that does not eliminate parasites and bacteria from the water.

As seen in Table 66, the water used for sanitation needs comes from very similar sources as the consumed water, however, collecting it is much more laborious. Keeping the household water supply takes considerably more trips and longer time, especially, during the dry season when household members need to spend close to an hour in the process.

Table 64: Water cooking and drinking

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Water source: Piped water into dwelling	30.8 (2.345)	31.7 (4.385)	31.6 (4.069)	29.0 (3.638)	0.16	0.85	3595
Water source: Public tap/stand pipe	25.1 (1.912)	26.4 (2.946)	25.7 (3.903)	23.0 (2.984)	0.36	0.70	3595
Water source: Tubewell or borewell	22.5 (1.715)	20.0 (2.500)	22.3 (2.658)	25.4 (3.622)	0.75	0.47	3595
Water source: Handpump	21.3 (2.091)	21.5 (3.907)	19.3 (3.231)	23.1 (3.639)	0.31	0.73	3595
Water source: Protected dug well	14.7 (1.591)	13.5 (2.846)	13.8 (2.607)	16.8 (2.773)	0.42	0.66	3595
HH has water all year from a source	82.6 (1.677)	83.0 (2.908)	82.3 (2.761)	82.4 (3.040)	0.020	0.98	3594
HH has water in own dwelling/yard	38.3 (2.303)	38.4 (4.343)	38.3 (3.901)	38.3 (3.643)	0.00013	1.00	3594
Distance water source (mtrs)	141.7 (7.790)	140.8 (12.28)	129.6 (12.08)	154.4 (15.53)	0.80	0.45	2729
Distance water source (mins)	20.5 (0.605)	21.8 (1.192)	20.6 (0.906)	19.0* (0.952)	1.80	0.17	2729
Trips per day water source	11.6 (0.318)	12.2 (0.532)	11.5 (0.580)	10.9* (0.515)	1.44	0.24	2729
Other water source dry season	20.9 (1.266)	18.3 (2.009)	23.4* (2.257)	21.1 (2.236)	1.45	0.24	3594

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Household Questionnaire.

5.3.10 Health care utilisation

The final section focussing on the household as a whole concerns health care utilisation by all its members.

Our statistics in Table 67 indicate that 17% of the households had at least one member receiving medical attention in the last month which did not require hospitalisation. The mean number of times this happened within this period was 1.46. The main cause for visiting a health facility or being visited by a health worker the last time were digestive problems (38%), followed by respiratory problems (31%), other undefined types of pain or medical reasons account for most of the remaining causes.

The average total amount incurred by households in need of this kind of health care during the last month²⁶ is of Rs. 4,444 (~USD 71), a large cost given the levels of income seen earlier.

²⁶The survey enquired about the last three visits in the past month and only three households reported as many visits, thus, the above can be readily understood as a measure of total cost of outpatient care for our sampled households in a given month.

Table 65: Purifying and paying for water

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
HH stores water	63.2 (1.544)	62.6 (2.803)	66.0 (2.304)	61.0 (2.812)	1.03	0.36	3594
HH pays for water wet season	65.8 (1.710)	63.5 (3.256)	66.2 (2.786)	67.9 (2.736)	0.53	0.59	3592
HH pays for water dry season	64.6 (1.671)	62.8 (3.213)	64.9 (2.649)	66.2 (2.726)	0.34	0.72	3593
HH purifies water	78.4 (1.561)	76.1 (2.784)	82.0* (2.151)	77.4 (3.018)	1.64	0.20	3595
Purification method: Filtering w/ cloth	95.3 (0.578)	94.7 (1.058)	95.1 (1.109)	96.2 (0.750)	0.85	0.43	2820
Purification method: Water filter	4.50 (0.599)	4.46 (1.107)	5.24 (1.121)	3.76 (0.819)	0.58	0.56	2820

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Household Questionnaire.

Table 67: Outpatient medical care

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Medical care received	16.7 (1.099)	15.7 (1.921)	16.6 (1.965)	17.9 (1.797)	0.33	0.72	3574
Nr of visits HH	1.46 (0.0856)	1.53 (0.150)	1.46 (0.181)	1.40 (0.0937)	0.29	0.75	392
Reason med. care: Digestive Problems	38.0 (2.763)	44.8 (4.950)	35.3 (4.329)	34.1 (4.756)	1.46	0.24	405
Reason med. care: Respiratory problems	31.4 (2.801)	36.6 (4.842)	28.6 (4.109)	29.0 (5.300)	0.90	0.41	405
Reason med. care: Preventive medical exam/check-up	9.63 (1.543)	10.4 (2.683)	10.5 (2.856)	7.97 (2.380)	0.33	0.72	405
Reason med. care: Skin problems	3.46 (0.836)	3.73 (1.579)	3.01 (1.347)	3.62 (1.419)	0.076	0.93	405
Reason med. care: Other acute pain	12.3 (1.652)	12.7 (3.176)	10.5 (2.436)	13.8 (2.897)	0.39	0.68	405
Reason med. care: Other medical problem	21.0 (2.605)	14.9 (4.000)	22.6 (5.676)	25.4* (3.543)	1.94	0.15	405
Cost of outpatient care	4444 (729.3)	4444 (1897.8)	3843 (819.3)	4513 (727.8)	0.26	0.77	403

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Household Questionnaire.

In the last year 25% of households in our sample sought inpatient medical care (Table 68), although this was a relatively infrequent event for most of them, with the number of admissions averaging little above one. The main reason for being admitted into a hospital, clinic or other health centre was some type of illness, typically affecting the digestive system. The incidence of digestive problems seems to be higher for the GK+ND group at a 10 % significance level, but this is the only case of unbalancedness among the health care utilisation data. As expected, the cost of having household members hospitalised is markedly higher than receiving outpatient care and amounts to more than fifteen thousand rupees (~USD 240), a serious blow to the economy of our average household in our study area.

Table 66: Water for sanitation

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Dry season: Piped water into dwelling	23.2 (1.976)	24.3 (3.968)	22.3 (3.164)	23.1 (2.977)	0.082	0.92	3594
Dry season: Public tap/stand pipe	18.7 (1.617)	19.7 (2.468)	20.7 (3.313)	15.6 (2.473)	1.05	0.35	3594
Dry season: Tubewell or borewell	20.4 (1.527)	19.1 (2.225)	20.3 (2.534)	22.0 (3.120)	0.29	0.75	3594
Dry season: Handpump	18.3 (1.800)	18.7 (3.159)	16.1 (2.659)	20 (3.453)	0.44	0.64	3594
Dry season: Protected dug well	15.7 (1.505)	15.6 (2.669)	16.4 (2.609)	15.3 (2.526)	0.045	0.96	3594
Wet season: Piped water into dwelling	31.7 (2.253)	32.6 (4.275)	30.4 (3.838)	31.9 (3.493)	0.078	0.92	3594
Wet season: Public tap/stand pipe	21.6 (1.642)	23.3 (2.647)	22.5 (3.275)	18.8 (2.494)	0.87	0.42	3594
Wet season: Tubewell or borewell	18.4 (1.486)	17.0 (2.190)	18.5 (2.346)	19.7 (3.114)	0.29	0.75	3594
Wet season: Handpump	15.8 (1.731)	15.8 (3.217)	15.1 (2.732)	16.4 (3.000)	0.053	0.95	3594
Wet season: Protected dug well	10.3 (1.258)	9.45 (2.206)	10.3 (2.156)	11.2 (2.161)	0.16	0.85	3594
Distance (mins): Dry season	55.9 (4.360)	57.8 (8.037)	53.7 (6.607)	56.1 (7.874)	0.078	0.92	3594
Distance (mins): Wet season	38.5 (2.875)	38.7 (4.877)	35.9 (4.721)	41.1 (5.291)	0.28	0.76	3594
Trips per day: Dry season	15.5 (0.423)	16.3 (0.709)	14.9 (0.738)	15.2 (0.730)	0.98	0.38	3073
Trips per day: Wet season	12.4 (0.323)	13.2 (0.538)	12.0 (0.598)	12.0 (0.517)	1.56	0.21	2873

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Household Questionnaire.

Table 68: Inpatient medical care

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Hospitalisation HH member	24.5 (1.051)	24.2 (1.783)	22.6 (1.583)	26.7 (2.010)	1.28	0.28	3577
Nr of admissions HH	1.21 (0.0315)	1.22 (0.0458)	1.21 (0.0764)	1.21 (0.0408)	0.048	0.95	877
Reason hosp.: Illness (digestive problems)	32.7 (1.721)	30.1 (3.176)	29.7 (2.807)	37.8* (2.711)	2.68	0.073*	877
Reason hosp.: Illness (other)	39.6 (1.949)	41.8 (3.504)	41.0 (3.034)	36.2 (3.442)	0.78	0.46	877
Reason hosp.: Accident	3.42 (0.571)	3.68 (0.987)	3.38 (1.038)	3.21 (0.945)	0.061	0.94	877
Reason hosp.: Physical aggression (violence)	3.99 (0.750)	3.34 (0.955)	4.51 (1.199)	4.17 (1.593)	0.31	0.73	877
Reason hosp.: Surgery	7.07 (1.004)	6.69 (1.412)	4.14 (1.183)	9.94 (2.166)	3.00	0.054*	877
Reason hosp.: Medical analysis or studies	1.14 (0.382)	0.33 (0.324)	1.88 (0.957)	1.28 (0.616)	1.82	0.17	877
Reason hosp.: Other	7.18 (1.029)	7.69 (1.663)	6.77 (1.784)	7.05 (1.864)	0.076	0.93	877
Cost of inpatient care	15974 (1347.8)	15974 (2918.8)	12619* (1556.6)	15807 (1943.7)	2.20	0.11	877

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Household Questionnaire.

During the survey participants without a toilet were asked to predict qualitatively the change in their health expenditures under the hypothetical scenario that they owned a toilet (Table 69). Whereas most households consider that they would see a reduction in their health expenditures (45%) as a result of having a sanitation facility available, many believe that they would not appreciate any variation in their medical care costs (25%), or that these may even rise (15%).

Table 69: Hypothetical health costs

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
We would spend more	14.6 (1.562)	15.4 (2.911)	16.0 (3.036)	12.5 (2.051)	0.58	0.56	2497
We would spend less	44.5 (2.388)	43.4 (4.194)	45.8 (4.519)	44.6 (3.691)	0.072	0.93	2497
We would spend the same	25.2 (1.102)	24.4 (1.808)	24.4 (2.019)	26.8 (1.901)	0.52	0.60	2497

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Household Questionnaire.

5.4 Women versus men in study sample (individual woman and man survey)

In this section we look at and compare information collected from an individual adult woman and man of the household, who were either the mother/father of a child less than 6 years old (preferably the youngest child), or the head/spouse of head of the household. We refer to Figure 12 in Section 4.3.4 for more details on the respondent selection process.

Several questions (not all) were asked of both man and woman in the individual surveys, which allows for a comparison. However, if the respondent of the individual man questionnaire was the same as for the household survey, several sections of the man questionnaire were skipped to avoid survey fatigue. Where possible, responses from the household survey were used in order to compare the male responses to the female responses, but this was not always possible. As a result, the sample size for male respondent is considerably reduced for some variables.

Given that the main use of these individual data will involve making intra-household comparisons, we will present the baseline results of the sample restricted to those households for which we have data at all levels: household level, individual man and woman level. This reduces the individual survey sample size to 3,434 respondents (a 5% drop from the total).

5.4.1 Background and status

We first look at the background and the status of the interviewed woman and men in the household presenting information about their age, education and that of their parents (if available) in Table 70 for women and Table 71 for men.

Table 70: Background and status women

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Age of respondent	30.7 (0.206)	30.1 (0.363)	30.9* (0.288)	31.0* (0.394)	1.96	0.14	3433
Years of schooling	5.90 (0.119)	6.00 (0.192)	6.16 (0.196)	5.54 (0.214)	2.40	0.095*	3297
Years of schooling father	2.35 (0.0874)	2.36 (0.115)	2.31 (0.164)	2.39 (0.175)	0.052	0.95	3059
Years of schooling mother	0.71 (0.0456)	0.65 (0.0550)	0.79 (0.0961)	0.69 (0.0815)	0.82	0.44	3015
Living in village all her life	8.07 (0.567)	7.69 (0.895)	7.91 (0.886)	8.61 (1.147)	0.21	0.81	3434

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Woman Questionnaire.

Female respondents, presumably with young children, are little over 30 years of age and have received around 6 years of formal education. The parents of the interviewed women tend to be significantly less educated than their daughter, especially the mothers, who are mostly illiterate (86%) and received on average less than one year of schooling. The majority of women come from a different village than the one where they currently

Table 71: Background and status men

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Age of respondent	43.4 (0.306)	43.0 (0.569)	43.6 (0.475)	43.5 (0.536)	0.29	0.75	3430
Years of schooling	7.66 (0.122)	7.70 (0.189)	8.02 (0.179)	7.24 (0.244)	3.31	0.040**	3341
Years of schooling father	3.37 (0.179)	3.20 (0.317)	3.76 (0.324)	3.18 (0.271)	1.11	0.33	659
Years of schooling mother	1.14 (0.126)	1.20 (0.221)	1.18 (0.199)	1.02 (0.230)	0.20	0.82	633
Living in village all his life	97.6 (0.587)	96.2 (1.323)	99.1** (0.595)	97.5 (0.953)	2.54	0.084*	703

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Man Questionnaire.

live, most likely as a result of moving into their husband's home after getting married. In line, almost half of the women in our sample live with their in-laws, whereas little more than 1% share a roof with their own parents (Table 72).

Table 72: General information female respondent

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Married	99.1 (0.331)	99.2 (0.602)	98.3 (0.754)	99.6 (0.169)	1.60	0.21	3434
Age when married	17.7 (0.0502)	17.7 (0.0784)	17.8 (0.0893)	17.7 (0.0909)	1.03	0.36	3366
Arranged marriage	98.9 (0.261)	98.4 (0.458)	99.3 (0.290)	99.0 (0.556)	1.19	0.31	3402
Sanitation facilities mattered	75.0 (2.256)	73.0 (4.119)	75.6 (3.595)	76.5 (3.958)	0.19	0.82	3402
Dowry paid	72.9 (1.253)	76.0 (1.816)	70.8 (2.700)	71.8 (1.800)	1.91	0.15	3391
Living with in-laws	45.1 (1.213)	46.2 (2.178)	44.1 (2.193)	45.0 (1.903)	0.24	0.79	3434
Living with parents	1.11 (0.183)	1.11 (0.351)	0.70 (0.253)	1.51 (0.323)	1.96	0.15	3434

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Woman Questionnaire.

Male respondents are considerably older than their female counterparts with almost 40 years of age and are also significantly more educated, having little less than 8 years of schooling. To some extent, this higher education level also extends to their parents, although illiteracy is still predominant among them. Unlike our female respondents, almost the entirety of our male sample has always lived in the same village where they were born. This is especially true for members of the GK treatment group, who have a significantly higher proportion of permanent residents >

As expected given the selection criteria for the respondents to the women questionnaire, 99% of respondents are married. Their marriages took place when they were young, about 18 years old, in an arranged ceremony (99%), which typically involved

the payment of a dowry (73%). A large proportion of women (75%) reported that, when they were looking for a husband, it mattered whether their husband could offer them a home with sanitation facilities.

5.4.2 Women empowerment

We next describe how the most important decisions in the household are taken and to what extent women are involved in these matters. In Table 73 we show the average proportion of women who have a say in some key aspects of the household life as per their own perspective. 67% of our female respondents report being able to decide by themselves or jointly with their husbands about having another child, 68% can influence whether the child goes to the doctor and 61% take part in deciding whether to send the children to school. With regard to monetary matters and employment, the women in our sample seem to have less power to decide, and approximately only half of them report involvement in deciding whether they work, what they do with the money they earn and whether the household should make a significant household purchase.

Table 73: Women’s involvement decision making in the household: Women’s perspective

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Having another child	66.7 (1.745)	69.3 (2.669)	61.9* (3.487)	68.7 (2.717)	1.64	0.20	3434
Taking children to doctor	68.2 (1.164)	68.4 (1.781)	67.5 (2.340)	68.7 (1.893)	0.078	0.93	3434
Children schooling	61.2 (1.340)	61.9 (2.019)	59.8 (2.793)	62.0 (2.063)	0.22	0.80	3434
Money earned	54.4 (1.329)	56.8 (1.882)	52.9 (2.504)	53.4 (2.426)	1.05	0.35	3434
Buying large HH item	56.9 (1.347)	57.4 (2.097)	57.1 (2.496)	56.2 (2.392)	0.067	0.93	3434
Working	52.6 (1.426)	53.4 (2.254)	51.4 (2.789)	52.8 (2.325)	0.16	0.85	3434

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Woman Questionnaire.

It is interesting to compare women’s reported power of their degree of intra-household decision making with what their husband’s give them credit for. Such information is shown in Table 74. Men affirm, to a greater extent than their own spouses (82% versus 67%), that the latter’s opinion count when it comes to deciding about adding another member to the family. This trend is however reversed for the rest of the decision areas. According to the results, 61% of men think that their wives’ views regarding children medical needs matter; Only 23% and 32% of men report that their spouse has a any say on whether they should work and what to do with their earnings, respectively, and 34% consult their spouses when considering to buy an expensive household item.

Table 74: Women's involvement decision making in the household: Men's perspective

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Having another child	81.5 (1.822)	79.1 (3.669)	81.9 (2.774)	83.5 (2.861)	0.46	0.63	703
Taking children to doctor	61.5 (2.459)	58.6 (3.990)	65.6 (4.345)	60.3 (4.350)	0.75	0.47	703
Children schooling	50.4 (2.685)	48.1 (4.231)	56.4 (4.375)	46.8 (5.040)	1.32	0.27	703
Money earned	32.4 (2.440)	27.2 (3.693)	38.8** (4.413)	31.6 (4.335)	2.02	0.14	703
Buying large HH item	33.7 (2.414)	33.5 (3.471)	36.1 (4.476)	31.6 (4.504)	0.25	0.78	703
Working	23.5 (2.613)	21.8 (3.757)	25.6 (5.100)	23.2 (4.653)	0.18	0.84	703

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Man Questionnaire.

Lastly, Table 75 presents an overview of different indicators attempting to capture the extent of women's degree of autonomy. Most women (about 80%) feel free to visit neighbours, the temple or fetch water on their own. The majority of them (66%) also feel permitted to visit unaccompanied friends and relatives within the village where they live, but only 45% feel permitted to do so outside the village. Visiting the local market or going to the doctor unaccompanied is only allowed to only 44% and 47% of women, respectively. Their control over money is also rather limited, with just about half of the sample being able to decide about buying basic personal items (e.g. clothes and medicines) and food by themselves. Whereas 15% of women in the sampled area are part of a lending group, only 59% of them claim being able to participate in the decision of borrowing from the group.

Table 75: Women's autonomy

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Go alone: local market	44.2 (1.630)	47.5 (2.575)	43.9 (2.757)	41.1 (3.040)	1.33	0.27	3317
Go alone: Health centre/doctor	46.9 (1.895)	50.7 (3.037)	46.6 (3.233)	43.3 (3.461)	1.28	0.28	3363
Go alone: Visit neighbours	79.4 (1.373)	81.8 (2.126)	77.4 (2.729)	78.8 (2.200)	0.95	0.39	3394
Go alone: Visit friends/relatives in village	66.2 (1.705)	69.9 (2.451)	64.4 (2.841)	64.2 (3.438)	1.42	0.24	3388
Go alone: Visit friends/relatives out village	44.6 (1.950)	47.4 (3.317)	45.4 (3.301)	40.9 (3.425)	0.98	0.38	3376
Go alone: The Shrine / Mosque / Temple	76.2 (1.294)	78.3 (1.956)	74.0 (2.576)	76.2 (2.121)	0.90	0.41	3390
Go alone: Collect water	82.9 (1.232)	84.2 (1.962)	81.8 (2.595)	82.8 (1.733)	0.30	0.74	3328
Control money: Fruits or Vegetables	50.7 (1.885)	52.7 (3.158)	49.0 (3.325)	50.5 (3.279)	0.34	0.71	3226
Control money: Other food items	47.2 (1.729)	48.8 (2.898)	45.9 (3.074)	46.9 (2.990)	0.25	0.78	3232
Control money: Clothes for yourself	51.9 (2.134)	54.1 (3.711)	50.7 (3.652)	50.9 (3.685)	0.27	0.76	3262
Control money: Medicine for yourself	49.9 (2.368)	51.6 (4.192)	50.0 (3.899)	48.2 (4.166)	0.16	0.85	3260
Control money: Toiletries for yourself	62.0 (2.456)	64.8 (4.416)	60.8 (4.234)	60.2 (4.040)	0.33	0.72	3270
Control money: Clothes and medicine for your kids	45.5 (2.101)	47.3 (3.754)	44.3 (3.277)	44.9 (3.822)	0.19	0.83	3243
Member of lending group	14.8 (1.040)	14.5 (1.635)	14.1 (1.693)	15.6 (2.055)	0.16	0.85	3434
Decision loan lending group	59.4 (3.022)	58.8 (4.595)	64.6 (4.669)	55.1 (6.039)	0.84	0.43	507

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Woman Questionnaire.

5.4.3 Credit and savings

In this section we look at the individual credit history and savings of the man and woman interviewed. Given the comprehensive overview of the household finances provided in the corresponding subsections of the household part of this report, we only enquire about whether each member has savings and borrowed money last year. The answer to these questions are presented in Table 76. 14% of women have savings that they consider their own, not at the disposal of the rest of the family. By contrast, 26% of the male respondents have accumulated some money that they they consider their own. Whereas we learnt from the household level survey results that 20% of the households reported to have taken a loan, here we observe that 6% of respondents - both man and woman - report to have taken a loan for themselves in the last 12 months. The discrepancy with the household survey is likely driven by the fact that the individual questionnaire is not necessarily administered to the household head, who is most likely in the household to take credit, loans that would count towards the reported borrowing percentage in the household survey.

Table 76: Savings and credit man and women

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Woman: Has savings	14.0 (0.987)	13.7 (1.667)	15.5 (1.808)	12.9 (1.615)	0.58	0.56	3434
Woman: Taken loan last year	6.29 (0.619)	5.90 (1.202)	6.94 (1.065)	6.04 (0.907)	0.28	0.76	3434
Man: Has savings	26.0 (2.080)	25.5 (3.558)	26.0 (3.483)	26.6 (3.749)	0.021	0.98	703
Man: Taken loan last year	6.83 (1.101)	6.69 (2.163)	7.49 (1.857)	6.33 (1.665)	0.11	0.90	703

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Woman Questionnaire.

5.4.4 Social networks, group membership and political activity

As part of the survey administered to male respondents, we enquired about their political affiliations and the degree of their political involvement. Table 77 shows that while our respondents are rarely members of any community group themselves (only 2% belong to one, usually a religious or social group), they seem to be heavily engaged in community politics. Almost without exception, respondents voted in the last Lok Sabha election, and, at the more local level, 56% of them participated in a meeting called by the gram panchayat within the last year.

Table 77: Group membership and political activity

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Member of group	4.69 (0.615)	4.19 (1.326)	4.92 (0.822)	4.97 (0.960)	0.13	0.88	3434
Voted Lok Sabha election	98.0 (0.324)	97.9 (0.498)	98.7 (0.461)	97.6 (0.690)	1.14	0.32	3434
Attended meeting called by GP	56.1 (1.847)	53.8 (3.514)	59.8 (3.228)	55.0 (2.685)	0.95	0.39	3434

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Woman Questionnaire.

We also asked about social networks of our male respondents, and by extension his household. Specifically, we ask whether they knew anyone outside their household who could come and help them to make house improvements for free. As far as sanitation is considered a home improvement, this social network hence tells us something about relationships that may help the household in the construction of a toilet (Table 78). We find that this does not seem to be the case for 98% of the sample. On the other hand, a majority of the respondents (53%) seem to have someone to whom they can resort to if in need of borrowing Rs. 1000 urgently.

Table 78: Social networks

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Help dwelling improvements	2.39 (0.368)	2.48 (0.641)	1.93 (0.570)	2.75 (0.690)	0.46	0.63	3434
Borrowing Rs. 1000 readily	53.1 (1.694)	53.9 (2.917)	51.2 (2.712)	54.0 (3.145)	0.31	0.73	3434

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Man Questionnaire.

5.4.5 Sanitation

In this section, we provide a comprehensive overview of the men and women’s practices, expectations, perceived benefits and beliefs related to sanitation

Sanitation practices

The sanitation practices of our respondents are summarised in Tables 79 and 80. 29% of women report typically using their own toilet to defecate, however, the most common location used among them is an open area more than five minutes from their residence (70%). These practices are mirrored by men, of which 65% usually go for OD more than five minutes from their home. We do however in addition find a larger percentage of men performing OD near their residence (7%) than women (2%). Almost everyone report to be content with possessing a private toilet for doing their needs, whereas less than 20% of the respondents are happy with the open areas where they typically defecate. The large majority of women (82%) feel unsafe visiting these places, particularly when they are far from home. Despite the low levels of satisfaction the open locations are perceived sufficiently comfortable for about 32% of women and 44% of men. Other locations (such as neighbour’s toilet or community toilet) are little used in our sample and are not reported here.²⁷

²⁷The statistics in these tables do not present any significant differences across treatment arms with only two exceptions which have very small sample sizes. Balancedness among the different treatment groups also holds for the rest of the data on sanitation practices.

Table 79: OD practice, satisfaction, comfort and safety: Women

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Own toilet	28.9 (1.716)	26.8 (2.878)	33.4* (2.629)	26.4 (3.257)	1.98	0.14	3434
OD (close home)	2.04 (0.316)	2.22 (0.461)	1.23* (0.372)	2.66 (0.729)	2.29	0.11	3434
OD (>5 mins home)	69.5 (1.654)	71.1 (2.824)	65.6 (2.546)	71.7 (3.111)	1.52	0.22	3434
Satisfied with Own toilet	93.6 (1.202)	93.3 (2.134)	92.9 (2.196)	94.9 (1.718)	0.33	0.72	991
Satisfied with OD (>5 mins home)	15.2 (1.661)	16.6 (3.096)	17.3 (2.882)	11.8 (2.510)	1.26	0.29	2386
Comfortable visiting Own toilet	98.9 (0.350)	99.4 (0.444)	98.7 (0.669)	98.7 (0.628)	0.59	0.55	991
Comfortable visiting OD (>5 mins home)	32.2 (1.876)	32.2 (3.537)	29.5 (3.095)	34.8 (2.993)	0.76	0.47	2373
Safe visiting Own toilet	98.8 (0.389)	99.7 (0.315)	98.4 (0.706)	98.3 (0.849)	2.15	0.12	991
Safe visiting OD (>5 mins home)	18.8 (1.584)	18.1 (2.762)	17.5 (3.050)	20.7 (2.371)	0.42	0.66	2386

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Woman Questionnaire.

Table 80: OD practice, satisfaction, comfort and safety: Men

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Own toilet	29.0 (1.880)	26.1 (3.032)	33.7* (2.856)	27.2 (3.721)	1.91	0.15	2923
OD (>5 mins home)	64.8 (1.877)	66.6 (3.136)	60.9 (2.729)	66.8 (3.762)	1.27	0.29	2923
Satisfied with Own toilet	88.8 (3.051)	87.3 (5.331)	88.5 (5.377)	90.4 (5.032)	0.087	0.92	241
Satisfied with OD (>5 mins home)	20 (2.693)	20.9 (4.291)	21.1 (5.166)	18.0 (4.598)	0.14	0.87	410
Comfortable visiting Own toilet	91.1 (2.662)	87.3 (5.235)	88.1 (4.970)	97.5* (2.519)	2.45	0.093*	235
Comfortable visiting OD (close home)	43.5 (5.623)	54.2 (9.861)	41.7 (9.952)	33.3* (6.833)	1.47	0.24	69
Comfortable visiting OD (>5 mins home)	47.4 (3.747)	44.1 (7.153)	54.8 (5.993)	44.7 (5.756)	0.95	0.39	390

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Man Questionnaire.

It comes at no surprise that water and soap is rarely available at the site of open defecation (Tables 81 and 82). For households with toilets, we find that more than one third of men (38%) and one quarter of women (27%) report not having access to water at their dwelling. This is in line with findings reported above on access to water for sanitation, with only about a third of households reporting to have piped water into the house and having to undertake considerable average walking times to collect water. Of those who do have water available, the majority also seem to have soap (88% for woman and 75% for men). In spite of these difficulties to find water and soap at the site of defecation, the vast majority of men (91%) and women (83%) claim to always wash their hands after defecating, and usually with soap and water, 87% of men and 83% of women.

Table 81: Washing hands: Women

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Water washing hands at site Own toilet	72.9 (3.353)	80.5 (4.459)	73.7 (4.720)	63.9* (7.152)	1.99	0.14	987
Water washing hands at site OD (>5 mins home)	1.45 (0.504)	1.82 (1.029)	0.96 (0.577)	1.51 (0.890)	0.33	0.72	2351
Soap at site Own toilet	88.0 (2.115)	90.3 (2.657)	91.0 (2.415)	81.1 (5.719)	1.30	0.28	826
Soap at site OD (>5 mins home)	5.81 (1.339)	4.94 (1.190)	11.9* (3.921)	2.05** (0.723)	4.75	0.011**	775
Hand washing: Never	0.73 (0.165)	0.60 (0.238)	0.88 (0.327)	0.71 (0.287)	0.24	0.79	3434
Hand washing: Sometimes	5.74 (0.726)	6.07 (1.319)	4.66 (1.129)	6.48 (1.293)	0.64	0.53	3434
Hand washing: Most of the time	9.64 (1.510)	9.06 (2.628)	11.3 (2.843)	8.53 (2.324)	0.31	0.73	3434
Hand washing: Always	83.9 (1.742)	84.3 (3.431)	83.1 (2.931)	84.3 (2.594)	0.051	0.95	3434
Hand washing manner: Soap and water	82.6 (1.245)	83.3 (2.113)	82.9 (2.271)	81.7 (2.067)	0.16	0.85	3434
Hand washing manner: Ash and water	8.12 (0.791)	7.86 (1.436)	8.35 (1.262)	8.17 (1.402)	0.032	0.97	3434
Hand washing manner: Water only	7.28 (0.848)	6.92 (1.335)	6.41 (1.335)	8.53 (1.690)	0.50	0.61	3434
Hand washing manner: Wipe in soil	1.11 (0.287)	1.20 (0.527)	1.49 (0.608)	0.62 (0.277)	1.11	0.33	3434
Hand washing manner: Wipe in cloth	0.61 (0.223)	0.51 (0.316)	0.62 (0.433)	0.71 (0.400)	0.077	0.93	3434
Wash hands before cooking	99.6 (0.120)	99.7 (0.143)	99.6 (0.258)	99.6 (0.211)	0.22	0.80	3427

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Woman Questionnaire.

Table 82: Washing hands: Men

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Water washing hands at site Own toilet	62.1 (2.940)	64.1 (4.506)	62.3 (4.867)	59.6 (5.768)	0.19	0.83	817
Water washing hands at site OD (>5 mins home)	5.07 (0.698)	5.27 (1.166)	5.68 (1.376)	4.27 (1.082)	0.37	0.69	1757
Soap at site Own toilet	75.4 (3.719)	67.6 (6.975)	76.3 (5.752)	82.1 (5.660)	1.28	0.28	211
Soap at site OD (>5 mins home)	4.10 (1.542)	0.99 (0.998)	9.46* (4.350)	3.23 (2.437)	2.02	0.14	268
Hand washing: Never	0.14 (0.0680)	0.20 (0.138)	0.10 (0.103)	0.11 (0.106)	0.18	0.84	2923
Hand washing: Sometimes	3.73 (0.443)	3.27 (0.614)	3.69 (0.894)	4.26 (0.768)	0.51	0.60	2923
Hand washing: Most of the time	5.30 (0.818)	5.15 (1.516)	6.36 (1.536)	4.37 (1.118)	0.55	0.58	2923
Hand washing: Always	90.8 (1.008)	91.3 (1.746)	89.8 (1.987)	91.3 (1.436)	0.20	0.82	2923
Hand washing manner: Soap and water	87.0 (1.532)	88.7 (2.376)	88.9 (2.550)	83.5 (2.775)	1.29	0.28	702
Hand washing manner: Ash and water	4.70 (0.876)	4.60 (1.324)	3.10 (1.040)	6.33 (1.933)	1.20	0.31	702
Hand washing manner: Water only	8.26 (1.208)	6.69 (1.801)	7.96 (2.268)	10.1 (2.081)	0.78	0.46	702

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Man Questionnaire.

Regarding the location where people in the area usually go to urinate (Tables 83 and 84), the proportion of women using their own toilet is almost identical as for defecation (29%), yet it is some ten percentage points lower for men. The majority of men (55%) and a sizeable proportion of women (30%) opt to use a private bathroom enclosure (without toilet) to urinate. Urinating in an open area is still relatively common for women (42%) and men (30%). The comfort associated with urinating in the open pales in comparison with that provided by private urination sites, especially in the case of women, who report feeling comfortable using a private toilet in more than 90%²⁸ of the cases compared to less than 40% for open urination.

Table 83: Urination location: Women

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Own toilet	28.8 (1.690)	27.5 (2.814)	32.6 (2.620)	26.3 (3.233)	1.43	0.24	3434
Open area (close home)	37.7 (2.327)	35.7 (3.896)	35.0 (3.815)	42.4 (4.265)	0.97	0.38	3434
Open area (>5 mins home)	4.14 (0.644)	5.13 (1.451)	2.72 (0.478)	4.53 (1.118)	2.08	0.13	3434
Own bathroom/bathing enclosure	30.5 (2.088)	31.5 (3.843)	31.1 (3.416)	28.9 (3.546)	0.16	0.85	3434
Comfortable visiting Own toilet	97.9 (0.577)	98.8 (0.612)	97.0 (1.148)	98.0 (1.061)	0.93	0.40	989
Comfortable visiting Open area (close home)	38.8 (3.100)	33.4 (5.553)	36.1 (4.923)	45.7 (5.196)	1.50	0.23	1264
Comfortable visiting Open area (>5 mins home)	32.6 (4.878)	32.2 (6.985)	35.5 (11.02)	31.4 (8.219)	0.046	0.95	141
Comfortable visiting Own bathroom/bathing enclosure	91.7 (1.368)	92.4 (1.558)	89.0 (3.296)	93.8 (1.632)	0.88	0.42	1046

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Woman Questionnaire.

Table 84: Urination location: Men

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Own toilet	17.1 (1.246)	13.9 (1.675)	20.2** (2.126)	17.5 (2.497)	2.83	0.063*	2923
Open area (close home)	20.5 (1.288)	22.5 (2.361)	16.8* (1.997)	22.1 (2.214)	2.25	0.11	2923
Open area (>5 mins home)	9.48 (0.739)	8.91 (1.299)	9.64 (1.212)	9.91 (1.327)	0.16	0.85	2923
Own bathroom/bathing enclosure	54.5 (1.334)	55.7 (2.455)	55.0 (2.056)	52.6 (2.373)	0.49	0.61	2923
Comfortable visiting Own toilet	94.0 (1.375)	95.1 (1.820)	92.4 (2.747)	95.1 (1.996)	0.40	0.67	504
Comfortable visiting Open area (close home)	59.5 (2.902)	57.9 (4.758)	66.5 (4.544)	55.7 (5.293)	1.42	0.25	565
Comfortable visiting Open area (>5 mins home)	46.3 (4.233)	46.1 (7.081)	42.4 (6.881)	50 (7.656)	0.27	0.76	281
Comfortable visiting Own bathroom/bathing enclosure	73.4 (2.591)	73.8 (4.360)	75.1 (3.954)	70.9 (5.114)	0.21	0.81	1588

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Man Questionnaire.

²⁸Which however still leaves almost 10% of women feel uncomfortable using the own toilet.

As can be seen in tables 85 and 86, there is a much greater variety in bathing locations than places used for defecation. The typical bathing site for the majority of respondents is a closed bathroom, located either inside the house (as for 28% of women and 35% of men) or outside (22% and 27%, respectively). Indoor closed bathrooms seem to be significantly more common among female members of the GK group with respect to the other groups, as indicated by the low p-values registered for the equal means test. The same applies to the percentage of men who bathe in a shielded structure outside the house, which is less common among respondents of the GK treatment area (6%) and more so for those in the control group (14%). For all other bathing locations the descriptives are statistically equivalent for all groups.

Table 85: Bathing location: Women

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Closed bathroom inside house	27.9 (1.739)	24.0 (2.667)	34.4** (3.333)	25.5 (2.753)	3.21	0.044**	3434
Bathroom enclosure inside the house	11.8 (0.839)	11.8 (1.206)	11.3 (1.410)	12.3 (1.711)	0.087	0.92	3434
Part of a room inside house	3.96 (0.643)	4.87 (1.195)	3.16 (1.069)	3.82 (1.043)	0.57	0.57	3434
Closed bathroom outside house	21.7 (1.722)	22.3 (2.487)	20.7 (3.183)	22.0 (3.247)	0.079	0.92	3434
Bathroom enclosure outside house	10.3 (0.763)	11.5 (1.317)	9.05 (1.161)	10.3 (1.446)	0.94	0.39	3434
Shielded structure outside house	17.2 (1.251)	18.2 (2.139)	15.6 (2.174)	17.7 (2.164)	0.42	0.66	3434
Space outside house not enclosed	2.30 (0.348)	2.74 (0.540)	1.58 (0.601)	2.58 (0.648)	1.13	0.33	3434
Open bathing space outside house	5.68 (0.711)	5.73 (1.338)	4.75 (0.809)	6.57 (1.434)	0.68	0.51	3434

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Woman Questionnaire.

Table 86: Bathing location: Men

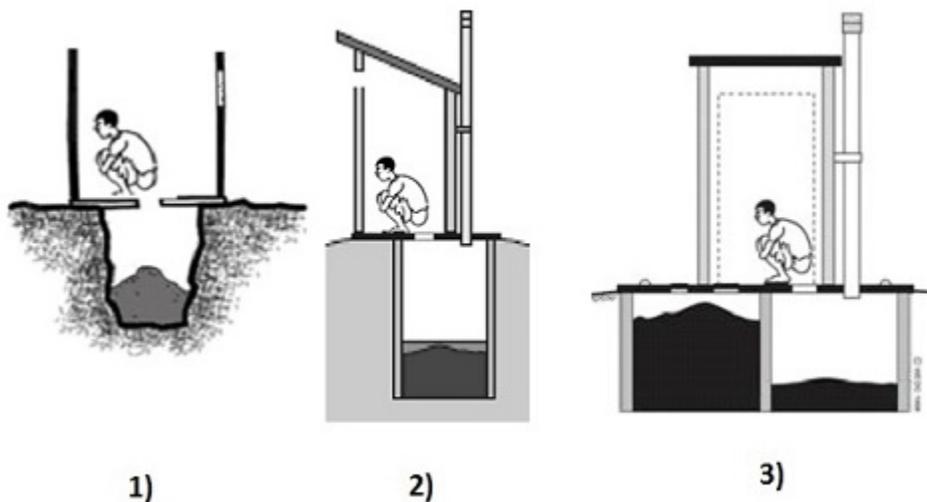
	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Closed bathroom inside house	34.9 (2.400)	33.5 (4.141)	35.7 (4.384)	35.4 (3.941)	0.084	0.92	703
Bathroom enclosure inside the house	20.1 (1.960)	16.7 (2.765)	22.9 (3.883)	20.7 (3.457)	0.94	0.39	703
Part of a room inside house	2.70 (0.687)	2.51 (1.258)	3.08 (1.159)	2.53 (1.137)	0.077	0.93	703
Closed bathroom outside house	27.0 (2.611)	25.1 (4.013)	29.1 (4.378)	27.0 (5.066)	0.22	0.80	703
Bathroom enclosure outside house	7.97 (1.228)	7.95 (2.165)	6.17 (1.747)	9.70 (2.381)	0.74	0.48	703
Shielded structure outside house	8.96 (1.406)	13.8 (2.938)	5.73** (1.664)	7.17* (2.165)	2.87	0.061*	703
Space outside house not enclosed	2.28 (0.582)	2.09 (0.841)	2.20 (1.059)	2.53 (1.109)	0.051	0.95	703
Open bathing space outside house	2.70 (0.703)	2.09 (1.235)	2.20 (0.909)	3.80 (1.407)	0.53	0.59	703

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Man Questionnaire.

Sanitation expectations, beliefs and perceived benefits

This subsection attempts to shed light on the perceptions and beliefs that people in the area hold about sanitation. The aim is to unravel possible reasons for why ownership of sanitation facilities is not more widespread in our study area.

Figure 14: Toilet Types



We start by examining respondents' price and duration expectations for three different types of toilets, ranging from the most basic to a more sophisticated installation. A visual representation of the three type of toilets was shown to respondents and is here replicated in Figure 14. These types were further described before asking questions related to them: Toilet type 1 is a simple toilet, with a hole dug out to be a pit (no lining) and a simple structure that ensures that one cannot be seen by others when using it. This structure could be made of plastic sheets, palm leave or other locally available materials. The second kind of toilet is one with a lined single pit, and a structure that provides privacy and protects from the elements. Finally, the third option is a toilet displaying a twin pit toilet with a pucca super-structure (i.e. walls, roof, door), which ensures privacy. This toilet can be locked and has a vent-pipe for aeration; The latter toilet is often interpreted by people in that area as being a septic tank (and therefore this picture was also described to them as such), even though in practice it is not.

Respondents were shown these pictures and asked to estimate the minimum and maximum cost of each toilet type in turn, their answers are summarised in Tables 87 and 88. Both men and women consider that the most basic type of toilet ranges from about seven thousand (\sim USD 112) to nine thousand rupees (\sim USD 145). The more advanced second option is considered to imply a substantial jump in prices in the mind of our respondents, costing a minimum of Rs. 20,000 (\sim USD 312) and a maximum of about Rs. 26,000 (\sim USD 430). The price of the third and highest quality installation more than doubles with respect to the second type. Respondents reckon that this kind of toilet is worth at least Rs. 43,000 (\sim USD 691) and that its cost can reach up to Rs. 52,000 (\sim USD 835). While both genders report relatively similar costs, it

is worth noting that women report throughout slightly lower estimates than men. In general though, it is fair to say that, independent of the gender, respondents believe that building a toilet requires large financial commitments, which can be a difficult obstacle to overcome judging by what we have learned so far about income levels and access to finance for people in our study area.

Table 87: Expected cost and duration of toilets: Women

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Min. cost: Toilet type 1	6331 (185.4)	6331 (294.0)	6345 (298.8)	6591 (362.7)	0.65	0.53	3434
Min. cost: Toilet type 2	19343 (495.1)	19343 (874.9)	19294 (871.1)	19413 (823.2)	0.0054	0.99	3434
Min. cost: Toilet type 3	43983 (1189.2)	43983 (2067.4)	43779 (2100.9)	44256 (2006.1)	0.014	0.99	3434
Max. cost: Toilet type 1	8544 (249.4)	8544 (388.3)	8585 (394.2)	8924 (498.7)	0.82	0.44	3434
Max. cost: Toilet type 2	25472 (680.2)	25472 (1183.7)	25559 (1231.3)	25474 (1114.1)	0.0052	0.99	3434
Max. cost: Toilet type 3	53808 (1495.7)	53808 (2496.2)	53013 (2548.0)	55070 (2713.1)	0.17	0.84	3434
Min. duration: Toilet type 1	5.27 (0.168)	5.07 (0.248)	5.34 (0.327)	5.41 (0.291)	0.45	0.64	3434
Min. duration: Toilet type 2	18.3 (0.506)	18.2 (0.796)	17.4 (0.833)	19.3 (0.964)	1.14	0.32	3434
Min. duration: Toilet type 3	37.1 (0.816)	37.2 (1.436)	35.8 (1.294)	38.3 (1.473)	0.79	0.46	3434
Max. duration: Toilet type 1	7.17 (0.232)	6.80 (0.323)	7.29 (0.455)	7.42 (0.413)	0.82	0.44	3434
Max. duration: Toilet type 2	23.5 (0.610)	23.4 (0.922)	22.6 (1.042)	24.6 (1.172)	0.83	0.44	3434
Max. duration: Toilet type 3	45.2 (0.976)	45.2 (1.774)	43.9 (1.513)	46.4 (1.740)	0.62	0.54	3432

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Woman Questionnaire.

Table 88: Expected cost and duration of toilets: Men

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Min. cost: Toilet type 1	7284 (241.0)	7284 (411.4)	6829 (371.5)	7597 (455.7)	1.02	0.36	3434
Min. cost: Toilet type 2	22155 (511.2)	22155 (802.6)	21990 (916.2)	22254 (936.3)	0.025	0.98	3434
Min. cost: Toilet type 3	43704 (800.9)	43704 (1052.0)	42295* (1395.8)	43402 (1633.9)	1.65	0.20	3434
Max. cost: Toilet type 1	9929 (333.2)	9929 (565.0)	9238 (498.8)	10488 (641.3)	1.31	0.27	3434
Max. cost: Toilet type 2	28447 (608.6)	28447 (919.4)	28055 (1070.4)	28557 (1165.3)	0.12	0.89	3434
Max. cost: Toilet type 3	52058 (895.3)	52058 (1146.6)	50067** (1527.8)	51898 (1854.3)	2.35	0.100*	3434
Min. duration: Toilet type 1	4.50 (0.121)	4.42 (0.224)	4.47 (0.192)	4.62 (0.210)	0.25	0.78	3434
Min. duration: Toilet type 2	14.7 (0.562)	14.4 (0.779)	14.8 (1.052)	15.0 (1.069)	0.12	0.89	3434
Min. duration: Toilet type 3	29.4 (1.140)	29.0 (1.743)	29.4 (2.017)	30.0 (2.147)	0.062	0.94	3434
Max. duration: Toilet type 1	6.50 (0.166)	6.41 (0.300)	6.36 (0.258)	6.74 (0.296)	0.52	0.60	3434
Max. duration: Toilet type 2	19.3 (0.683)	18.9 (0.996)	19.2 (1.238)	19.8 (1.296)	0.15	0.86	3434
Max. duration: Toilet type 3	36.8 (1.351)	36.3 (2.166)	36.4 (2.292)	37.6 (2.549)	0.087	0.92	3424

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Man Questionnaire.

The costs estimated by respondents stand in stark contrast with the calculations done by the government and the GK microfinance institute, who believe that a decent quality two pit toilet technology can be constructed at a costs of at most Rs. 12,000-15,000. This suggests that there is widespread misperception among the population about the actual costs of making sanitation investments. One of the objectives of the ND intervention in our study is to better inform the people in this regard.

The average estimated time that the household can use the toilet without any hassle (repair or having pit fill up) raises substantially from the first to the last type of toilet. While toilet type 1 is expected to last no more than seven years, the second one is expected be in service for at least sixteen years and up to twenty according to participants. Lastly, the third installation is estimated to be operational a minimum of three times the years of the most basic toilet, and could last over 40 years.

Although women estimated costs slightly lower than men, they tend to be more optimistic than men regarding the longevity of the toilets. No significant differences are observed in the expectations of the different treatment groups.

We now turn to the beliefs held about sanitation by people in our study area. Tables 89 and 90 depict whether respondents agree (or strongly agree) with a wide array of statements. Responses for men and women are fairly similar, and no significant differences are found across the treatment groups. The great majority of the people interviewed (close to 90%) consider that only poor people struggle to find the funding

for constructing a toilet, and more than 60% agree with the notion that such financial barriers affect the whole of the community. There is widespread support (~65%) to the idea that sanitation is the responsibility of the government.

Table 89: Beliefs about sanitation: Women

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Sanitation responsibility of government	69.0 (2.128)	70.9 (3.641)	65.2 (4.098)	71.0 (3.185)	0.73	0.48	3434
All community members struggle to fund building toilet	59.4 (2.159)	61.2 (3.822)	54.6 (3.541)	62.5 (3.709)	1.39	0.25	3434
Only poor members struggle to fund building toilet	88.4 (1.825)	89.0 (3.150)	85.9 (3.507)	90.1 (2.738)	0.45	0.64	3434
Even if HH has toilet, HH members don't use it	22.5 (2.080)	23.9 (3.960)	22.8 (3.550)	20.7 (3.200)	0.22	0.80	3434
Even if HH has toilet, men don't use it	22.9 (2.047)	24.7 (4.010)	23.0 (3.345)	21.0 (3.149)	0.27	0.77	3434
Even if HH has toilet, children don't use it	20.6 (2.151)	21.4 (4.169)	22.0 (3.634)	18.4 (3.256)	0.31	0.73	3434
Even if HH has toilet, older HH members don't use it	21.6 (2.172)	23.2 (4.182)	22.6 (3.607)	19.1 (3.378)	0.38	0.69	3434
If HH has toilet, neighbours will come to use it	16.5 (1.973)	16.2 (3.731)	18.2 (3.450)	15.1 (2.990)	0.23	0.79	3434
If neighbours don't use toilet, more likely I fall sick	67.5 (1.724)	63.0 (3.044)	68.6 (2.495)	71.0* (3.186)	1.82	0.17	3434
If neighbours use toilet, environment I live cleaner	76.2 (1.734)	72.9 (3.236)	78.5 (2.651)	77.2 (2.990)	0.92	0.40	3434
Acceptable to OD as ancestors did it	21.0 (2.113)	23.4 (3.917)	23.5 (3.800)	15.9 (3.034)	1.71	0.19	3434
If people OD, nobody minds as is common habit	26.4 (2.232)	27.7 (4.211)	30.1 (3.770)	21.2 (3.399)	1.67	0.19	3434

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Woman Questionnaire.

Table 90: Beliefs about sanitation: Men

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Sanitation responsibility of government	63.1 (1.537)	66.8 (2.143)	60.0* (2.656)	62.4 (3.014)	2.06	0.13	3434
All community members struggle to fund building toilet	64.2 (1.615)	67.4 (2.604)	60.5* (2.812)	64.6 (2.851)	1.66	0.19	3434
Only poor members struggle to fund building toilet	91.4 (0.875)	94.4 (0.979)	90.8* (1.527)	88.9*** (1.823)	4.39	0.014**	3434
Even if HH has toilet, HH members don't use it	20.0 (2.348)	22.5 (4.498)	15.6 (3.543)	21.8 (3.982)	1.01	0.37	3434
Even if HH has toilet, men don't use it	20.9 (2.344)	23.1 (4.532)	16.6 (3.504)	23.0 (3.960)	0.97	0.38	3434
Even if HH has toilet, children don't use it	17.8 (2.235)	20.5 (4.306)	12.1 (2.999)	20.7 (3.996)	2.05	0.13	3434
Even if HH has toilet, older HH members don't use it	18.3 (2.209)	20.8 (4.244)	13.5 (3.169)	20.6 (3.840)	1.41	0.25	3434
If HH has toilet, neighbours will come to use it	15.4 (1.955)	18.7 (3.654)	11.2 (2.916)	16.3 (3.403)	1.46	0.24	3434
If neighbours don't use toilet, more likely I fall sick	74.9 (1.678)	73.2 (3.089)	75.7 (3.023)	75.7 (2.529)	0.23	0.80	3434
If neighbours use toilet, environment I live cleaner	85.1 (1.400)	82.4 (2.616)	86.1 (2.402)	87.0 (2.133)	1.00	0.37	3434
Acceptable to OD as ancestors did it	26.6 (1.427)	25.3 (2.053)	24.9 (2.382)	29.6 (2.870)	0.94	0.39	3434
If people OD, nobody minds as is common habit	29.2 (1.464)	28.1 (2.359)	26.4 (2.359)	33.2 (2.758)	1.85	0.16	3434

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Man Questionnaire.

Under the scenario hypothetical for many - that the household owns a toilet- one

Table 91: Toilet construction responsibility: Women

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Every HH should be responsible own toilet	17.1 (1.352)	13.3 (1.649)	19.6** (2.375)	18.6 (2.774)	2.88	0.060*	3434
Government should provide toilets for all	37.3 (1.766)	36.3 (3.377)	39.4 (3.301)	36.3 (2.326)	0.32	0.73	3434
Government should just support the poor	31.6 (1.888)	34.3 (3.450)	27.4 (3.232)	33.0 (2.990)	1.25	0.29	3434
Gram Panchayat should be responsible	13.7 (1.206)	16.1 (2.357)	13.2 (1.939)	11.9 (1.852)	0.97	0.38	3434
Other should be responsible	0.17 (0.108)	0 (0)	0.35 (0.274)	0.18 (0.176)	1.34	0.27	3434

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Woman Questionnaire.

Table 92: Toilet construction responsibility: Men

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Every HH should be responsible own toilet	32.9 (1.665)	31.5 (2.342)	34.9 (2.779)	32.4 (3.431)	0.43	0.65	3434
Government should provide toilets for all	29.8 (1.246)	31.1 (2.063)	28.8 (2.205)	29.3 (2.187)	0.33	0.72	3434
Government should just support the poor	30.8 (1.580)	29.9 (2.476)	32.2 (2.839)	30.5 (2.883)	0.19	0.83	3434
Gram Panchayat should be responsible	6.32 (1.157)	7.18 (2.114)	4.13 (1.635)	7.64 (2.165)	1.09	0.34	3434
Other should be responsible	0.029 (0.0291)	0.085 (0.0849)	0 (0)	0 (0)	1.01	0.32	3434

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Man Questionnaire.

out of five respondents reports that household members will not use it. Importantly, almost 30% of people still hold the view that open defecation is justified on the grounds that it is commonly practiced and/or that it was done by past generations. Most respondents (over 70%) are however aware that improved sanitation practices by the neighbours would reduce the incidence of disease suffered by their own household and would improve their living environment.

Consistent with what we observed above, Tables 91 and 92 about the question of who should be responsible for the construction of toilets show that more than 60% of respondents believe that the government should take the lead in this effort, either supporting only the poor or every household without a toilet. Interestingly, a larger proportion of men (32%) than women (17%) think that households should in fact be ultimately responsible for building their own toilets.

Finally, Tables 93 and 94 examine the benefits that our respondents consider a toilet may bring. Almost every man and woman interviewed agree that a toilet would make their family happier and healthier, enable household members to work more, improve the social status of the family, save them time in doing their needs and make women in the household feel safer. Note, however, that over one quarter of the respondents think that having a toilet has some negative implications, such as unhealthy smells, less physical exercise, less time to socialise with others and/or negative health impacts.

Importantly, over 60% of the women (compared to only 44% of men) state that having a toilet means that the family has to spend more time fetching water.

Table 93: Benefits of sanitation: Women

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Family will be happier	94.9 (0.704)	95.5 (0.993)	95.4 (0.971)	93.9 (1.594)	0.41	0.66	3434
Family will be healthier	94.0 (0.749)	95.0 (1.001)	93.8 (1.173)	93.3 (1.639)	0.49	0.61	3434
HH members able to work more	94.2 (0.627)	94.8 (0.791)	94.8 (0.996)	93.0 (1.384)	0.72	0.49	3434
Family less embarrassed when family and friends come	93.7 (0.671)	94.1 (0.939)	94.2 (1.006)	92.6 (1.471)	0.44	0.64	3434
Family less embarrassed when guests from outside visit	93.4 (0.693)	94.0 (0.985)	94.0 (1.022)	92.2 (1.517)	0.60	0.55	3434
Women in family safer	96.0 (0.623)	96.0 (0.866)	96.1 (1.017)	95.7 (1.321)	0.028	0.97	3434
Family will save time	94.3 (0.663)	95.4 (0.898)	93.8 (1.181)	93.5 (1.317)	0.91	0.40	3434
Toilets are unhealthy because they stink	25.0 (1.318)	23.2 (2.352)	27.4 (2.301)	24.3 (2.143)	0.88	0.42	3434
Family members will get less exercise	36.4 (2.419)	35.9 (4.175)	42.0 (4.118)	31.3 (4.109)	1.69	0.19	3434
Family members will miss out spending time w/others	35.5 (2.147)	33.7 (3.709)	40.9 (3.571)	32.0 (3.734)	1.69	0.19	3434
Family members might get sick more easily	26.4 (1.921)	24.1 (3.321)	31.3 (3.411)	23.9 (3.113)	1.58	0.21	3434
Family will have to spend more time fetching water	62.5 (1.891)	60.3 (3.397)	67.3 (2.932)	59.9 (3.318)	1.83	0.16	3434

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Woman Questionnaire.

Table 94: Benefits of sanitation: Men

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Family will be happier	94.6 (0.724)	94.5 (1.316)	93.8 (1.308)	95.4 (1.107)	0.41	0.67	3434
Family will be healthier	93.0 (0.860)	92.9 (1.569)	92.6 (1.502)	93.4 (1.380)	0.082	0.92	3434
HH members able to work more	91.8 (0.932)	92.3 (1.425)	91.6 (1.478)	91.4 (1.911)	0.099	0.91	3434
Family less embarrassed when family and friends come	91.3 (1.022)	92.8 (1.582)	91.3 (1.537)	89.8 (2.125)	0.68	0.51	3434
Family less embarrassed when guests from outside visit	91.2 (0.982)	92.6 (1.520)	90.3 (1.682)	90.7 (1.875)	0.61	0.54	3434
Women in family safer	95.1 (0.683)	94.8 (1.244)	95.4 (0.972)	95.0 (1.304)	0.089	0.92	3434
Family will save time	93.4 (0.818)	93.0 (1.381)	93.7 (1.334)	93.7 (1.530)	0.082	0.92	3434
Toilets are unhealthy because they stink	17.3 (1.176)	16.2 (1.900)	15.7 (1.918)	20.0 (2.206)	1.23	0.30	3434
Family members will get less exercise	26.1 (1.686)	25.4 (2.984)	24.2 (2.852)	28.9 (2.855)	0.73	0.49	3434
Family members will miss out spending time w/others	25.4 (1.729)	24.3 (2.835)	23.9 (3.104)	28.1 (2.990)	0.59	0.55	3434
Family members might get sick more easily	21.0 (1.210)	20 (1.828)	19.4 (2.178)	23.5 (2.203)	1.07	0.35	3434
Family will have to spend more time fetching water	44.2 (1.894)	43.8 (2.966)	45.1 (3.651)	43.8 (3.191)	0.044	0.96	3434

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Man Questionnaire.

5.4.6 Personal hygiene

In addition to the detailed information about their sanitation practices, women individually interviewed by female interviewers were also enquired about their intimate hygiene practices. Their answers are summarised in Table 95. Most of the 80% of female respondents who still menstruate use a cloth to prevent the bloodstains associated with the period from becoming apparent (82%) . Only 15% use a sanitary napkin. The majority of the women dispose of their protection by burning it (60%) and/or wash and re-use it (60%). Ten percent still simply throw it away into the environment.

Table 95: Personal hygiene women

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Gets menstruation	80.4 (0.944)	82.2 (1.444)	80.8 (1.583)	78.2* (1.810)	1.55	0.22	3434
Use cloth	82.4 (1.408)	81.3 (2.557)	82.4 (2.097)	83.6 (2.614)	0.21	0.81	2762
Use cotton	9.12 (1.399)	10.5 (2.512)	9.02 (2.093)	7.73 (2.607)	0.29	0.75	2762
Use sanitary napkin	14.8 (1.022)	14.4 (1.332)	15.5 (2.323)	14.3 (1.499)	0.11	0.90	2762
Disposal: throw it in toilet pit	2.24 (0.407)	2.60 (0.695)	2.72 (0.803)	1.36 (0.561)	1.40	0.25	2762
Disposal: throw it in field	0.80 (0.178)	1.14 (0.293)	0.33** (0.183)	0.91 (0.403)	3.14	0.047**	2762
Disposal: throw it away along other rubbish	10.2 (1.132)	9.56 (1.737)	13.2 (2.410)	7.95 (1.478)	1.69	0.19	2762
Disposal: burn it	59.9 (1.398)	64.3 (2.407)	57.4** (2.180)	57.7* (2.486)	2.74	0.069*	2762
Disposal: wash and re-use it	59.5 (2.290)	58.1 (4.080)	56.8 (3.527)	63.9 (4.148)	0.89	0.41	2762

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Woman Questionnaire.

5.4.7 Health

The focus of this section is on individuals' health condition and the problems that women face when seeking treatment, and evaluate their awareness of the causes of diarrhoea and how to prevent it.

Health condition

Starting with the perception of the sampled participants about their own health in Tables 96 and 97, we see that most people consider themselves to be in good health (about 60%), or at least have fair health (approx. 27%). This is not only similar between men and women, but also across groups by treatment status. We observe a statistically significance difference in the proportion of women reporting to be in extremely poor health but in terms of magnitude this difference is very small (less than 0.26 percentage points).

Table 96: Health condition: Women

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Extremely poor health	0.17 (0.0697)	0.26 (0.143)	0* (0)	0.27 (0.149)	3.22	0.043**	3434
Poor health	3.38 (0.400)	3.50 (0.719)	2.99 (0.607)	3.64 (0.741)	0.28	0.76	3434
Fair health	27.9 (1.524)	26.8 (2.164)	31.5 (3.119)	25.5 (2.471)	1.23	0.30	3434
Good health	60.9 (1.661)	62.6 (2.404)	57.6 (3.383)	62.7 (2.703)	0.88	0.42	3434
Very good health	6.90 (0.739)	6.50 (1.192)	7.03 (1.409)	7.19 (1.230)	0.090	0.91	3434
Excellent health	0.67 (0.166)	0.43 (0.179)	0.88 (0.343)	0.71 (0.314)	0.82	0.44	3434

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Woman Questionnaire.

Table 97: Health condition: Men

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Extremely poor health	0.31 (0.0988)	0.099 (0.0981)	0.41 (0.196)	0.43 (0.201)	1.74	0.18	2923
Poor health	1.92 (0.261)	2.38 (0.559)	1.44 (0.339)	1.92 (0.405)	1.14	0.32	2923
Fair health	27.3 (1.599)	26.1 (2.604)	26.6 (2.658)	29.4 (3.041)	0.38	0.69	2923
Good health	58.0 (1.641)	60.8 (2.519)	56.8 (2.921)	56.2 (3.052)	0.86	0.43	2923
Very good health	9.61 (1.101)	8.81 (1.687)	10.8 (2.057)	9.28 (1.961)	0.28	0.76	2923
Excellent health	2.84 (0.504)	1.78 (0.595)	4* (1.142)	2.77 (0.763)	1.64	0.20	2923

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Man Questionnaire.

Having said that, almost a quarter of the women interviewed suffered from stomach ache or fever in the last four weeks, close to 18% had been coughing during the period, and around 15% experienced fatigue or shakiness. Our restricted sub-sample of men, likely younger than the overall sample, seem not to have undergone any kind of sickness in the last month other than perhaps a fever, reported by 8% of them.

Table 98: Suffered from disease: Women

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Stomach ache	22.5 (1.365)	21.3 (2.514)	23.6 (2.213)	22.8 (2.329)	0.25	0.78	3433
Nausea/vomiting	8.54 (0.895)	9.32 (1.721)	7.21 (1.305)	9.07 (1.559)	0.65	0.52	3432
Diarrhoea	5.24 (0.646)	5.47 (1.202)	4.31 (1.151)	5.96 (0.964)	0.61	0.54	3432
Cough	17.6 (0.994)	18.4 (1.898)	18.2 (1.907)	16.2 (1.219)	0.67	0.51	3431
Fever	26.9 (1.237)	26.9 (2.012)	28.1 (2.406)	25.7 (1.965)	0.32	0.73	3430
Toothache	7.96 (0.693)	7.95 (1.190)	8.17 (1.198)	7.75 (1.213)	0.030	0.97	3430
A burning feeling during urination	5.89 (0.604)	5.81 (1.040)	6.50 (1.074)	5.34 (1.012)	0.31	0.73	3431
Frequent and intense urge to urinate	4.98 (0.693)	5.21 (1.355)	4.75 (0.939)	4.98 (1.255)	0.042	0.96	3432
Fatigue or shakiness	14.5 (1.429)	17.2 (2.986)	13.2 (2.028)	12.9 (2.210)	0.77	0.46	3430
A bad smelling genital discharge	5.21 (1.167)	6.67 (2.520)	4.92 (1.519)	4 (1.826)	0.37	0.69	3433
Genital itching	5.13 (1.045)	5.81 (2.153)	4.75 (1.447)	4.80 (1.727)	0.093	0.91	3432

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Woman Questionnaire.

Table 99: Suffered from disease: Men

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Stomach ache	5.12 (1.077)	4.18 (1.275)	3.52 (1.716)	7.59 (2.376)	1.04	0.36	703
Nausea/vomiting	1.14 (0.389)	1.67 (0.783)	0.88 (0.623)	0.84 (0.578)	0.42	0.66	703
Diarrhoea	1.14 (0.392)	1.26 (0.701)	0.44 (0.437)	1.69 (0.813)	1.13	0.33	703
Cough	3.56 (0.782)	2.93 (1.346)	2.64 (1.219)	5.06 (1.392)	0.96	0.38	703
Fever	7.54 (1.137)	6.69 (1.623)	7.49 (2.040)	8.44 (2.187)	0.21	0.81	703
Toothache	0.85 (0.333)	0.84 (0.567)	1.32 (0.718)	0.42 (0.407)	0.64	0.53	703

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Man Questionnaire.

Problems getting medical treatment

We have seen above how women struggle to have their voice heard when it comes to important households decisions. These problems extend to the domain of health care as well. The majority (59%) of women find it difficult to cover the distance to the health provider, in line with about half of the sample not having access to any health facility within their village (reported in Section 4.4.2) And that is only if they get permission to make the visit (a challenge for 41% of them) and gather the money to meet the costs of the treatment, which is reported to be no easy task for more than half of the sample. Additional barriers arise from the limitations posed by the health care system in our study area, which in the eyes of our female respondents is usually understaffed (68%), especially with female health workers who can make the visit more comfortable, and undersupply of drugs (70%).

Table 100: Problems in getting medical treatment

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Getting permission	40.5 (2.857)	34.6 (4.721)	46.6* (5.144)	40.4 (4.810)	1.47	0.23	3434
Getting money for treatment	52.9 (2.838)	47.5 (4.819)	59.1* (4.833)	52.4 (4.937)	1.44	0.24	3434
Distance to health facility	59.2 (2.680)	53.9 (4.542)	63.9 (4.594)	60.0 (4.679)	1.21	0.30	3434
Transport issues	62.4 (2.658)	56.7 (4.470)	65.1 (4.712)	65.5 (4.500)	1.23	0.30	3434
Finding someone to go with you	52.1 (2.706)	50.1 (4.749)	53.3 (4.977)	53.1 (4.281)	0.14	0.87	3434
Concern no female health provider	63.4 (2.384)	60.3 (3.885)	64.3 (4.432)	65.7 (4.016)	0.51	0.60	3434
Concern no health provider	68.2 (2.131)	66.3 (3.543)	68.7 (3.886)	69.6 (3.623)	0.23	0.80	3434
Concern no drugs available	70.4 (1.998)	68.0 (3.175)	71.4 (3.781)	72.0 (3.379)	0.42	0.66	3434

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Woman Questionnaire.

Health knowledge

To conclude this subsection, we assess the understanding of women about what the main causes of diarrhoea in children are (Table 101) and how it can be prevented (Table 102). 80% of women correctly state that water can carry diseases, implying that 2 out of 10 women are not aware that water can carry diseases. They see dirty water (60%), eating food gone bad (about 50%) and an unbalanced diet (47%) as the main carriers of diarrhoea. A lower proportion of respondents mention poor hygiene, in the form of not washing one hands (37%), open defecation (22%), unwashed food (29%) or dirtiness of the residence (20%), as responsible for children contracting the disease.

Table 101: Health knowledge: Causes diarrhoea

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Thinks water can carry diseases	79.8 (1.158)	78.5 (2.098)	80.3 (2.124)	80.7 (1.749)	0.34	0.72	3433
Dirty water	60.3 (2.174)	57.7 (3.432)	59.4 (3.953)	64.0 (3.833)	0.79	0.46	3433
Eating spoiled food	54.3 (2.421)	55.6 (3.921)	51.6 (4.228)	55.6 (4.391)	0.29	0.75	3433
Eating food touched by flies	47.3 (2.322)	45.6 (4.195)	48.2 (4.238)	48.2 (3.571)	0.14	0.87	3433
An unbalanced diet	46.9 (1.584)	50.9 (2.553)	46.1 (2.929)	43.6** (2.628)	2.07	0.13	3433
Not washing one's hands	33.6 (1.783)	33.7 (3.117)	33.2 (3.099)	33.8 (3.042)	0.013	0.99	3433
Teething/New teeth arriving	32.1 (1.736)	35.4 (3.269)	31.4 (2.572)	29.5 (3.058)	0.90	0.41	3433
Exposure to the sun	11.7 (1.223)	12.2 (2.218)	11.3 (2.063)	11.5 (2.057)	0.045	0.96	3433
Certain types of vaccinations	9.26 (1.314)	11.2 (2.822)	9.41 (1.892)	7.10 (1.881)	0.82	0.44	3433
Unwashed food	28.8 (1.608)	26.2 (2.724)	31.4 (2.678)	28.8 (2.903)	0.94	0.39	3433
Changing weather	31.5 (2.267)	30.6 (3.984)	31.8 (3.974)	32.0 (3.807)	0.037	0.96	3433
Mother's milk	5.83 (0.677)	5.64 (1.165)	4.93 (1.055)	6.93 (1.264)	0.74	0.48	3433
Bottle feeding	9.50 (0.967)	9.32 (1.606)	9.85 (1.703)	9.33 (1.718)	0.033	0.97	3433
Eating raw food	21.5 (1.933)	21.7 (3.450)	22.9 (3.415)	19.9 (3.141)	0.21	0.81	3433
Using dirty toilets	14.7 (1.414)	12.7 (2.427)	16.3 (2.647)	15.1 (2.221)	0.52	0.60	3433
Open defecation	21.8 (1.719)	17.9 (2.606)	22.3 (3.408)	25.4** (2.754)	2.00	0.14	3433
Dirtiness of house	19.5 (1.553)	18.5 (2.470)	17.9 (2.939)	22.1 (2.597)	0.75	0.47	3433
Dirtiness of neighbours/area	7.25 (0.894)	5.38 (1.364)	6.86 (1.591)	9.59** (1.615)	1.99	0.14	3433

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Woman Questionnaire.

Regarding the ways in which children can be prevented from having diarrhoea, most importantly, our respondents cite good personal hygiene (66%) and practices, like

hand washing (50%) and protecting the food (54%) and water (57%) consumed, as key elements.

Table 102: Health knowledge: Prevention diarrhoea

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Protect environment	35.1 (1.817)	32.8 (2.761)	31.3 (3.074)	41.3* (3.344)	2.78	0.066*	3433
Protect food	54.4 (2.407)	54.1 (4.041)	53.0 (4.357)	56.1 (4.093)	0.14	0.87	3433
Protect water	56.7 (1.662)	56.4 (2.559)	54.0 (3.036)	59.7 (2.952)	0.91	0.41	3433
Good personal hygiene	65.9 (1.804)	63.2 (3.323)	66.2 (3.277)	68.5 (2.641)	0.76	0.47	3433
Wash hands before eating	49.9 (2.220)	50.6 (3.739)	50.1 (4.227)	48.9 (3.526)	0.056	0.95	3433
Wash hands before cooking	43.6 (2.525)	42.2 (4.281)	44.0 (4.784)	44.8 (4.014)	0.096	0.91	3433
Wash hands before serving	36.2 (2.484)	34.4 (4.218)	39.3 (4.650)	34.9 (3.983)	0.36	0.70	3433
Wash hands after defecation	44.1 (2.228)	43.1 (3.968)	44.2 (3.972)	45.1 (3.614)	0.072	0.93	3433
Wash hands after removing faeces	21.4 (1.973)	22.2 (3.612)	24.0 (3.591)	17.9 (2.907)	0.96	0.39	3433
Eat less	7.25 (0.923)	8.29 (1.886)	7.04 (1.512)	6.39 (1.316)	0.34	0.71	3433
Sterilise drinking water	41.0 (2.076)	39.0 (3.651)	42.9 (3.336)	41.1 (3.760)	0.32	0.73	3433
Avoid raw fruit	10.9 (1.241)	11.6 (2.296)	10.6 (1.877)	10.6 (2.240)	0.077	0.93	3433
Cow dung on the floor	8.30 (1.001)	8.38 (1.808)	8.09 (1.652)	8.44 (1.731)	0.012	0.99	3433
Causes diarrhoea: Nothing, it's God's will	1.46 (0.296)	1.97 (0.616)	1.23 (0.404)	1.15 (0.478)	0.63	0.54	3433
Avoid spicy food	20.0 (1.891)	21.9 (3.622)	21.1 (3.425)	17.1 (2.612)	0.76	0.47	3433

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Woman Questionnaire.

5.5 Children in study sample (individual woman survey)

Children are possibly the members of the household who stand more to gain from improved sanitation since they are more vulnerable to disease and also are more severely affected by it, which is why we dedicated a set of questions focuses on young children in the household. Our examination of this elicited information completes our analysis of data collected through household and individual level questionnaires. We provide an overview of the child's health, how it is fed, the manner its hygiene is dealt with for the youngest child bore by the interviewee, and anthropometrics for all children below two years of age.

5.5.1 General characteristics children

We begin by summarising the profile of the female respondent in terms of her fertility (Table 103). The proportion of women with children less than six years of age is 54%, and on average they currently have only one child. Only 44% of the female respondents have babies younger than 2 years. The age of the youngest child is about 19 months on average and 55% of them are boys. There is a slight imbalance in the fraction of households with young children and in their number, which echoes the finding from the household data where the same occurs. Women in the treatment group GK have significantly fewer children and in a lesser proportion in statistical terms, however, the differences in practice are only of a maximum of four percentage points among groups. As it has been consistently the case with the rest of the data, the descriptives for children are for the most part statistically equivalent for the three comparison groups with very few exceptions to note.

Table 103: Children in the household

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Children <6 years	54.3 (0.921)	56.6 (1.408)	52.0** (1.400)	54.2 (1.876)	2.66	0.074*	3567
Nr children <6 years	0.85 (0.0162)	0.88 (0.0219)	0.80** (0.0270)	0.87 (0.0333)	2.81	0.064*	3567
Babies(<2 years)	43.9 (0.871)	45.4 (1.501)	42.1* (1.304)	44.1 (1.657)	1.43	0.24	3567
Nr babies(<2 years)	0.49 (0.0113)	0.50 (0.0197)	0.47 (0.0204)	0.49 (0.0179)	0.55	0.58	3567
Age child (months)	18.6 (0.393)	18.0 (0.627)	18.7 (0.662)	19.1 (0.741)	0.68	0.51	1936
Gender youngest child (fraction male)	55.3 (1.169)	55.2 (1.770)	58.2 (1.905)	52.4 (2.272)	1.97	0.14	1936

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Woman Questionnaire.

5.5.2 Health

Most women with children believe that their youngest offspring are as healthy as any other child of his/her age (60%) or for 33% of respondents even healthier than average (Table 104). As seen in Table 105, almost all children have received the BCG vaccine against tuberculosis, the three doses against polio and the DPT to avoid infection by diphtheria, pertussis or tetanus. The proportion of vaccinated children drops against measles, and almost none of them have been administered any of the doses of the pentavalent vaccine (which provides additional protection against hepatitis B and diphtheria).

Table 104: Child health status

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Healthier	32.5 (1.869)	32.6 (3.195)	31.5 (3.325)	33.5 (3.173)	0.10	0.90	1936
Same health level	59.8 (1.885)	58.0 (3.183)	62.3 (3.275)	59.4 (3.298)	0.47	0.63	1936
Less healthy	7.39 (0.681)	9.04 (1.262)	5.87* (1.088)	7.03 (1.082)	1.81	0.17	1936

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
 Source: Woman Questionnaire.

Table 105: Child vaccination

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
BCG, 1st dose at zero months	99.0 (0.233)	99.0 (0.410)	99.0 (0.368)	98.9 (0.424)	0.035	0.97	1930
POLIO, doses 1-3 at 1.5-3.5 months	97.2 (0.384)	97.0 (0.742)	98.0 (0.523)	96.6 (0.660)	1.59	0.21	1929
DPT, doses 1-3	90.7 (0.801)	89.7 (1.339)	91.6 (1.392)	91.0 (1.421)	0.51	0.60	1921
MEASLES, 1 dose	70.3 (1.322)	68.6 (2.444)	70.3 (2.153)	72.2 (2.153)	0.62	0.54	1923
Pentavalent, 1 dose	0.17 (0.0993)	0.17 (0.164)	0.18 (0.181)	0.17 (0.172)	0.0025	1.00	1726

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
 Source: Woman Questionnaire.

According to the descriptives in Table 106 not less than 1 out of 4 children in each of our three study groups suffered from fever in the last 7 days. One out of ten of the children had had periods of vomiting and 6% of the kids had had diarrhoea. These findings bear witness to how vulnerable children are to disease in the study area. Diarrheic children had to endure it for close to four days out of the seven in the week, and 80% of their mothers sought advice about how to treat their condition.

Table 106: Child diseases

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Vomiting	12.9 (1.421)	15.2 (3.055)	12.3 (2.177)	11.1 (1.692)	0.72	0.49	1931
Fever	24.9 (1.677)	26.4 (3.200)	25.9 (2.726)	22.3 (2.587)	0.67	0.51	1931
Skin rashes	3.68 (0.827)	5.60 (1.773)	3.61 (1.485)	1.60** (0.539)	2.89	0.060*	1930
Itching sores on feet and legs	3.11 (0.600)	3.88 (1.224)	3.11 (1.022)	2.24 (0.737)	0.72	0.49	1930
Indigestion	4.04 (0.629)	3.45 (1.034)	4.75 (1.263)	4.01 (0.953)	0.31	0.73	1930
Stomach Pain	4.05 (0.618)	3.88 (1.080)	4.10 (1.317)	4.17 (0.733)	0.024	0.98	1928
Unusual tiredness	2.75 (0.547)	3.59 (1.187)	2.62 (0.819)	1.93 (0.646)	0.81	0.45	1929
Unusually paleness	2.85 (0.574)	3.16 (1.124)	3.11 (1.044)	2.25 (0.724)	0.36	0.70	1929
Diarrhoea	5.86 (0.626)	5.19 (0.964)	6.72 (1.212)	5.78 (1.061)	0.49	0.61	1927
Nr of days diarrhoea last week	3.71 (0.169)	3.53 (0.343)	3.71 (0.254)	3.89 (0.274)	0.34	0.71	113
Sought advice diarrhoea	79.6 (4.435)	69.4 (8.425)	85.4 (6.982)	83.3 (6.794)	1.17	0.32	113

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Woman Questionnaire.

5.5.3 Nutrition

Table 107 starts by showing the fraction of two-year olds who are currently breastfed, which accounts for 88% of the sample. The remainder were fed with their mother's milk in the past but not anymore. Children typically drink plain water (79%), and watery items, like dal or rice water. 24% of them are also given other types of milk, such as, tinned or animal milk. Commercially produced liquids are hardly consumed by children in the area of study, with little more than 5% of mothers feeding them to their babies.

Table 107: Child nutrition: Liquids

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Not breastfed anymore	11.3 (1.405)	9.78 (2.360)	12.6 (2.218)	11.9 (2.670)	0.38	0.68	1527
Currently breastfed	88.2 (1.414)	89.7 (2.399)	87.2 (2.239)	87.5 (2.660)	0.31	0.73	1527
Never breastfed	0.46 (0.168)	0.54 (0.300)	0.21 (0.208)	0.60 (0.334)	0.71	0.49	1527
Plain Water	78.8 (1.371)	78.0 (2.347)	81.8 (2.112)	76.9 (2.534)	1.28	0.28	1521
Commercially produced infant formula	5.07 (0.697)	5.81 (1.428)	4.61 (1.117)	4.67 (0.933)	0.26	0.77	1520
Watery items (e.g Dal/Rice water)	52.7 (1.451)	51.7 (1.980)	57.0 (2.799)	49.6 (2.672)	1.97	0.14	1520
Tinned/powdered/animal milk	23.7 (1.752)	24.5 (3.104)	22.0 (2.741)	24.5 (3.168)	0.25	0.78	1521
Juice or juice drinks	5.14 (1.358)	7.27 (2.718)	3.14 (1.066)	4.68 (2.630)	1.06	0.35	1518
Butter milk/beaten curd	5.07 (1.179)	6.38 (1.914)	3.77 (1.324)	4.88 (2.619)	0.63	0.54	1518
Any other liquids	57.6 (10.21)	46.7 (14.27)	72.7 (15.90)	57.1 (22.86)	0.70	0.51	33

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Woman Questionnaire.

The diet of the children in our sample (Table 108) seems to be dominated by some specific items, like grains (e.g. bread, roti, chapati, rice, kitchdi, noodles) consumed by 63% of them, any kind of milk (38%), leafy vegetables (18%) and fruit (15%). About half of them also get to eat biscuits and other sweets

Table 108: Child nutrition: Food

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Any commercially fortified baby food	9.62 (1.024)	9.07 (1.700)	11.6 (2.014)	8.35 (1.569)	0.83	0.44	1517
Any bread/rice/roti/other food from grains	63.0 (1.751)	63.9 (2.682)	64 (3.278)	61.2 (3.156)	0.27	0.77	1518
Any pumpkin, carrots, or sweet potatoes	6.19 (1.245)	7.08 (2.106)	5.26 (1.369)	6.10 (2.717)	0.27	0.77	1518
Any white potatoes, white yams, cassava	6.65 (1.273)	6.72 (2.405)	7.16 (1.535)	6.10 (2.444)	0.069	0.93	1518
Any dark green, leafy vegetables	18.1 (1.514)	19.1 (2.567)	17.3 (2.261)	17.7 (2.952)	0.14	0.87	1518
Any mangoes/papayas/cantaloupe/jackfruit	6.06 (1.103)	6.17 (2.057)	7.16 (1.514)	4.88 (2.021)	0.41	0.66	1518
Any other fruits or vegetables	15.0 (1.255)	14.2 (2.013)	15.8 (1.939)	15.0 (2.518)	0.17	0.85	1517
Is the child non-vegetarian	3.16 (0.537)	4 (1.038)	3.16 (1.017)	2.24 (0.625)	1.14	0.32	1517
Any meat	1.45 (0.364)	2.00 (0.739)	1.26 (0.568)	1.02 (0.514)	0.59	0.55	1516
Any eggs	1.25 (0.339)	1.45 (0.650)	1.26 (0.569)	1.02 (0.512)	0.15	0.87	1518
Any fresh or dried fish or shellfish	0.73 (0.264)	0.91 (0.529)	0.63 (0.352)	0.61 (0.443)	0.11	0.89	1516
Any pulses/lentils/beans or other legumes	4.48 (0.803)	4.54 (1.424)	5.89 (1.475)	3.05 (1.223)	1.12	0.33	1518
Any nuts or foods made from nuts	9.29 (1.222)	9.09 (2.206)	10.7 (1.958)	8.13 (2.119)	0.42	0.66	1517
Any milk	37.6 (2.135)	38.5 (3.748)	37.7 (3.437)	36.6 (3.829)	0.063	0.94	1518
Any cheese, yoghurt or other milk products	11.7 (1.906)	11.3 (3.344)	13.3 (3.222)	10.6 (3.292)	0.19	0.83	1518
Any food made with oil, fat, ghee or butter	15.6 (2.188)	15.8 (3.929)	17.7 (3.576)	13.4 (3.727)	0.34	0.71	1518
Any biscuits or other sweets	54.3 (1.662)	52.1 (2.631)	54.9 (2.670)	56.1 (3.295)	0.53	0.59	1518
Any other solid or semi-solid food	13.1 (1.747)	16.7 (3.350)	13.9 (2.999)	8.13** (2.334)	2.55	0.082*	1517

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Woman Questionnaire.

We also asked mothers of the youngest children about their care and feeding practices when the child contracts diarrhoea (Tables 109 and 110). Their answers reveal that contrarily to the usual medical advise, children drink commonly the same (37%) or even less (40%), instead of drinking more to rehydrate the loss of liquids. In a similar way, they eat less (39%) or about the same (37%).

Table 109: Diarrhoea: Drinking

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Child drank: Much less	1.00 (0.288)	0.58 (0.351)	0.83 (0.533)	1.64 (0.591)	1.18	0.31	1902
Child drank: Somewhat less	40.5 (2.188)	41.9 (3.715)	40.2 (4.062)	39.1 (3.560)	0.15	0.86	1902
Child drank: About the same	36.9 (2.445)	35.8 (3.951)	41.2 (4.233)	33.9 (4.415)	0.79	0.45	1902
Child drank: More	0.53 (0.159)	0.44 (0.242)	0* (0)	1.15 (0.395)	5.83	0.0038***	1902
Child drank: Nothing to drink	5.99 (1.109)	5.53 (1.580)	6.13 (2.055)	6.38 (2.146)	0.058	0.94	1902
Child drank: Not Applicable	13.0 (1.448)	14.1 (2.460)	9.44 (1.963)	15.2 (2.854)	1.85	0.16	1902
Child drank: Don't know	2.10 (0.514)	1.60 (0.689)	2.15 (0.820)	2.62 (1.129)	0.33	0.72	1902

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Woman Questionnaire.

Table 110: Diarrhoea: Food

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Child ate: Much less	1.00 (0.270)	0.87 (0.330)	1.16 (0.574)	0.98 (0.497)	0.095	0.91	1902
Child ate: Somewhat less	38.9 (2.124)	38.9 (3.722)	39.2 (4.045)	38.6 (3.195)	0.0071	0.99	1902
Child ate: About the same	37.1 (2.024)	34.5 (3.338)	41.9 (3.528)	35.4 (3.513)	1.35	0.26	1902
Child ate: More	2.05 (0.827)	3.49 (1.810)	2.15 (1.518)	0.33* (0.228)	2.17	0.12	1902
Child ate: Nothing to eat	1.79 (0.355)	1.89 (0.576)	1.16 (0.477)	2.29 (0.743)	0.99	0.38	1902
Child ate: Not Applicable	16.4 (1.525)	18.5 (2.559)	11.8** (2.220)	18.7 (2.922)	2.68	0.073*	1902
Child ate: Don't know	2.73 (0.609)	1.89 (0.760)	2.65 (1.043)	3.76 (1.311)	0.79	0.46	1902

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Woman Questionnaire.

5.5.4 Hygiene

Following the pattern of the rest of the survey, which has elicited comprehensive information about sanitation through all its survey instruments, mothers of children below five were asked in detail about how they dealt with their child's hygiene.

Table 111 informs us about the type of diaper used by children and how their mothers handle them. The majority of children (55%) wear cloth nappies to retain stools, as opposed to cotton (18%) or disposable diapers (2%), and 24% do not wear anything. 84% of those mothers using cloth nappies usually wash them with soap (in 83% of the cases) and (or) water. This water is typically thrown away along with other rubbish (46%) or rinsed into a drain (37%).

Table 111: Child hygiene: Diaper

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Cloth nappy	54.6 (2.379)	55.6 (4.289)	54.4 (4.257)	53.7 (3.702)	0.058	0.94	1903
Cotton nappy	18.8 (1.874)	16.2 (3.000)	23.5 (3.574)	17.0 (3.059)	1.39	0.25	1903
Disposable nappy/diaper	1.84 (0.504)	2.47 (1.080)	2.48 (0.918)	0.49* (0.272)	3.50	0.033**	1903
Nothing	24.8 (2.329)	25.8 (4.167)	19.7 (3.696)	28.6 (4.055)	1.42	0.25	1903
Washed nappy: Water	70.6 (2.961)	69.4 (4.852)	70.2 (5.475)	72.6 (5.080)	0.11	0.90	1039
Washed nappy: Soap / soap powder	83.9 (2.099)	84.0 (3.284)	86.6 (3.264)	81.1 (4.304)	0.53	0.59	1039
Washed nappy: Bleaching powder	6.06 (1.405)	3.66 (1.459)	11.2* (3.548)	3.66 (1.647)	2.08	0.13	1039
Washed nappy: Other	0.19 (0.136)	0.26 (0.262)	0.30 (0.300)	0 (0)	1.01	0.37	1039
Disposal water: Thrown toilet pit	10.6 (1.179)	9.69 (1.915)	12.5 (2.383)	9.76 (1.768)	0.51	0.60	1039
Disposal water: Thrown in field	0.67 (0.280)	0.79 (0.434)	0* (0)	1.22 (0.710)	3.11	0.048**	1039
Disposal water: Thrown along other rubbish	46.0 (3.160)	45.8 (5.837)	44.7 (5.231)	47.6 (5.108)	0.079	0.92	1039
Disposal water: Rinsed into drain/ditch	37.1 (2.783)	36.9 (4.999)	35.9 (4.728)	38.4 (4.603)	0.075	0.93	1039
Disposal water: Other	5.68 (1.615)	6.81 (3.002)	6.99 (3.385)	3.05 (1.274)	1.10	0.34	1039

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Woman Questionnaire.

Table 112 shows that most of the children under the age of two years (85%) did not use neither a toilet nor a potty last time they defecated. Lack of toilet use among babies is more common in the GK+ND area than in the other two by a non-negligible margin. To dispose of the stools mothers threw them in the garbage (43%) or left them in the open (24%). Few took the precaution of having them rinsed into the a drain or toilet (19%) or burying them (20%). The stools were handled with a cloth or leaves 41% of the times, 38% used scrap material for the purpose and 24% just took them with their hands to remove them.

After the child defecates, Table 113 tells us that the mother cleaned the child's bottom almost always. After doing this the majority - but not all - of our respondents reported to wash their hands every time, with soap (73%) or at least water (26%).

Table 112: Child hygiene: Defecation

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Child used toilet or latrine	5.68 (0.692)	5.39 (1.170)	5.45 (1.132)	6.22 (1.277)	0.14	0.87	1903
Child used child potty	9.56 (1.501)	15.3 (2.964)	8.43* (2.532)	4.26*** (1.563)	5.60	0.0047***	1903
Child did not use toilet/latrine/potty	84.8 (1.654)	79.3 (3.286)	86.1 (2.707)	89.5*** (2.027)	3.50	0.033**	1903
Disp stools: Child used toilet or latrine	4.78 (0.599)	5.09 (1.000)	4.46 (0.945)	4.75 (1.156)	0.11	0.90	1903
Disp stools: Put/rinsed into toilet or latrine	4.26 (0.708)	5.09 (1.382)	4.79 (1.345)	2.78 (0.776)	1.53	0.22	1903
Disp stools: Put/rinsed into drain or ditch	14.8 (1.308)	12.7 (1.683)	16.7 (2.506)	15.2 (2.571)	0.99	0.37	1903
Disp stools: Thrown into garage	42.5 (2.064)	43.8 (3.438)	38.3 (3.151)	45.0 (3.967)	1.10	0.34	1903
Disp stools: Buried	10.3 (1.841)	10.3 (3.537)	10.4 (2.898)	10.1 (2.966)	0.0021	1.00	1903
Disp stools: Left in the open	22.8 (1.578)	22.4 (2.624)	24.8 (2.567)	21.3 (2.956)	0.44	0.65	1903
Handling stools: Hands only/bare hands	23.9 (2.251)	23.4 (3.856)	25.6 (4.450)	22.6 (3.322)	0.15	0.86	1903
Handling stools: Cloth/paper/leaves	41.1 (2.346)	41.6 (4.057)	39.7 (4.265)	41.9 (3.839)	0.086	0.92	1903
Handling stools: Scrap material to scoop feces	38.3 (2.352)	39.4 (4.358)	38.3 (4.267)	36.8 (3.407)	0.12	0.89	1903
Handling stools: Did not handle the feces	14.9 (1.290)	14.6 (2.053)	16.2 (2.543)	14.1 (2.075)	0.22	0.80	1903
Handling stools: Other	0.42 (0.246)	0 (0)	0.83 (0.681)	0.49 (0.358)	1.67	0.19	1903

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Woman Questionnaire.

Table 113: Child hygiene: After defecation

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Cleaned bottom	93.3 (0.931)	91.7 (1.706)	95.2 (1.464)	93.3 (1.551)	1.24	0.29	1903
No one cleaned child's bottom	4.73 (0.785)	6.40 (1.465)	2.64** (1.139)	4.91 (1.321)	2.19	0.12	1903
Child cleaned bottom him/herself	1.94 (0.399)	1.89 (0.719)	2.15 (0.785)	1.80 (0.530)	0.068	0.93	1903
Washed hands after the cleaning	99.0 (0.268)	98.8 (0.510)	99.0 (0.498)	99.2 (0.350)	0.16	0.85	1902
Wash hands: Soap	72.9 (2.813)	70.7 (5.054)	80.6 (3.685)	67.7 (5.254)	2.49	0.087*	1883
Wash hands: Water	26.2 (2.809)	28.3 (5.074)	18.2 (3.674)	31.7 (5.207)	2.67	0.073*	1883

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Woman Questionnaire.

5.5.5 Anthropometrics

In this section we present objective measures with the aim of assessing the health and nutritional status of children under two years of age. In order to construct these indices we collected anthropometric data for the child's weight and height employing Seca scales and following World Health Organisation (WHO) recommendations. General descriptives for all the children who could be measured by enumerators (in some cases belonging to the same household and having the same mothers) appear in Table 114. On average these infants are a bit over 1 year old (13.5 months of age), they weigh 8 kilograms and are 72 centimetres tall. Table 114 also contains information on three important measures to evaluate the health and nutritional state of the child population in our sample, these are:

- **Weight for age** (Underweight). This is probably the most common assessment of child nutrition status. It represents a suitable combination of both linear growth and body proportion and thus can be used for the diagnosis of underweight children.
- **Weight for height** (Wasting). This is a measure of current body mass. It is generally seen as a measure of acute or short-term inadequate nutrition and/or poor health status. It is the best index to use to reflect wasting malnutrition, when it is difficult to determine the exact ages of the children being measured.
- **Height for age** (Stunting). This is a measure of linear growth. Stunting refers to shortness. A deficit in height for age is generally assumed to indicate exposure to an unhealthy environment, such as poor nutrition, unhygienic environment or disease in the past and hence captures long-term, cumulative effects.

More specifically, we look at the z-scores of these measures. A z-score describes how much a point (such as the weight for height for a specific child) deviates from a reference point. The reference point is in this case the WHO Child Growth Standards. Details on these standards and how they were constructed can be found in WHO Multicentre Growth Reference Study Group (2006,2007).²⁹ Looking at the summary statistics of the z-scores of the variables has two purposes. First, we can simultaneously look at whether the sample is balanced between treatments and control with respect to these measures. Second, the summary statistics give us an idea of the nutritional status of the sample population we are considering. With regard to the former aim, we do not find any imbalances among the z-scores for the three groups. Furthermore, considering also the hypotheses tests carried out in Table 115, this extends to all the anthropometric data and derived indices, leading us to conclude that children are statistically similar across groups in anthropometric terms. We will also employ the z-scores to draw conclusions about how healthy infants are in the sample area. As we can see, the mean of the distributions for weight for age as well as weight for height are below zero. This implies

²⁹Following the guidelines of these studies, we removed those observations deemed extreme (i.e. biologically implausible) and, therefore likely a byproduct of measurement error, using their established criteria before estimating the z-scores.

that in our sample most children are affected by health and nutrition problems to a certain extent. We continue investigating this below.

Table 114: Anthropometrics

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Age (months)	13.5 (0.195)	13.4 (0.322)	13.5 (0.317)	13.4 (0.375)	0.055	0.95	1449
Weight (kg)	8.12 (0.0628)	8.05 (0.107)	8.19 (0.105)	8.13 (0.112)	0.40	0.67	1449
Height (in cm)	72.3 (0.266)	72.0 (0.482)	72.7 (0.385)	72.2 (0.483)	0.69	0.50	1449
Body Mass Index (BMI)	15.4 (0.0738)	15.3 (0.115)	15.4 (0.129)	15.5 (0.139)	0.46	0.63	1449
Weight-for-height z-score	-0.89 (0.0550)	-0.93 (0.0827)	-0.95 (0.108)	-0.79 (0.0944)	0.85	0.43	1449
BMI for age z-score	-0.81 (0.0569)	-0.86 (0.0890)	-0.85 (0.108)	-0.71 (0.0977)	0.71	0.49	1449
Height-for-age z-score	-1.24 (0.0668)	-1.30 (0.118)	-1.17 (0.112)	-1.24 (0.113)	0.32	0.73	1449
Weight-for-age z-score	-1.34 (0.0437)	-1.41 (0.0715)	-1.33 (0.0818)	-1.26 (0.0717)	0.99	0.37	1449

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Woman Questionnaire.

Table 115 displays the relevant descriptives. Prevalence based reporting is done using a cut-off point, which is consistent with clinical screening. Prevalence of malnutrition (measured by weight for age) is for example the percentage of children in the age range of our infants (0 to 2 years) whose weight for age measure is more than two standard deviations below the median of the international reference population for the same age group. For the measure weight for age, the WHO Global Database on Child Growth and Malnutrition uses values of less than three standard deviations below the norm ($< -3SD$) as severe undernutrition and a value of less than two standard deviations below the norm ($< -2SD$) as low weight for age. Values of less than two standard deviations for height for age and weight for age are considered as moderate to severe.

Knowing this, we can now scrutinise the data and conclude that the messages conveyed by the z-scores statistics are confirmed by the prevalence data. 20% of children in our sample are moderately to severely undernourished, and almost 9% suffer from severe undernutrition. The seemingly poor health situation of infants on our study area is further backed by the high incidence of low height-for-age in the sample, with 18% of the children moderately stunted and as many as 14% severely so. Consistently with this, a similar proportion of children are wasted or severely wasted.

Table 115: Prevalence variables

	Treatment Status				F-stat	P-value	N
	Whole Sample	Control	GK	GK + ND			
Underweight (-3 SD <weight-for-age <-2 SD)	19.9 (1.223)	22.3 (2.075)	18.5 (2.181)	18.6 (1.985)	1.09	0.34	1449
Severely underweight (weight-for-age <-3 SD)	8.90 (0.851)	10.1 (1.190)	9.47 (1.866)	6.93* (1.319)	1.70	0.19	1449
Stunted (-3 SD <length-for-age <-2 SD)	18.2 (1.222)	18.8 (2.178)	18.9 (2.187)	16.7 (1.912)	0.40	0.67	1449
Severely stunted (length-for-age <-3 SD)	14.3 (1.190)	15.6 (1.989)	13.4 (1.853)	13.6 (2.297)	0.35	0.70	1449
Wasted (-3 SD <weight-for-height <-2 SD)	15.3 (1.097)	15.6 (1.766)	16.3 (1.954)	14.1 (1.974)	0.34	0.72	1449
Severely wasted (weight-for-height <-3 SD)	7.18 (0.751)	7.32 (1.224)	8.37 (1.612)	5.84 (0.993)	1.03	0.36	1449

Note: Standard Errors in parenthesis, clustered at the gram panchayat, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Woman Questionnaire.

6 Conclusions

This report discussed the activities and findings of the baseline data collection for this RCT component of the project titled “Understanding the Links and Interactions between Low Sanitation and Health Insurance in India”, funded through the Strategic Impact Evaluation Fund (SIEF).

We provided an overview of data collection activities, including a description of challenges faced and how they were met. We have further shown formal tests comparing all important characteristics collected at baseline, across treatments and control. This is an important exercise since it allows us to see whether, indeed, the randomisation was successful at creating three groups (two treatment and one control) that appear similar on all dimensions. The only difference between these groups will then be that treatment arms are eligible for and do receive the intervention.

Our results from the treatment versus control comparison are generally very encouraging. The number of imbalances found fall well within the expected proportion of ten percent and those imbalances observed are for the most part significant only at the 5 or mostly at the 10% significance level. We further find no noteworthy significant differences in key variables across treatment and control units, particularly in the sanitation situation and health outcomes. This is important since it implies that our treatment and control households were not significantly different in terms of their sanitation practices and starting level of their health status before the start of the intervention.

Despite these encouraging overall results, it is worth pointing out that we observe more imbalances between the GK treatment arm and the control group than we do between the GK+ND treatment arm and the control group. Findings suggest that GK communities are somewhat better off with respect to some assets they own. Households in GK communities are for example more likely to own their dwelling, the dwelling type owned is more likely to be of strong (pucca) structure, and they own more land than households in the control communities. It is likely that we observe a number of imbalances due to correlations between some of the variables. Likely related to the three variables just mentioned is for example the length of time households lived in a village.

In line we find that one of the other significant differences observed is that households in the GK treatment arm are more likely to have been born in the village they now live in. We also note though that, while households in GK communities seem to be richer in terms of some assets, they do report to earn the same levels of income and have similar yearly consumption expenditures. These results suggest that households residing in the GK treatment arm are not necessarily systematically richer.

In any event, it will be important to account for characteristics for which imbalances are observed when we conduct the impact analysis and check for robustness of findings to the in- and ex-clusion of these characteristics.

References

- A.D. Dangour, L. Watson, O. Cumming, S. Boisson, and et al. Che, Y. Interventions to improve water quality and supply, sanitation and hygiene practices, and their effects on the nutritional status of children. *Cochrane Database Syst Rev* 8: CD009382., 2013.
- GoI. Swach bharat mission (gramin) guidelines, 2014. URL http://www.mdws.gov.in/sites/upload_files/ddws/files/pdf/SwachBharatGuidlines.pdf.
- WHO Multicentre Growth Reference Study Group. 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., 2006.
- WHO Multicentre Growth Reference Study Group. Child growth standards: Head circumference-for-age, arm circumference-for-age, triceps skinfold-for-age and subscapular skinfold-for-age: Methods and development. 2007.
- A. Hueso and B. Bell. An untold story of policy failure: the total sanitation campaign in india. *Water Policy*, 15(1001-17), 2013.
- JH Humphrey. Child undernutrition, tropical enteropathy, toilets, and handwashing. *Lancet*, 374:1032–1035, 2009.
- M.W. Jenkins and B. Scott. Behavioral indicators of household decision-making and demand for sanitation and potential gains from social marketing in ghana. *Soc. Sci. Med.*, 64(12):2427–2442, 2007.
- M.W. Jenkins and Sugden. Rethinking sanitation: Lessons and innovations for sustainability in the new millennium., 2006.
- S. Kumar and S. Vollmer. Does access to improved sanitation reduce childhood diarrhea in rural india. *Health Economics*, 2012.
- S.K. Pattanayak, J.C. Yang, K.L. Dickinson, C. Poulos, S.R. Patil, R.K. Mallick, J.L. Blitstein, and P. Praharaj. Shame or subsidy revisited: social mobilization for sanitation in orissa, india., 2009.

- A. Pruss-Unstun, R. Bos, F. Gore, and J. Bartram. Safer water, better health: costs, benefits and sustainability of interventions to protect and promote health., 2008.
- R. Raman and S. Tremolet. Financing on-site sanitation for the poor: A six country comparative review and analysis, 2010. URL http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2010/10/07/000333038_20101007004230/Rendered/PDF/569430WP0finan101public10BOX353742B.pdf.
- D. Spears. Effects of rural sanitation on infant mortality and human capital: Evidence from india’s total sanitation campaign, 2012.
- D. Spears. The nutritional value of toilets: How much international variation in child height can sanitation explain? *RICE Working Paper*, 2013. URL <http://riceinstitute.org/wordpress/wpcontent/uploads/downloads/2013/07/Spearsheight-and-sanitation-6-2013.pdf>.
- A. Tarozzi, A. Mahajan, B. Blackburn, D. Kopf, L. Krishnan, and J. Yoong. Micro-loans, bednets and malaria: Evidence from a randomized control trial in orissa (india)., 2011.
- UNICEF. Impact assessment of nirmal gram puraskar awarded panchayats: Final report, 2008.
- WSP (Water and Sanitation Program). The economic impacts of inadequate sanitation in india, 2010. URL <http://www.zaragoza.es/contenidos/medioambiente/onu/765-eng.pdf>.
- WHO-UNICEF. Progress on drinking water and sanitation, 2014. URL http://www.unicef.org/media/files/JMP_2014_Update.pdf.

A Power analysis

The key outcome discussed in the power analysis in the submitted proposal was the number of health insurance claims. We will not be able to update this analysis until the health insurance component is implemented. We will therefore discuss here only the power analysis related to sanitation uptake.

The exact variable we consider is the percentage of households owning a toilet. We will use the same approach as followed for and presented in the proposal, computing minimum detectable effects (MDEs) for a power of 0.8 and 0.9, and a significance level of 0.05, assuming two-sided t-tests and a sample of 80 communities (40 treatment and 40 control) and 30 interviewed households within each community. At the time of proposal-writing, we assumed a baseline sanitation uptake of 0.15 percent. Since we had only indicative information for the intra-cluster correlation (ICC) of the outcome, we presented three scenarios, one with an ICC of 0, one with an ICC of 0.15 and the third with an ICC of 0.3. The MDEs we calculated under these assumptions are >5.7 percentage points (ppt), 11-13ppt and 15-18 ppt for the respective ICCs. These numbers can be seen in Table 116 (columns 1a, 1b, and 1c).

Table 116: Updated power analysis - sanitation uptake

	(1a)	(1b)	(1c)	(2)	(3)	(4)
	Proposal			Full	GK	Non-
	Assumption:			sample	clients	clients
	1	2	3			
Sample size	2,400	2,400	2,400	2,400	1,200	1,200
Clusters	80	80	80	80	80	80
Power	0.8-0.9	0.8-0.9	0.8-0.9	0.8-0.9	0.8-0.9	0.8-0.9
Mean of control BL	0.15	0.15	0.15	0.28	0.245	0.316
ICC - raw outcome	0	0.15	0.3	0.143	0.173	0.147
ICC - residual				0.105	0.116	0.096
MDE (with lowest ICC) - ppt	>5.7	11-13	15-18	10.9-12.3	12.0-14.3	12.0-14.0

Notes: The Table shows statistics for the power analysis, repeating numbers presented in the proposal (columns "Proposal", under three different scenarios) and showing updated numbers for different samples: The full sample, GK client households only, and Non-client households only. 'ICC' stands for Intra-Cluster-Correlation, 'ICC - raw outcome' refers to the ICC calculated on the original baseline data, 'ICC - residual' refers to the ICC calculated based on the residual of a regression of the outcome on a number of household-level characteristics. 'MDE' stands for minimum detectable effect.

With the baseline data at hand we are able to refine this analysis, being able to calculate the actual baseline mean of our outcome variable of interest in the study area, and calculating the ICC. We present two ICCs, one calculated on the raw data, the second calculated on the residual of a regression of the outcome on a number of household characteristics³⁰. We use this latter ICC in the calculation of MDEs, as it

³⁰The household characteristics used in the regression analysis are the household's religion, caste,

is a more precise measure, increasing our power. The updated mean that relates to the numbers presented in the proposal can be seen in column (2) of Table 116. The actual baseline mean is higher than was anticipated, namely 28% of households own a toilet in the study area. The raw ICC is 0.143, which is reduced to 0.105 by taking household characteristics into account. Given this mean and ICC and keeping all other assumptions constant, we will be able to detect increases of about 10.9-12.3 percentage points. This assumption compares well with our preferred scenario discussed in the proposal.

Table 116 presents two additional scenarios: One where we look only at the sample of GK clients and the other where we look at non-clients only. It is worth noting that the changes we are able to detect for these sub-samples are about 1-2 percentage points larger than those using the whole sample, which is driven by the reduced sample size. For the final analysis, we will calculate weights that allow us to make use of the full sample, accounting for the fact that we have two distinct sub-sample groups.

We further note that we will also be able to pool the two treatment arms and thereby increasing our sample size (and hence power). Under this scenario we have 80 treatment communities and 40 control. With this increased sample size, the MDEs we are able to detect decrease by roughly 1-2 percentage points. However, we note an important caveat: The two treatments are likely to have differential impacts, which would imply different variances in our two treatment samples, which would change the exact MDE we are able to detect.

B More details on the segmentation procedure

The household listing exercise involved preparing up-to-date location and layout sketch maps (each of which are available), assigning and marking a number to each structure in the GP, recording addresses of these structures, identifying residential structures, and then listing the names of heads of households and collecting other information (if required). Two different procedures were followed for segmentation, depending on the type of the sampled village:

1. Segmentation using *notional* segments (most common approach):
 - Step 1 - On the first day of a listing exercise in a particular GP, the field team reached the village, conducted a transect walk and identified the boundaries of the village (including the hamlets, which is a sub-division of the village). This exercise was done with the help of knowledgeable persons in the village, which included Panchayat Secretary members/ teachers and others.
 - Step 2 – After identification of the village boundary, the supervisor prepared a rough layout map of the village by marking the directions as north, south,

household composition, gender, marital status, age, main activity and education level of the household head, ownership status of the dwelling as well as some indicators about the dwelling such as materials of the floor and also the cooking material used by the household, and whether the households owns a benefit card or not.

east and west. While preparing the layout map, the supervisor again took help from local knowledgeable persons to identify different physical features (hillock, water body, stream), road network and important landmarks in the village.

- Step 3 – After preparing the layout map, the supervisor requested community members/knowledgeable persons in the village to advise them on how to segment the village in equal parts (based on latest available population data & growth rate).
 - Step 4 – After the segments were drawn on the map and the boundary/start point and end point of each segment were identified, the team went back to the field once again to confirm the segment and to identify more landmarks.
 - Step 5 – Once the team identified all segments with all identification details, sample segments got selected by following a simple random sampling method, conditional on having at least one segment with a GK kendra located in it (this information was provided by the GK office).
2. Segmentation using *natural* segments: In case of densely populated large villages with number of hamlets located outside the main village, an alternative method for segmentation was followed. In this case, segmentation makes use of the natural layout of the village (internal road network) and all segments may not be necessarily of same size. For such villages, the field team also conducted a transect walk and discussed with various knowledgeable persons in the village, identifying the segment based on the normal spread of the village. For example, in case of a village with 700 households, there might have been six segments with the following number of households: 200, 150, 150, 100, 50 and 50. In such case a systematic random sampling technique was followed to identify the study segments and effort were made to select the required number of segments so that the approximate number of households required for listing matches with the selection of clusters.

We refer to the training manual of the listing survey for more details on the segmentation procedure.

C More details on the baseline sampling strategy

Strategy followed in listed GPs

Sampling non-GK clients

For 107 out of the 120 sampled GPs we have listing data available which we could use to sample 15 households of non-GK clients, stratified by whether or not the household had kids less than 2 years old. Specifically, we randomly sampled 15 households (plus 8 back-up households) from the listed population excluding those clients which were successfully matched to GK clients. Given the fact that we did not manage to match all

clients, it is possible that some of the households in the pool from which we sampled non-client households actually have client members, and if we were unlucky it is possible that we sampled those. We do not expect this to have happened too often (given the small number of clients in the village relative to the village size) and if it happened we expect this to have happened as (un-)frequently in the control group as in the treatment group. This is something we can verify after (re-)matching households from the baseline survey to the GK client database, using the detailed information provided in the household roster. This exercise will take some time and had not yet been completed at the time of writing up this baseline report.

Sampling GK clients

Given imperfect matching of the GK clients to the listing data, an alternative sampling approach was adopted for the sampling of GK clients. In particular, the survey teams took a list of all GK clients (matched and unmatched) living in a particular GP to the GP office, where they met with GP officials and anganwadi workers³¹ in order to obtain details about the GK clients (without knowing that these were GK clients). Specifically, they collected information about (i) whether or not there were any kids below 2 years old or 5 years old living in the household of the client and (ii) about the location of the household. This information was pre-populated for those GK clients which had been successfully matched to the listing data.

The names on the list were sorted in a random order, such that after all information was collected the field supervisor could use the list to randomly sample GK clients by stratum (kids/no kids), starting from the top of the list and going down until the required number of clients in each stratum was reached. In GPs with more than 2 kendras the list started with listing, in a random order, all clients of two randomly sampled GPs. After this, the list continued with listing the clients of the remaining kendras in the GP (again in a random order). In segmented GPs and/or in GPs with more than one village and with more than 1 kendra, kendras with zero matches were listed at the end of the list (giving priority to kendras which are more likely to be located in the segmented/listed areas).

Strategy followed in non-listed GPs

In the 13 study GPs that were added to the study sample only after listing had taken place, the procedure for the selection of households of GK clients was exactly the same as described in the previous Section. That is, the supervisor identified respondents of client households following a randomised list of clients after consultation with GP officials and anganwadi workers.

However, given that no listing data were available for these 13 GPs, we could not follow the same sampling procedure as the one we followed in the 107 originally selected GPs for the selection of non-GK client households. Instead, the survey teams followed the following procedure for the selection of non-GK clients in those GPs:

³¹ Anganwadi workers are local welfare officers who provide child related health care and educational services to households in their communities, with the aim of combatting child malnutrition

- They first prepared a rough map of the GP in line with the maps prepared for all the original set of GPs when these were visited for listing (see Appendix B). In case there were more than 480 households in a GP (or sampled village within the GP), the GP was segmented. When doing the segmentation, the survey teams followed exactly the sample procedure as the one followed when segmenting villages for listing (see Appendix B).
- In small GPs (less than or equal to 480 households), the GP was divided in 4 notional clusters. Each of these four clusters was further divided into two sub-clusters to cover either households with kids or households without kids. The allocation of each sub-cluster to the kids/no kids stratum was done at random. In each sub-cluster, the total number of households was divided by the required sample to arrive at an interval. The team covered every n^{th} household starting from the north west point of the cluster. The starting household was selected using a computer generated random number. Thus, in each sub-cluster around two households, either with or without kids, was covered. In case there was not a sufficient number of households in a particular stratum (kids/no kids) available in a sub-cluster designated for that particular stratum, households belonging to that stratum were randomly sampled in sub-clusters that had originally been designated to the other stratum.
- In bigger GP (more than 480 households) notional segments of 120 households each were made and four segments were selected using a systematic random sampling procedure. The total sample of 15 sample households was divided equally in these four selected segments. Within each segment a similar procedure was followed to select households as that followed within a cluster in the smaller GP.

D Details study GPs

Table 117: Baseline details of study GPs

District	Branch	Block	GP	Treatment	# villages	Sampled village	# HHs	Listing	Segmented	# kendras	# clients	% toilet	% OD
Latur	Ahmadpur	Ahmadpur	Andhori	GK	1	Andhori	992	Yes	Yes	1	34	31%	49%
Latur	Ahmadpur	Ahmadpur	Belur	GK + ND	1	Belur	352	Yes	No	1	32	41%	42%
Latur	Ahmadpur	Ahmadpur	Chilkha	Control	1	Chilkha	213	Yes	No	1	34	10%	78%
Latur	Ahmadpur	Ahmadpur	Dhalegaon	GK	1	Dhalegaon	690	Yes	Yes	1	28	40%	46%
Latur	Ahmadpur	Ahmadpur	Hadolti	GK + ND	1	Hadolti	1637	Yes	Yes	4	139	33%	48%
Latur	Ahmadpur	Ahmadpur	Hipparga Kajal	Control	1	Hipparga Kajal	448	Yes	No	1	41	32%	64%
Latur	Ahmadpur	Ahmadpur	Kingaon	Control	1	Kingaon	2265	No	N/A	5	141	N/A	N/A
Latur	Ahmadpur	Ahmadpur	Lanji	GK	2	Lanji	328	Yes	No	1	30	35%	63%
Latur	Ahmadpur	Ahmadpur	Sangavi (Su)	GK + ND	2	Sangavi (Sunegaon)	381	No	N/A	1	38	N/A	N/A
Latur	Ahmadpur	Ahmadpur	Tembhurni	GK	1	Tembhurni	257	Yes	No	2	40	23%	68%
Latur	Ahmadpur	Ahmadpur	Thodga	Control	1	Thodga	563	Yes	No	3	77	6%	76%
Latur	Ahmadpur	Ahmadpur	Vilegaon	Control	2	Vilegaon	396	No	N/A	1	31	N/A	N/A
Latur	Nilanga	Nilanga	Ambulga Vi	Control	4	Ambulga	342	Yes	No	1	22	7%	74%
Latur	Nilanga	Nilanga	Ansarwada	GK + ND	1	Ansarwada	600	Yes	No	2	57	33%	56%
Latur	Nilanga	Nilanga	Dapka	GK + ND	2	Dapka	1054	No	N/A	2	70	N/A	N/A
Latur	Nilanga	Nilanga	Dhanora	Control	2	Dhanora	405	Yes	No	2	51	30%	59%
Latur	Nilanga	Nilanga	Hadga	Control	1	Hadga	570	Yes	Yes	1	28	38%	52%
Latur	Nilanga	Nilanga	Halgara	GK	2	Halgara	1344	Yes	Yes	5	103	26%	67%
Latur	Nilanga	Nilanga	Handral	GK + ND	1	Handral	129	Yes	Yes	1	40	52%	23%
Latur	Nilanga	Nilanga	Hanmantwadi Ab	GK	1	Hanmantwadi Ab	480	No	N/A	1	27	N/A	N/A
Latur	Nilanga	Nilanga	Harijawalga	GK	2	Harijawalga	374	No	N/A	1	40	N/A	N/A
Latur	Nilanga	Nilanga	Jajnoor	GK + ND	1	Jajnoor	453	Yes	No	2	53	17%	73%
Latur	Nilanga	Nilanga	Jau	GK	1	Jau	227	Yes	No	1	36	58%	36%
Latur	Nilanga	Nilanga	Jewari	GK + ND	3	Jewari	320	Yes	No	1	25	10%	71%
Latur	Nilanga	Nilanga	Kelgaon	GK + ND	1	Kelgaon	868	Yes	Yes	4	133	18%	73%
Latur	Nilanga	Ausa	Kharosa	Control	1	Kharosa	1246	Yes	Yes	4	119	14%	82%
Latur	Nilanga	Nilanga	Kokalgaon	GK	1	Kokalgaon	691	Yes	Yes	1	25	18%	56%
Latur	Nilanga	Nilanga	Lambota	Control	1	Lambota	742	Yes	Yes	1	39	22%	73%
Latur	Nilanga	Nilanga	Madansuri	GK	1	Madansuri	972	Yes	Yes	2	58	25%	27%
Latur	Nilanga	Nilanga	Makni	Control	1	Makni	466	Yes	No	1	47	35%	50%
Latur	Nilanga	Nilanga	Manathpur	GK + ND	1	Palapur	180	Yes	No	2	58	10%	73%
Latur	Nilanga	Nilanga	Nanand	Control	1	Nanand	860	Yes	Yes	1	40	12%	83%
Latur	Nilanga	Nilanga	Nitur	Control	6	Nitur	1445	Yes	Yes	4	95	16%	80%
Latur	Nilanga	Nilanga	Sarwadi	GK	1	Sarwadi	673	Yes	Yes	1	43	84%	6%
Latur	Nilanga	Nilanga	Sindhkhed	GK	1	Sindhkhed	421	Yes	No	1	45	15%	73%
Latur	Nilanga	Nilanga	Umarga H	GK + ND	1	Umarga H	545	Yes	No	1	23	22%	69%
Latur	Nilanga	Deoni	Walandi	GK	1	Walandi	1026	Yes	Yes	3	85	35%	60%
Latur	Nilanga	Nilanga	Yelnoor	Control	1	Yelnoor	447	Yes	No	1	45	10%	81%
Latur	Udgir	Udgir	Banshelki	GK	1	Banshelki	486	No	N/A	1	32	N/A	N/A
Latur	Udgir	Udgir	Bhakaskheda	Control	1	Bhakaskheda	499	Yes	No	1	25	63%	23%
Latur	Udgir	Udgir	Chandegaon	GK	1	Chandegaon	311	Yes	No	1	20	22%	60%
Latur	Udgir	Udgir	Dawangaon	GK + ND	1	Dawangaon	525	Yes	No	1	30	25%	45%
Latur	Udgir	Udgir	Dewarjan	GK + ND	3	Dewarjan	931	Yes	Yes	1	39	25%	55%
Latur	Udgir	Udgir	Ekurga Road	Control	1	Ekurga Road	335	Yes	No	1	20	31%	25%
Latur	Udgir	Deoni	Hanchanal	GK	1	Hanchanal	293	Yes	No	1	19	27%	65%
Latur	Udgir	Udgir	Kallur	GK + ND	1	Kallur	345	Yes	No	1	27	46%	15%
Latur	Udgir	Deoni	Kamalwadi	GK	1	Kamalwadi	286	Yes	No	2	46	26%	66%
Latur	Udgir	Udgir	Karkheli	GK + ND	1	Karkheli	263	Yes	No	1	26	46%	18%
Latur	Udgir	Udgir	Kiniyalladevi	GK + ND	1	Kiniyalladevi	486	Yes	No	1	23	32%	33%
Latur	Udgir	Udgir	Kumtha Kh	Control	1	Kumtha Kh	659	Yes	Yes	1	37	31%	38%
Latur	Udgir	Udgir	Lohara	Control	1	Lohara	1283	Yes	Yes	2	67	21%	53%
Latur	Udgir	Udgir	Loni	GK	1	Loni	765	No	N/A	5	134	N/A	N/A
Latur	Udgir	Udgir	Malkapur	GK	1	Malkapur	1059	Yes	Yes	2	55	67%	17%
Latur	Udgir	Udgir	Nagalgaon	GK	4	Nagalgaon	677	Yes	Yes	3	100	33%	30%
Latur	Udgir	Udgir	Nalgir	GK + ND	2	Nalgir	985	Yes	Yes	3	84	27%	35%
Latur	Udgir	Udgir	Netragaon	GK + ND	1	Netragaon	225	Yes	No	1	25	21%	40%
Latur	Udgir	Udgir	Nideban	GK + ND	1	Nideban	1024	Yes	Yes	2	53	65%	14%

Baseline details of study GPs - continued

District	Branch	Block	GP	Treatment	# villages	Sampled village	# HHs	Listing	Segmented	# kendras	# clients	% toilet	% OD
Latur	Udgir	Udgir	Pimpari	Control	1	Pimpari	335	Yes	No	1	30	49%	34%
Latur	Udgir	Udgir	Shekapur	Control	1	Shekapur	377	Yes	No	1	26	64%	24%
Latur	Udgir	Udgir	Shelhal	GK	1	Shelhal	467	Yes	No	1	29	55%	20%
Latur	Udgir	Udgir	Shirol	Control	1	Shirol	548	Yes	No	1	23	27%	62%
Latur	Udgir	Udgir	Somnathpur	Control	1	Somnathpur	1061	Yes	Yes	1	29	34%	59%
Latur	Udgir	Udgir	Wadhona (Bk)	GK	1	Wadhawana (Bk)	1713	No	N/A	3	68	N/A	N/A
Nanded	Degloor	Degloor	Achegaon	GK + ND	1	Achegaon	279	Yes	No	1	26	26%	70%
Nanded	Degloor	Degloor	Alur	GK + ND	1	Alur	587	Yes	No	1	29	15%	81%
Nanded	Degloor	Degloor	Amdapur	GK	1	Amdapur	169	Yes	No	1	17	14%	81%
Nanded	Degloor	Degloor	Bagantakli	GK	1	Bagantakli	266	Yes	No	1	35	35%	54%
Nanded	Degloor	Degloor	Bhayegaon	GK	1	Bhayegaon	235	Yes	No	2	45	18%	81%
Nanded	Degloor	Degloor	Borgaon	Control	1	Borgaon	160	Yes	No	1	19	14%	82%
Nanded	Degloor	Degloor	Davangir	GK + ND	1	Davangir	338	Yes	No	1	26	10%	89%
Nanded	Degloor	Degloor	Dhosni	Control	1	Dhosni	228	Yes	No	1	17	8%	32%
Nanded	Degloor	Mukhed	Gojegaon	GK	1	Gojegaon	712	Yes	Yes	1	35	14%	85%
Nanded	Degloor	Degloor	Hanegaon	Control	2	Hanegaon	1336	Yes	Yes	2	31	7%	35%
Nanded	Degloor	Degloor	Hottal	Control	1	Hottal	359	Yes	Yes	1	29	21%	38%
Nanded	Degloor	Mukhed	Jahur	GK	1	Jahur	738	Yes	Yes	2	41	9%	86%
Nanded	Degloor	Degloor	Karegaon	GK + ND	2	Karegaon	196	Yes	No	1	18	22%	26%
Nanded	Degloor	Degloor	Khanapur	GK	1	Khanapur	1204	Yes	Yes	2	37	15%	20%
Nanded	Degloor	Degloor	Kini	GK + ND	2	Kini	412	Yes	No	1	18	8%	28%
Nanded	Degloor	Degloor	Kushawadi	GK	1	Kushawadi	277	Yes	No	1	41	21%	20%
Nanded	Degloor	Degloor	Loni	GK + ND	3	Loni	778	Yes	Yes	3	85	7%	88%
Nanded	Degloor	Degloor	Malegaon (M)	Control	1	Malegaon (M)	557	No	N/A	2	41	N/A	N/A
Nanded	Degloor	Degloor	Markhel	Control	1	Markhel	1217	Yes	Yes	3	88	13%	32%
Nanded	Degloor	Degloor	Martoli	Control	1	Martoli	379	Yes	No	2	52	13%	84%
Nanded	Degloor	Mukhed	Mukramabad	GK + ND	1	Mukramabad	1652	Yes	Yes	2	59	15%	83%
Nanded	Degloor	Degloor	Narangal Bk	Control	1	Narangal Bk	708	Yes	Yes	2	40	22%	69%
Nanded	Degloor	Degloor	Shahapur	GK	2	Shahapur	1432	Yes	Yes	1	37	31%	58%
Nanded	Degloor	Mukhed	Undri Pd	GK + ND	1	Undri Pd	509	Yes	No	1	19	7%	89%
Nanded	Degloor	Degloor	Walag	GK + ND	1	Walag	529	Yes	No	3	74	21%	69%
Nanded	Degloor	Degloor	Wazarga	GK + ND	1	Wazarga	251	Yes	No	1	15	14%	82%
Nanded	Degloor	Degloor	Zari	Control	2	Zari	414	Yes	No	1	26	14%	76%
Nanded	Naigaon	Biloli	Arjapur	Control	1	Arjapur	432	Yes	No	1	27	22%	52%
Nanded	Naigaon	Biloli	Belkoni Bk	GK + ND	1	Belkoni Bk	567	Yes	Yes	1	23	19%	57%
Nanded	Naigaon	Naigaon	Betakbiloli	GK + ND	1	Betakbiloli	494	Yes	No	1	20	6%	85%
Nanded	Naigaon	Naigaon	Bhopala	GK + ND	1	Bhopala	423	Yes	No	1	20	10%	75%
Nanded	Naigaon	Biloli	Bijur	GK	1	Bijur	230	Yes	No	1	15	10%	64%
Nanded	Naigaon	Biloli	Chinchala	Control	1	Chinchala	377	No	N/A	1	38	N/A	N/A
Nanded	Naigaon	Naigaon	Degaon	GK	1	Degaon	826	Yes	Yes	1	32	18%	71%
Nanded	Naigaon	Naigaon	Gadga	Control	1	Gadga	664	Yes	Yes	1	19	19%	59%
Nanded	Naigaon	Naigaon	Ghungrala	GK	1	Ghungrala	524	Yes	No	1	34	19%	66%
Nanded	Naigaon	Naigaon	Kahala Bk	GK + ND	1	Kahala Bk	438	Yes	No	3	124	19%	68%
Nanded	Naigaon	Naigaon	Kandala	GK + ND	1	Kandala	377	Yes	No	4	116	6%	77%
Nanded	Naigaon	Biloli	Kasrali	Control	3	Kasrali	1121	Yes	Yes	5	127	24%	55%
Nanded	Naigaon	Naigaon	Khairgaon	Control	1	Khairgaon	390	Yes	No	3	93	16%	77%
Nanded	Naigaon	Biloli	Kinhala	GK + ND	1	Kinhala	438	Yes	No	1	30	16%	40%
Nanded	Naigaon	Biloli	Kolheborgaon	GK + ND	1	Kolheborgaon	573	Yes	No	3	67	12%	48%
Nanded	Naigaon	Naigaon	Krushnur	Control	1	Krushnur	916	Yes	Yes	2	45	18%	69%
Nanded	Naigaon	Naigaon	Kuntur	GK + ND	1	Kuntur	1342	Yes	Yes	3	82	16%	76%
Nanded	Naigaon	Biloli	Lohgaon	GK + ND	1	Lohgaon	1106	Yes	Yes	2	70	19%	60%
Nanded	Naigaon	Naigaon	Manjaram	GK	2	Manjaram	1232	Yes	Yes	4	102	12%	78%
Nanded	Naigaon	Loha	Martala	Control	1	Martala	547	No	N/A	2	60	N/A	N/A
Nanded	Naigaon	Naigaon	Narsi	GK	1	Narsi	2190	Yes	Yes	13	366	17%	78%
Nanded	Naigaon	Naigaon	Palasgaon	Control	2	Palasgaon	383	Yes	No	1	29	14%	70%
Nanded	Naigaon	Naigaon	Pimpalgaon	GK + ND	1	Pimpalgaon	357	Yes	No	3	54	12%	72%
Nanded	Naigaon	Biloli	Ramtirth	GK	1	Ramtirth	551	Yes	No	1	29	7%	29%

Baseline details of study GPs - continued

District	Branch	Block	GP	Treatment	# villages	Sampled village	# HHs	Listing	Segmented	# kendras	# clients	% toilet	% OD
Nanded	Naigaon	Mukhed	Salgara Kh	GK	1	Salgara Kh	404	No	N/A	2	50	N/A	N/A
Nanded	Naigaon	Naigaon	Sawarkhed	Control	1	Sawarkhed	435	Yes	No	2	35	10%	76%
Nanded	Naigaon	Naigaon	Somthana	GK	1	Somthana	476	Yes	No	1	29	19%	67%
Nanded	Naigaon	Naigaon	Talbid	GK	1	Talbid	159	Yes	No	1	13	29%	59%
Nanded	Naigaon	Biloli	Talni	GK	1	Talni	399	Yes	No	1	29	25%	22%
Nanded	Naigaon	Mukhed	Umardari	Control	1	Umardari	516	Yes	No	2	57	20%	72%

Table 118: Sample size and composition (by GP)

District	Branch	Block	GP	Treatment	Actual sample size			Sample shortfall/surplus		
					Total	GK clients	non-GK clients	Total	GK clients	non-GK clients
Latur	Ahmadpur	Ahmadpur	Andhori	GK	36	20	16	6	5	1
Latur	Ahmadpur	Ahmadpur	Belur	GK + ND	30	15	15	0	0	0
Latur	Ahmadpur	Ahmadpur	Chilkha	Control	32	16	16	2	1	1
Latur	Ahmadpur	Ahmadpur	Dhalegaon	GK	30	15	15	0	0	0
Latur	Ahmadpur	Ahmadpur	Hadolti	GK + ND	30	15	15	0	0	0
Latur	Ahmadpur	Ahmadpur	Hipparga Kajal	Control	32	16	16	2	1	1
Latur	Ahmadpur	Ahmadpur	Kingaon	Control	29	15	14	-1	0	-1
Latur	Ahmadpur	Ahmadpur	Lanji	GK	29	14	15	-1	-1	0
Latur	Ahmadpur	Ahmadpur	Sangavi (Su)	GK + ND	30	15	15	0	0	0
Latur	Ahmadpur	Ahmadpur	Tembhurni	GK	30	15	15	0	0	0
Latur	Ahmadpur	Ahmadpur	Thodga	Control	30	15	15	0	0	0
Latur	Ahmadpur	Ahmadpur	Vilegaon	Control	32	15	17	2	0	2
Latur	Nilanga	Nilanga	Ambulga Vi	Control	30	15	15	0	0	0
Latur	Nilanga	Nilanga	Ansarwada	GK + ND	31	16	15	1	1	0
Latur	Nilanga	Nilanga	Dapka	GK + ND	32	17	15	2	2	0
Latur	Nilanga	Nilanga	Dhanora	Control	34	19	15	4	4	0
Latur	Nilanga	Nilanga	Ekurga Road	Control	30	15	15	0	0	0
Latur	Nilanga	Nilanga	Hadga	Control	30	15	15	0	0	0
Latur	Nilanga	Nilanga	Halgara	GK	30	15	15	0	0	0
Latur	Nilanga	Nilanga	Handral	GK + ND	29	14	15	-1	-1	0
Latur	Nilanga	Nilanga	Hanmantwadi Ab	GK	24	9	15	-6	-6	0
Latur	Nilanga	Nilanga	Harijawalga	GK	30	15	15	0	0	0
Latur	Nilanga	Nilanga	Jajnoor	GK + ND	30	15	15	0	0	0
Latur	Nilanga	Nilanga	Jau	GK	30	15	15	0	0	0
Latur	Nilanga	Nilanga	Jewari	GK + ND	30	15	15	0	0	0
Latur	Nilanga	Nilanga	Kelgaon	GK + ND	28	15	13	-2	0	-2
Latur	Nilanga	Ausa	Kharosa	Control	30	15	15	0	0	0
Latur	Nilanga	Nilanga	Kokalgaon	GK	30	15	15	0	0	0
Latur	Nilanga	Nilanga	Lambota	Control	30	15	15	0	0	0
Latur	Nilanga	Nilanga	Madansuri	GK	30	15	15	0	0	0
Latur	Nilanga	Nilanga	Makni	Control	30	15	15	0	0	0
Latur	Nilanga	Nilanga	Mannat hpur	GK + ND	30	15	15	0	0	0
Latur	Nilanga	Nilanga	Nanand	Control	30	15	15	0	0	0
Latur	Nilanga	Nilanga	Nitur	Control	28	15	13	-2	0	-2
Latur	Nilanga	Nilanga	Sarwadi	GK	30	15	15	0	0	0
Latur	Nilanga	Nilanga	Sindhkhed	GK	28	13	15	-2	-2	0
Latur	Nilanga	Nilanga	Umarga H	GK + ND	30	15	15	0	0	0
Latur	Nilanga	Deoni	Walandi	GK	30	15	15	0	0	0
Latur	Nilanga	Nilanga	Yelnoor	Control	30	15	15	0	0	0
Latur	Udgir	Udgir	Banshelki	GK	30	15	15	0	0	0

Sample size and composition (by GP) - continued

District	Branch	Block	GP	Treatment	Actual sample size			Sample shortfall/surplus		
					Total	GK clients	non-GK clients	Total	GK clients	non-GK clients
Latur	Udgir	Udgir	Bhakaskheda	Control	30	15	15	0	0	0
Latur	Udgir	Udgir	Chandegaon	GK	30	15	15	0	0	0
Latur	Udgir	Udgir	Dawangaoon	GK + ND	30	15	15	0	0	0
Latur	Udgir	Udgir	Dewarjan	GK + ND	30	15	15	0	0	0
Latur	Udgir	Deoni	Hanchanal	GK	30	15	15	0	0	0
Latur	Udgir	Udgir	Kallur	GK + ND	30	15	15	0	0	0
Latur	Udgir	Deoni	Kamalwadi	GK	30	15	15	0	0	0
Latur	Udgir	Udgir	Karkheli	GK + ND	30	15	15	0	0	0
Latur	Udgir	Udgir	Kiniyalladevi	GK + ND	30	15	15	0	0	0
Latur	Udgir	Udgir	Kumtha Kh	Control	30	15	15	0	0	0
Latur	Udgir	Udgir	Lohara	Control	30	15	15	0	0	0
Latur	Udgir	Udgir	Loni	GK	30	15	15	0	0	0
Latur	Udgir	Udgir	Malkapur	GK	30	15	15	0	0	0
Latur	Udgir	Udgir	Nagalgaon	GK	30	15	15	0	0	0
Latur	Udgir	Udgir	Nalgir	GK + ND	29	15	14	-1	0	-1
Latur	Udgir	Udgir	Netragaon	GK + ND	30	15	15	0	0	0
Latur	Udgir	Udgir	Nideban	GK + ND	30	15	15	0	0	0
Latur	Udgir	Udgir	Pimpari	Control	30	15	15	0	0	0
Latur	Udgir	Udgir	Shekapur	Control	30	15	15	0	0	0
Latur	Udgir	Udgir	Shelhal	GK	30	15	15	0	0	0
Latur	Udgir	Udgir	Shirol	Control	30	15	15	0	0	0
Latur	Udgir	Udgir	Somnathpur	Control	30	15	15	0	0	0
Latur	Udgir	Udgir	Wadhona (Bk)	GK	30	15	15	0	0	0
Nanded	Degloor	Deglur	Achegaon	GK + ND	30	15	15	0	0	0
Nanded	Degloor	Deglur	Alur	GK + ND	30	15	15	0	0	0
Nanded	Degloor	Deglur	Amdapur	GK	30	15	15	0	0	0
Nanded	Degloor	Deglur	Bagantakli	GK	29	15	14	-1	0	-1
Nanded	Degloor	Deglur	Bhayegaon	GK	30	15	15	0	0	0
Nanded	Degloor	Deglur	Borgaon	Control	30	15	15	0	0	0
Nanded	Degloor	Deglur	Davangir	GK + ND	30	15	15	0	0	0
Nanded	Degloor	Deglur	Dhosni	Control	30	15	15	0	0	0
Nanded	Degloor	Mukhed	Gojegaon	GK	30	15	15	0	0	0
Nanded	Degloor	Deglur	Hanegaon	Control	30	15	15	0	0	0
Nanded	Degloor	Deglur	Hottal	Control	30	15	15	0	0	0
Nanded	Degloor	Mukhed	Jahur	GK	30	15	15	0	0	0
Nanded	Degloor	Deglur	Karegaon	GK + ND	30	15	15	0	0	0
Nanded	Degloor	Deglur	Khanapur	GK	30	15	15	0	0	0
Nanded	Degloor	Deglur	Kini	GK + ND	30	15	15	0	0	0
Nanded	Degloor	Deglur	Kushawadi	GK	30	15	15	0	0	0
Nanded	Degloor	Deglur	Loni	GK + ND	30	15	15	0	0	0

Sample size and composition (by GP) - continued

District	Branch	Block	GP	Treatment	Actual sample size			Sample shortfall/surplus		
					Total	GK clients	non-GK clients	Total	GK clients	non-GK clients
Nanded	Degloor	Deglur	Malegaon (M)	Control	30	15	15	0	0	0
Nanded	Degloor	Deglur	Markhel	Control	31	15	16	1	0	1
Nanded	Degloor	Deglur	Martoli	Control	30	15	15	0	0	0
Nanded	Degloor	Mukhed	Mukramabad	GK + ND	30	15	15	0	0	0
Nanded	Degloor	Deglur	Narangal Bk	Control	30	15	15	0	0	0
Nanded	Degloor	Deglur	Shahapur	GK	30	15	15	0	0	0
Nanded	Degloor	Mukhed	Undri Pd	GK + ND	30	15	15	0	0	0
Nanded	Degloor	Deglur	Walag	GK + ND	30	15	15	0	0	0
Nanded	Degloor	Deglur	Wazarga	GK + ND	30	15	15	0	0	0
Nanded	Degloor	Deglur	Zari	Control	30	15	15	0	0	0
Nanded	Naigaon	Biloli	Arjapur	Control	30	15	15	0	0	0
Nanded	Naigaon	Biloli	Belkoni Bk	GK + ND	30	15	15	0	0	0
Nanded	Naigaon	Naigaon	Betakbiloli	GK + ND	30	15	15	0	0	0
Nanded	Naigaon	Naigaon	Bhopala	GK + ND	25	10	15	-5	-5	0
Nanded	Naigaon	Biloli	Bijur	GK	26	11	15	-4	-4	0
Nanded	Naigaon	Biloli	Chinchala	Control	15	15	0	0	0	0
Nanded	Naigaon	Naigaon	Degaon	GK	30	15	15	0	0	0
Nanded	Naigaon	Naigaon	Gadga	Control	30	15	15	0	0	0
Nanded	Naigaon	Naigaon	Ghungrala	GK	31	15	16	1	0	1
Nanded	Naigaon	Naigaon	Kahala Bk	GK + ND	37	20	17	7	5	2
Nanded	Naigaon	Naigaon	Kahala Kh	Control	15	0	15	0	0	0
Nanded	Naigaon	Naigaon	Kandala	GK + ND	30	15	15	0	0	0
Nanded	Naigaon	Biloli	Kasrali	Control	30	15	15	0	0	0
Nanded	Naigaon	Naigaon	Khairgaon	Control	30	15	15	0	0	0
Nanded	Naigaon	Biloli	Kinhala	GK + ND	30	15	15	0	0	0
Nanded	Naigaon	Biloli	Kolheborgaon	GK + ND	30	15	15	0	0	0
Nanded	Naigaon	Naigaon	Krushnur	Control	30	15	15	0	0	0
Nanded	Naigaon	Naigaon	Kuntur	GK + ND	30	15	15	0	0	0
Nanded	Naigaon	Biloli	Lohgaon	GK + ND	29	14	15	-1	-1	0
Nanded	Naigaon	Naigaon	Manjaram	GK	28	14	14	-2	-1	-1
Nanded	Naigaon	loha	Martala	Control	30	15	15	0	0	0
Nanded	Naigaon	Naigaon	Narsi	GK	30	15	15	0	0	0
Nanded	Naigaon	Naigaon	Palasgaon	Control	30	15	15	0	0	0
Nanded	Naigaon	Naigaon	Pimpalgaon	GK + ND	30	15	15	0	0	0
Nanded	Naigaon	Biloli	Ramtirth	GK	30	15	15	0	0	0
Nanded	Naigaon	Mukhed	Salgara Kh	GK	30	15	15	0	0	0
Nanded	Naigaon	Naigaon	Sawarkhed	Control	30	15	15	0	0	0
Nanded	Naigaon	Naigaon	Somthana	GK	30	15	15	0	0	0
Nanded	Naigaon	Naigaon	Talbid	GK	27	12	15	-3	-3	0
Nanded	Naigaon	Biloli	Talni	GK	29	14	15	-1	-1	0

Sample size and composition (by GP) - continued

District	Branch	Block	GP	Treatment	Actual sample size			Sample shortfall/surplus		
					Total	GK clients	non-GK clients	Total	GK clients	non-GK clients
Nanded	Naigaon	Mukhed	Umardari	Control	30	15	15	0	0	0
TOTALS					3595	1794	1801	-5	-6	1