A Targeted Maximum Likelihood Estimation

A.1 Methods

We pursued a targeted maximum likelihood estimation (MLE) analysis for two reasons. First, we wanted to control for any residual confounding of the treatment effect. Although the matching in the design stage balanced a large number of covariates between the intervention and control groups, there were still some characteristics (e.g., cell phone and latrine ownership) that remained imbalanced at follow-up and could confound the treatment effect. Second, recent theoretical and simulation results suggest that targeted MLE has potential to increase the efficiency of the estimates when the likelihood is targeted for the parameter of interest. [1, 2]

The targeted MLE estimator is doubly robust, and is asymptotically equivalent to previously described double robust marginal structural model estimators based on estimating functions [1]. The estimator is considered double robust because it is consistent if the analyst correctly specifies the model for the outcome *or* the model for the treatment mechanism. One advantage of targeted MLE over other double robust estimators is that it is easier to implement using standard software. All analysis was conducted using R software (www.cran.r-project.org).

Let Y be an outcome of interest, A be the intervention status equal to 1 if a child lives in an intervention village and 0 otherwise, and W be a set of covariates that are potential confounders of the relationship between A and Y. We calculated adjusted estimates using the following steps (following notation from Bembom et al. [3]):

1. Estimate the conditional expectation of Y given A and W using a generalized linear model with maximum likelihood. We denote this initial estimate $Q_n^0(A, W)$.

2. Estimate the conditional probability of receiving the intervention, A, given W using a logit model. We denote this estimate $g_n^0(A|W)$.

3. For each child, calculate a covariate based on her observed values for A and W and using the estimate $g_n^0(A|W)$. We denote this covariate $h(A, W)$, where:

$$h(A, W) = \frac{I(A = 1)}{g_n^0(1|W)} - \frac{I(A = 0)}{g_n^0(0|W)} \quad (1)$$

4. Update the original regression by adding the covariate $h(A, W)$ and estimate the corresponding coefficient by maximum likelihood, holding the remaining coefficient estimates at their initial values. In practice, this is achieved by estimating a univariate regression of Y on $h(A, W)$ with $Q_n^0(A, W)$ as an offset with coefficient constrained to one. Let ϵ_n be the coefficient on $h(A, W)$. We denote this one-step updated regression $Q_n^1(A, W)$ where:

$$Q_n^1(A, W) = Q_n^0(A, W) + \epsilon_n h(A, W) \quad (2)$$

5. Evaluate the updated regression at $A = 1$ and $A = 0$ to get two predicted outcomes for each child. Take the empirical mean of the difference across the population to obtain a targeted estimate of the difference:

$$\theta_n^{T-MLE} = \frac{1}{n}\sum_{i=1}^{n} Q_n^1(1, W) - Q_n^1(0, W) \tag{3}$$

In our initial estimate of $Q_n^0(A, W)$ there are a potentially large number of covariates in W and the models' functional form is unknown. We initially considered village-level indicators and characteristics that were unlikely to have been affected by the intervention (i.e., they were pre-treatment). The covariates that we considered include (headings in **bold**):

Covariate	Abbreviation
Child's characteristics	
Sex	sex
Age (months)	age
Total months breast fed	bftot
Mother's characteristics	
Age (years)	mage
Works for money (yes/no)	mwork
Literate (yes/no)	mlit
Leaves village ≥ once per week (yes/no)	trips
Household characteristics	
Total persons living in home	totp
Num. children < 15 years	num15
Num. children < 5 years	num5
Electricity (yes/no)	elec
Dirt floor (yes/no)	dirt
Thatched roof (yes/no)	palm
Home ownership (yes/no)	homeown
Land ownership (yes/no)	landown
Use banking services (yes/no)	bank
Have relatives in USA (yes/no)	relus
Have relatives in the Capital (yes/no)	relguat
Travel time by car to the municipal capital (min)	ttime
Durable good ownership (yes/no)	
Refrigerator	refri
Radio	radio
Television	tv
Mobile phone	cell
Bicycle	bike
Automobile	car
Water supply	
Primary water source (factor)	watsource
Private tap	
Public tap	
Public well	
Spring	
Surface water (river/lake)	
Minutes per day retrieving water	wattime
Satisified with water quantity (yes/no)	watsat
Sanitation	
Latrine ownership (yes/no)	latrine
Animals in living vicinity of house (yes/no)	
Pigs	pigs
Chickens/ducks	birds
Dogs/cats	dogscats
Cows/horses/mules/donkeys	stock

We eliminated the number of children < 5 and < 15 due to collinearity with the total number of persons in each household. We also eliminated minutes per day retrieving water and satisfaction with water quantity due to collinearity with water source. We restricted the covariate set to those that were considered to be potential confounders by the authors [4] and had a strong association with the outcome based on a previously published backward deletion approach [5,6]. The backward deletion approach selects variables that, when removed from a multivariable specification including all candidate covariates, change the treatment coefficient by 5% or more.

After this dimension reduction step, we chose final model specifications using the Deletion/Substitution/Addition (D/S/A) selection algorithm allowing for two-way interactions, quadratic terms, and 15 total terms [7]. We ran the backward deletion and D/S/A model selection separately for each health outcome.

We also selected terms for the treatment mechanism, $g_n^0(A|W)$ by initially including covariates that had an absolute standardized mean difference greater than 20 (intervention minus control) and could not reasonably be influenced by the intervention. We chose a threshold standardized difference of 20 because it roughly corresponded to univariate t-statistic-based p-values of 0.01, and because it reduced the covariate set sufficiently for model selection. We selected our final treatment model using the D/S/A algorithm allowing for two-way interactions, quadratic terms, and 15 total terms. We used the same treatment model for all outcomes.

Finally, we calculated percentile-based 95% confidence intervals for the estimates using a bootstrap with matched village pairs as the sampling unit (to reflect the design) and 1,000 iterations [1].

Below, we present unadjusted and adjusted estimates for each child health outcome that we considered. For binary outcomes, the estimates are the marginal, population-averaged difference in the longitudinal prevalence for Intervention minus Control. For continuous Z-score outcomes, the estimates are the marginal, population-averaged difference in Z-score for Intervention minus Control. In each case we report the final covariates considered in $Q_n^0(A, W)$ and $g_n^0(A|W)$. We also summarize the distribution of $g_n^0(1|W)$ because the distribution of the predicted probability of treatment helps identify whether there exists common support on the covariates W between the treatment groups. Specifically, the parameter of interest is only well-defined if:

$$0 < P(A = 1|W) < 1 \tag{4}$$

which states that variation in treatment exists for each stratum of W [1].

A.2 Results

The model selection process successfully identified covariates to use in adjusted specifications, summarized in Table A.1 . In no case did the D/S/A algorithm select interactions of covariates in W with the treatment.

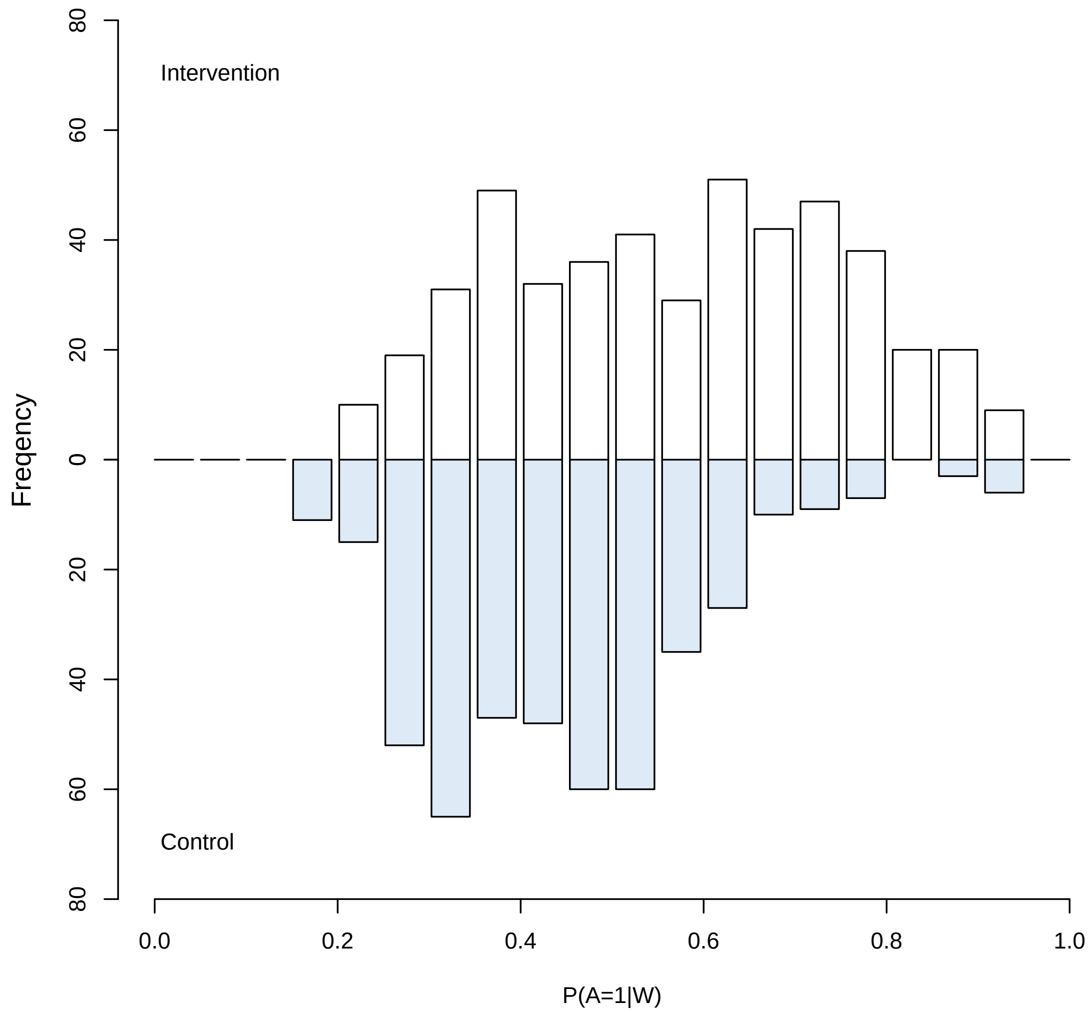
Table A.1 : **Summary of covariates included in targeted maximum likelihood estimation models for child health outcomes.**

Model, outcome	Covariates included in W
$g_n^0(A\|W)$	ttime, ttime^2, bike, cell, elec, latrine, watsource, pigs, stock
$Q_n^0(A,W)$, Diarrhea	age, bftot, dirt, watsource, age*watsource(public well), bftot*dirt, bftot*age
$Q_n^0(A,W)$, HCGI *	age, age^2, watsource, latrine, refri
$Q_n^0(A,W)$, Cough or diff. breathing	age
$Q_n^0(A,W)$, Congestion or coryza	age, age^2, dirt
$Q_n^0(A,W)$, ALRI †	mlit
$Q_n^0(A,W)$, Weight	age, age^2, tv, elec, palm, relus, watsource
$Q_n^0(A,W)$, Height	age, age^2, dirt, relguat
$Q_n^0(A,W)$, Weight-for-height	age, tv
$Q_n^0(A,W)$, Mid-upper arm circ.	age, age^2, tv, dirt, watsource

*Highly Credible Gastrointestinal Illness. The definition is included in the main text.
† Clinical Acute Lower Respiratory Infections. The definition is included in the main text.

The predicted probabilities from $g_n^0(A|W)$ indicate that there is common support for the covariates selected in W. The probabilities $g_n^0(1|W) = P(A = 1|W)$ are bounded away from 0 and 1, and range from 0.19 to 0.94 (median = 0.49, interquartile range = 0.44 – 0.69). Figure A.1 plots a histogram of the predicted probabilities of receiving the intervention for control and intervention children, and demonstrates good overlap in the distributions. This result helps confirm the usefulness of carefully selecting control villages in the design stage, and indicates that our parameters of interest are well-defined in this dataset.

Figure A.1 : **Histogram of predicted probabilities of receiving treatment, $g_n^0(1|W) = P(A = 1|W)$, for children in intervention and control villages using the specification in Table A.1 .**



In the tables below, we present the unadjusted results from the main manuscript along side the adjusted results obtained with targeted MLE. In all cases, the adjusted estimates are consistent with the unadjusted estimates and indicate that there is little or no residual confounding by the key demographic, socioeconomic and environmental covariates considered in our analysis. In almost all cases, the adjusted estimates are more efficient (have smaller standard errors), an important empirical finding that is consistent with earlier theoretical and simulation-based results [1].

Table A.2 : **Unadjusted and adjusted difference in longitudinal prevalence of illness in 929 children under age 5 following a three-year point-of-use water treatment and handwashing intervention, Camotán, Guatemala, 2007. Adjusted values were estimated using targeted maximum likelihood.**

Outcome	Control	Intervention	Unadjusted			Adjusted		
	Days Ill / Observed	Days Ill / Observed	LPD	SE *	(95% CI) *	LPD	SE *	(95% CI) *
Diarrhea	107/910	115/948	0.004	0.0288	(-0.051, 0.058)	0.007	0.0254	(-0.037, 0.059)
HCGI †	113/910	122/948	0.005	0.0308	(-0.054, 0.065)	0.010	0.0282	(-0.042, 0.068)
Cough or diff. breathing	268/910	291/948	0.012	0.0597	(-0.097, 0.137)	0.003	0.0592	(-0.111, 0.117)
Congestion or coryza	144/910	173/948	0.024	0.0249	(-0.026, 0.071)	0.023	0.0249	(-0.022, 0.075)
ALRI ‡	54/910	74/948	0.019	0.0278	(-0.028, 0.078)	0.018	0.0285	(-0.031, 0.077)

∗ Standard Errors and 95% confidence intervals calculated by bootstrap resampling matched village pairs with 1000 iterations.
† Highly Credible Gastrointestinal Illness. The definition is included in the main text.
‡ Clinical Acute Lower Respiratory Infections. The definition is included in the main text.

Table A.3 : **Unadjusted and adjusted difference in anthropometric Z-scores in children under age 5 following a three-year point-of-use water treatment and handwashing intervention, Camotán, Guatemala, 2007. Adjusted values were estimated using targeted maximum likelihood.**

Z-Score *	Control			Intervention			Unadjusted			Adjusted		
	N	Mean	SD	N	Mean	SD	Diff.	SE	(95% CI) †	Diff.	SE	(95% CI) †
Weight	423	-1.312	1.325	453	-1.365	1.219	-0.053	0.1368	(-0.331, 0.206)	-0.111	0.0768	(-0.254, 0.050)
Height	424	-2.177	1.880	453	-2.136	1.596	0.041	0.1605	(-0.305, 0.326)	-0.055	0.1338	(-0.332, 0.177)
Weight-for-height	421	-0.122	1.728	451	-0.187	1.421	-0.066	0.0967	(-0.248, 0.124)	-0.019	0.0837	(-0.174, 0.145)
Mid-upper-arm circ.	401	0.348	0.884	426	0.335	0.825	-0.014	0.0806	(-0.166, 0.145)	-0.057	0.0657	(-0.183, 0.079)

∗ Z-scores were calculated using a standard WHO Stata algorithm and 2006 world reference data.

† Standard Errors and 95% confidence intervals calculated by bootstrap resampling matched village pairs with 1000 iterations.

B Treatment Effects among Intervention Participants

B.1 Methods

In this section, we summarize the treatment effect estimates after defining the intervention population as the 147 households (49% of the intervention group) who reported participating in the intervention. This is analogous to a "treatment actually received" analysis in a randomized trial, and so there may be important self-selection into the treatment group that can lead to confounding bias. However, we present these exploratory results following the recommendations of Victora et al [8].

The self-reported measure of participation could be biased either up or down. Participants may have over-reported participation to please our field staff because they viewed it as a socially acceptable answer, or participants may under-report participation because they think that if they admit to having received assistance in the past it will hurt their chances for receiving additional assistance in future campaigns/relief programs. Without a detailed household-level record of program participation (none exists for the intervention we have evaluated) there is no way to reliably validate this measurement.

Our method of confounding adjustment is identical to the targeted MLE approach described in Appendix A, although we repeated all model selection routines for $g_n^0(A|W)$ and $Q_n^0(A, W)$ using the alternate treatment definition (self-reported participation). Since this alternate treatment definition varies at the household-level, our inference relies on a bootstrap that resamples children at the household level. This approach assumes that households are independent, which is a stronger assumption than in our primary analysis where we assume that children in separate villages are independent.

B.2 Results

Table B.1 : Water storage and treatment practices following a three-year point-of-use water treatment and handwashing intervention, Camotán, Guatemala, 2007. N=147 intervention and N=453 control households. Unlike the primary analysis, intervention treatment is assigned based on self-reported participation in the intervention.

Outcome	Control		Intervention		Risk Difference	
	%	(N)	%	(N)	(95% CI) *	
Water storage practices						
Stores drinking water in home	80.4	(364)	81.0	(119)	0.006	(-0.07, 0.08)
Excl. covered or narrow mouth	61.6	(279)	61.9	(91)	0.003	(-0.09, 0.10)
Exclusively covered	59.8	(271)	61.2	(90)	0.014	(-0.08, 0.09)
Self-reported water treatment						
Any method	23.8	(108)	37.4	(55)	0.136	(0.05, 0.23)
Boiling	17.9	(81)	19.7	(29)	0.018	(-0.06, 0.10)
SODIS †	1.1	(5)	23.1	(34)	0.220	(0.15, 0.29)
Chlorine	5.1	(23)	2.7	(4)	-0.024	(-0.06, 0.01)
Confirmed water treatment ‡						
Any method	4.9	(22)	9.5	(14)	0.047	(0.00, 0.10)
Boiling	3.8	(17)	4.1	(6)	0.003	(-0.03, 0.04)
SODIS	0.4	(2)	6.1	(9)	0.057	(0.02, 0.10)
Chlorine	1.1	(5)	0.0	(0)	-0.011	(-0.02, 0.00)

* 95%

Confidence Intervals calculated by bootstrap resampling households with 1000 iterations.

† SODIS: Solar Disinfection.

‡ Water treatment was confirmed if the family (i) self-reported treating water in the previous 7 days, (ii) had treated water at the time of the interview, and (iii) could produce the materials they used to treat water.

Table B.2 : Handwashing and hygiene conditions following a three-year point-of-use water treatment and handwashing intervention, Camotán, Guatemala, 2007. N=147 intervention and N=453 control households. Unlike the primary analysis, intervention treatment is assigned based on self-reported participation in the intervention.

Outcome	Control		Intervention		Risk Difference	
	%	(N)	%	(N)	(95% CI) *	
Self-reported handwashing †						
Before cooking	77.7	(352)	83.7	(123)	0.060	(-0.02, 0.13)
Before eating	35.1	(159)	28.6	(42)	-0.065	(-0.15, 0.02)
Before feeding children	18.8	(85)	17.0	(25)	-0.018	(-0.09, 0.06)
After defecation	52.3	(237)	49.0	(72)	-0.033	(-0.12, 0.06)
After changing baby	12.4	(56)	13.6	(20)	0.012	(-0.05, 0.08)
Spot check observations						
Mother's hands are clean	90.1	(408)	88.4	(130)	-0.016	(-0.08, 0.04)
Mother's nails are clean	72.2	(327)	74.1	(109)	0.020	(-0.06, 0.10)
Can produce a bar of soap	88.3	(400)	95.9	(141)	0.076	(0.04, 0.12)
Bar soap is in plain view	56.1	(254)	63.3	(93)	0.072	(-0.02, 0.16)
Food is covered	53.6	(243)	57.1	(84)	0.035	(-0.05, 0.12)
Garbage present inside home	54.1	(245)	47.6	(70)	-0.065	(-0.16, 0.03)
Feces observed in living area	73.5	(333)	74.1	(109)	0.006	(-0.08, 0.08)

* 95% Confidence Intervals calculated by bootstrap resampling households with 1000 iterations.

† Responses to an open-ended question about handwashing in the 24 hours before the interview.

Table B.3 : **Summary of covariates included in targeted maximum likelihood estimation models for child health outcomes. In this analysis treatment, A, was based on self-reported participation in the intervention.**

Model, outcome	Covariates included in W
$g_n^0(A\|W)$	ttime, dirt, watsource, pigs
$Q_n^0(A,W)$, Diarrhea	bftot, bftot^2, latrine, car, watsource
$Q_n^0(A,W)$, HCGI *	bftot, bftot^2, latrine, refri, watsource
$Q_n^0(A,W)$, Cough or diff. breathing	age
$Q_n^0(A,W)$, Congestion or coryza	dirt, pigs
$Q_n^0(A,W)$, ALRI †	mlit
$Q_n^0(A,W)$, Weight	age, age^2, totp, totp^2, trips, watsource
$Q_n^0(A,W)$, Height	bftot, bftot^2, trips, cell, latrine, relus, watsource
$Q_n^0(A,W)$, Weight-for-height	age, tv, mage
$Q_n^0(A,W)$, Mid-upper arm circ.	age, age^2, totp, totp^2, dirt, palm, mlit, tv, elec, relus, relguat, watsource

*Highly Credible Gastrointestinal Illness. The definition is included in the main text.

† Clinical Acute Lower Respiratory Infections. The definition is included in the main text.

Figure B.1 : **Histogram of predicted probabilities of receiving treatment, $g_n^0(1|W) = P(A = 1|W)$, for children in intervention and control villages using the specification in Table B.3 . The predicted probabilities range from 0.21 to 0.88 (median = 0.52).**

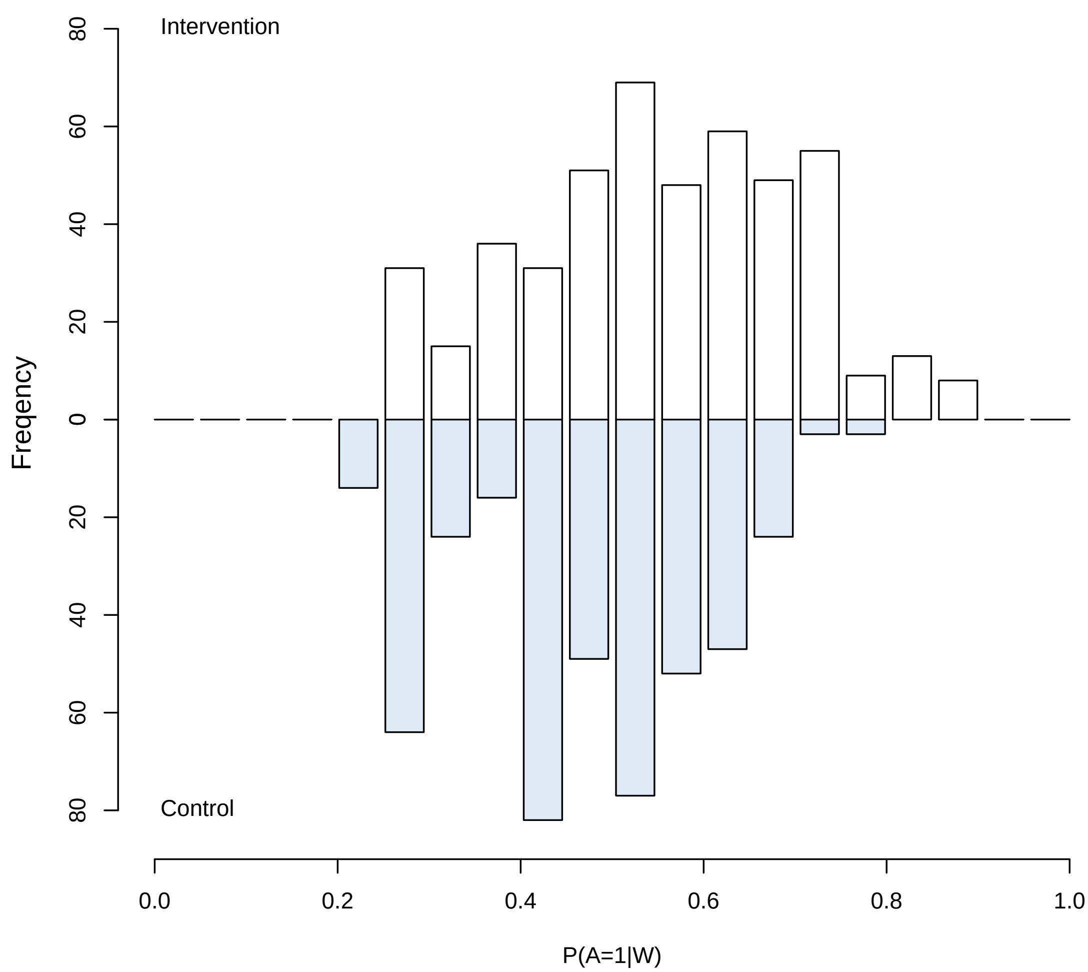


Table B.4 : **Unadjusted and adjusted difference in longitudinal prevalence of illness in 929 children under age 5 following a three-year point-of-use water treatment and handwashing intervention, Camotán, Guatemala, 2007. Adjusted values were estimated using targeted maximum likelihood.Unlike the primary analysis, intervention treatment is assigned based on self-reported participation in the intervention.**

Outcome	Control	Intervention	Unadjusted			Adjusted		
	Days Ill / Observed	Days Ill / Observed	LPD	SE *	(95% CI) *	LPD	SE *	(95% CI) *
Diarrhea	174/1404	48/454	-0.018	0.0242	(-0.064, 0.028)	0.012	0.0246	(-0.032, 0.060)
HCGI †	186/1404	49/454	-0.025	0.0238	(-0.069, 0.022)	0.013	0.0281	(-0.041, 0.068)
Cough or diff. breathing	405/1404	154/454	0.051	0.0395	(-0.029, 0.130)	0.003	0.0604	(-0.112, 0.119)
Congestion or coryza	220/1404	97/454	0.057	0.0346	(-0.012, 0.130)	0.034	0.0258	(-0.015, 0.087)
ALRI ‡	83/1404	45/454	0.040	0.0247	(-0.005, 0.090)	0.018	0.0285	(-0.031, 0.076)

∗ 95% Confidence Intervals calculated by bootstrap resampling households with 1000 iterations.
† Highly Credible Gastrointestinal Illness. The definition is included in the main text.
‡ Clinical Acute Lower Respiratory Infections. The definition is included in the main text.

Table B.5 : **Unadjusted and adjusted difference in anthropometric Z-scores in children under age 5 following a three-year point-of-use water treatment and handwashing intervention, Camotán, Guatemala, 2007. Adjusted values were estimated using targeted maximum likelihood.Unlike the primary analysis, intervention treatment is assigned based on self-reported participation in the intervention.**

Z-Score *	Control			Intervention			Unadjusted			Adjusted		
	N	Mean	SD	N	Mean	SD	Diff.	SE	(95% CI) †	Diff.	SE	(95% CI) †
Weight	659	-1.345	1.291	217	-1.322	1.208	0.024	0.1070	(-0.195, 0.236)	-0.071	0.0968	(-0.264, 0.116)
Height	659	-2.192	1.755	218	-2.048	1.686	0.144	0.1472	(-0.146, 0.434)	-0.094	0.1156	(-0.322, 0.136)
Weight-for-height	656	-0.149	1.646	216	-0.175	1.347	-0.026	0.1244	(-0.276, 0.204)	-0.037	0.0823	(-0.197, 0.111)
Mid-upper-arm circ.	623	0.341	0.853	204	0.342	0.856	0.000	0.0701	(-0.136, 0.143)	-0.061	0.0671	(-0.178, 0.084)

∗ Z-scores were calculated using a standard WHO Stata algorithm and 2006 world reference data.
† 95% Confidence Intervals calculated by bootstrap resampling households with 1000 iterations.

C Additional Information for Planning Studies: Intervention, Setting, Incidence and Intra Cluster Correlations

C.1 More details about the intervention

Between October 2003 and September 2006, two non-governmental organizations (NGOs), Caritas and Catholic Relief Services, implemented a large household water treatment and handwashing campaign in approximately 90 villages across three municipalities in rural eastern Guatemala. We conducted our evaluation in the municipality of Camotán because both the NGO records were more complete for that municipality compared to others (Jocotán and San Juan Ermita). The implementing organizations had oversight from the SODIS Foundation (www.fundacionsodis.org), who provided input into the training materials, social marketing messages and general implementation strategy. The NGOs promoted three water treatment methods: boiling, solar disinfection (SODIS), and chlorination using dilute bleach. Based on our exchanges with some of the Caritas technicians, the campaign likely emphasized the SODIS method over chlorination and boiling, but they encouraged families to use their own preferred method (or combination of methods). Handwashing education and social marketing included demonstrations of correct technique that emphasized using soap or detergent and scrubbing thoroughly. The promotion also emphasized critical times to wash hands that included: before cooking, before eating, before feeding children, after defecation and after changing babies.

All villages received the same intervention package and all activities were initiated at the same time (October 2003). The intervention program used a "train the trainer" model, where NGO technicians trained local community women to promote the behavior change through social marketing and household visits. The NGOs recruited approximately one community promoter per 25 participating households. The trained health promoters later visited households with children under age three or with pregnant mothers to promote water treatment and handwashing with soap. The visits occurred monthly or bi-monthly and lasted approximately 30 minutes each.

Promoters educated mothers about proper nutrition for their children, and at the end of each visit gave the family a small ration of rice, beans and oil. This nutritional component to the intervention was implemented at a regional scale in concert with many additional NGOs, UNICEF and Guatemala's National Plan for the Reduction of Chronic Malnutrition (a response to a drought and subsequent famine in 2001 that struck Camotán and adjacent Jocotán). This component was not unique to intervention villages in our sample (indeed, we confirmed that all villages in our sample – intervention and control – received food aid during the study period).

C.2 Setting

This study was conducted in the Camotán municipality in the mountainous state of Chiquimula, Guatemala near the eastern border with Honduras. Camotán is a dry, mountainous region with 94 rural villages located between 2 and 37 km from the municipal center, and typically accessed by dirt roads. The primary occupation is agriculture; corn and beans are the main crops with some coffee grown at higher elevations. There is typically only one crop per year during the wet season. More than 89% of people live in moderate or extreme poverty [9]. Water is obtained from mountain springs and surface water. Community and household taps, where available, are connected to gravity-fed spring networks, and water sources are typically contaminated with fecal organisms [10].

C.3 Cumulative Incidence

Table C.1 : **Incident illness over a 7 day period among children under age 5 following a three-year point-of-use water treatment and handwashing intervention, Camotán, Guatemala, 2007**

Outcome	Control		Intervention		Risk Difference	
	Children at Risk	New Episodes	Children at Risk	New Episodes	(95% CI) *	
Diarrhea	436	51	453	54	0.002	(-0.046, 0.053)
HCGI †	435	74	452	84	0.016	(-0.066, 0.094)
Cough or diff. breathing	387	74	395	74	-0.004	(-0.091, 0.081)
Congestion or coryza	418	42	433	49	0.013	(-0.025, 0.050)
ALRI ‡	387	10	395	15	0.012	(-0.024, 0.053)

∗ Standard Errors and 95% confidence intervals calculated by bootstrap resampling matched village pairs with 1000 iterations.

† Highly Credible Gastrointestinal Illness. The definition is included in the main text.

‡ Clinical Acute Lower Respiratory Infections. The definition is included in the main text.

C.4 ICC Calculations

We estimated the intra-cluster correlation of binary child health outcomes using the `aod` package in `R` with restricted maximum likelihood (REML) and 1000 Monte Carlo replicates to estimate the 95% confidence intervals. We estimated the intra-cluster correlation of continuous Z-scores using the `loneway` command in `Stata` (`www.stata.com`).

Table C.2 : Household- and village-level intra-cluster correlation estimates. The 95% confidence intervals for binary outcomes are based on Monte Carlo simulation. The 95% confidence intervals for Z-scores are based on asymptotic standard errors.

Outcome	Household-Level		Village-Level	
	ICC	(95% CI)	ICC	(95% CI)
Diarrhea	0.073	(0.016, 0.278)	0.015	(0.003, 0.063)
HCGI *	0.084	(0.022, 0.257)	0.026	(0.009, 0.076)
Cough or diff. breathing	0.389	(0.307, 0.475)	0.084	(0.043, 0.163)
Congestion or coryza	0.250	(0.166, 0.358)	0.015	(0.003, 0.078)
ALRI †	0.342	(0.260, 0.446)	0.055	(0.026, 0.123)
Weight ‡	0.293	(0.185, 0.401)	0.061	(0.015, 0.107)
Height ‡	0.176	(0.057, 0.295)	0.053	(0.010, 0.096)
Weight-for-height ‡	0.052	(0.000, 0.182)	0.012	(0.000, 0.035)
Mid upper arm circ ‡	0.291	(0.175, 0.406)	0.051	(0.008, 0.094)

* Highly Credible Gastrointestinal Illness. The definition is included in the main text.
† Clinical Acute Lower Respiratory Infections. The definition is included in the main text.
‡ Z-scores

References

[1] van der Laan M, Rubin D. Targeted Maximum Likelihood Learning. Int J Biostatistics. 2006;2(1):1–38.

[2] Moore KL, van der Laan M. Covariate Adjustment in Randomized Trials with Binary Outcomes: Targeted Maximum Likelihood Estimation; 2007. Available from: http://www.bepress.com/ucbbiostat/paper215.

[3] Bembom O, Petersen ML, Rhee SY, Fessel WJ, Sinisi SE, Shafer RW, et al. Biomarker Discovery Using Targeted Maximum Likelihood Estimation: Application to the Treatment of Antiretroviral Resistant HIV Infection; 2007. Available from: http://www.bepress.com/ucbbiostat/paper221.

[4] Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. Epidemiology. 1999;10(1):37–48.

[5] Rothman K, Greenland S. Modern Epidemiology. 2nd ed. Philadelphia: Lippincott-Raven Publishers; 1998.

[6] Colford J J M, Wade TJ, Schiff KC, Wright CC, Griffith JF, Sandhu SK, et al. Water quality indicators and the risk of illness at beaches with nonpoint sources of fecal contamination. Epidemiology. 2007;18(1):27–35.

[7] Sinisi SE, van der Laan MJ. Deletion/substitution/addition algorithm in learning with applications in genomics. Stat Appl Genet Mol Biol. 2004;3:Article18.

[8] Victora CG, Habicht JP, Bryce J. Evidence-based public health: moving beyond randomized trials. Am J Public Health. 2004;94(3):400–5.

[9] Mapas de pobreza en Guatemala al 2002. Guatemala City; 2006.

[10] Vaides Lopez O. Plan municipal de agua y saneamiento, Camotan, Chiquimula, 2005 - 2020; 2005.