

WASH Benefits Kenya: Pre-Specified Analysis Plan

Garret Christensen^{*1}, Clair Null² and Michael Kremer³

¹Swarthmore College, Department of Economics

²Emory University, Rollins School of Public Health

³Harvard University, Department of Economics

November 12, 2013

Abstract

PRELIMINARY DRAFT: PLEASE DO NOT CITE WITHOUT PERMISSION

The WASH Benefits project, a large water, sanitation, hygiene, and nutrition cluster randomized trial in Kenya and Bangladesh, seeks to address significant gaps in our knowledge of the effects of water, sanitation, hygiene, and nutrition on child health. In preparation for the randomized trial, the WASH Benefits Kenya project has prepared a draft pre-specified analysis plan to bind our hands when conducting analysis of the study data.

JEL Code: I15

1 Introduction

The WASH Benefits project is a large randomized control trial of water, sanitation, hygiene, and nutrition interventions in both Bangladesh and Kenya. The projects in Bangladesh and Kenya are ultimately run independently, although they are funded from the same grants¹ and there is significant cross-country collaboration and interaction. The teams have collaborated to produce a general project protocol and analysis plan, which can be seen in Arnold et al. [2013]. That paper provides a clear description of the project and the general design of the trial. This pre-specified plan here provides a more detailed analysis plan that is specific to the project in Kenya, specifically our interest in investigating the effects of externalities that would be of interest to policy makers. The project is designed to investigate the decision problem of whether a government should subsidize water, sanitation, hygiene, and nutrition interventions. This decision should be based on the biological efficacy of these interventions, their cost-effectiveness, and the benefits accruing to neighbors when an additional person nearby receives treatment.

This pre-specified analysis was written in order to bind the authors hands somewhat in the eventual analysis. By specifying functional forms of regression analysis, control variables, and subgroup analyses, we limit the potential for data mining or specification searching. This document

^{*}Send correspondence to gchrist1@swarthmore.edu. Check <http://www.ocf.berkeley.edu/~garret> for latest draft of paper.

¹Funded by a grant by the Bill & Melinda Gates Foundation to UC Berkeley. The Kenya trial was registered at clinicaltrials.gov (NCT01704105).

is not intended to completely prevent us from conducting analyses not specified herein, but by knowing which analyses were pre-specified and which were not, the reader gains extra insight into the potential reliability of our results. The practice of pre-specifying analysis has been common, if not required, of medical trials in the last decade, and there is a new movement towards making the practice more common in the social sciences. The first pre-specified plan in economics that we are aware of is Neumark [2001], from the minimum wage literature. In 2013 the American Economic Association started a registry at www.socialsciencetrials.org. See Casey et al. [2012] for an example of a recent pre-specified plan from the economics literature, which also includes an example of analysis where results could have gone either way if analysis had not been specified beforehand, and Humphreys et al. [2013] for a discussion of the topic as applied to political science. The Abdul Latif Jameel Poverty Action Lab now encourages its projects to register, and Innovations for Poverty Action (which is closely associated with J-PAL, and is implementing the WASH Benefits study) is clearly moving in the same direction. In accordance with the goals of a pre-specified analysis plan, this document was written in 2012 and 2013, begun when no data was available, and finalized when only data from a small (N=499) pilot was available and main study baseline data collection had begun. Collection of the follow-up data that will be used as described in this document will not be completed until 2015.

The three main hypotheses, or scientific objectives, are:

1. to determine if water, sanitation, and hygiene interventions aid in early child development
2. to determine if the combination of water, sanitation, and hygiene interventions are more beneficial (or cost effective) in early child development than a single intervention alone.
3. To determine if the combination of water, sanitation, and hygiene interventions plus nutrient supplements are more beneficial (or cost effective) than any of the interventions or supplements alone.

These will each be tested using a few different methods: non-parametric comparison of means, regression analysis controlling for and interacting with important baseline characteristics, and instrumental variable estimates of treatment-on-the-treated effects at a village level if we find no cross-village externalities. The null hypothesis for each intervention treatment is that there is no difference between treatment and control, and that there is no difference between combinations of treatment and a single treatment. The Kenya team will report the results of non-parametric unadjusted analyses, but will place a stronger emphasis on the results of analyses pre-specified below that include baseline control variables or control for individuals who already had the interventions delivered by the project with interactions.

The exact control variables and interactions are detailed in the remainder of this document, which is organized as follows: Section 2 summarizes the methods of analysis, Section 3 lists the outcomes and variables used in the analysis, and Section 4 lists the interactions for which we will test.

2 Analysis Methods

2.1 Unadjusted Non-Parametric Analysis

The most basic set of tests of our hypotheses consists of using completely non-parametric intention to treat (ITT) comparisons of the mean outcomes (such as height-for-age, diarrhea, child development

scores, listed exactly in section 3) in a given treatment arm to the mean outcome in the comparison group. (W, S, H, or WSH, compared to control for the first hypothesis, using the same double-sized control group in every comparison, then WSH vs. W, S, or H in the second hypothesis, and WSH+N compared to N in the third.) See the analysis plan section and Appendix 5 of Arnold et al. [2013] for more details on this method of analysis.

2.2 Basic Regression Analysis

Rather than the simple unadjusted analysis, we place primary emphasis on testing the same hypotheses using regression analysis controlling for all variables used in stratification (administrative location) as well as pre-specified baseline characteristics of sample subjects, such as age, gender, and indicators of health or environmental cleanliness before our interventions are delivered. Specifics of these regressions are provided here.

Our analysis will operate under the assumption that epidemiological externalities, both within and across villages, from treatment with our interventions are positive. (See Miguel and Kremer [2004] and Baird et al. [2011] for an example of significant positive externalities of deworming treatment.) Under this assumption, simple comparisons of individuals assigned to treatment to individuals assigned to control, as well as comparison of individuals assigned to combined interventions to individuals assigned to single interventions yield lower bounds on the true size of the effect of treatment. These lower bounds are useful to policy makers when evaluating the cost-effectiveness of a program, and will be useful in our circumstance because of the potential for externalities of treatment across individuals both within and across treatment villages. (Remember that treatments are mostly at the household or compound level, with the exception of the water intervention, which is a village-level intervention, and randomization is clustered by village.) Although we believe that cross-village externalities are likely to be quite small, we are interested in testing this assumption.

To do this we will add parameters to the simple ITT model that will control for the treatment status of individuals in surrounding villages. Under the assumption of no cross-village externalities, these parameters would have coefficients not significantly different than zero and the estimate of the effect of own-village treatment would be the same as in the simple ITT analysis, where the treatment status of neighboring villages is assumed to not affect outcomes in the village in question.

The simple ITT analysis that assumes, or does not allow for, externalities, is represented in equation 2.

$$Y_{ij} = \alpha + \sum_k (\beta_k \cdot T_{ik}) + X'_{ij} \delta + u_i + e_{ij} \quad (1)$$

Y_{ij} is the outcome of a given individual, where i refers to the village cluster and j to the individual; T_{ik} are village indicator variables for treatment arm assignment ($k \in W, S, H, N, WASH, WASH+$, with the control arm as the excluded group²); X_{ij} are baseline characteristics

²“Control group” refers to the double-sized active control group, which will have twice as many villages as any type of treatment, since power is improved easily by increasing the size of the control group—the group to which most of our comparisons are made. But it should be noted that the project also includes a (single-sized) passive control arm—passive in that it has no promoters. Respondents in the main (active) control arm receive regular visits from our health promoters, but they do not discuss our interventions, they only measure child growth by measuring mid-upper arm circumference (MUAC), the same as in treated arms. These visits are designed so that our estimates distinguish between the effect of visits from health quasi-professionals *per se* (e.g. cleaning up the yard because a visitor is stopping by) and the effect of the use of interventions as encouraged by the promoters, which we consider to be the far more important component of any resulting change. However, we will also test for a difference between

such as age and gender (listed more completely below); u_i is a village level error term; and e_{ij} is an individual error term. Standard errors will be clustered at the village level, the level of treatment, for all analyses.

But instead of this simple equation that assumes no cross-village externalities, we will instead place primary emphasis on the results of a model described in equation 2.

$$Y_{ij} = \alpha + \sum_k (\beta_k \cdot T_{ik}) + X'_{ij} \delta + \sum_d \sum_l (\gamma_{dl} \cdot N^l_{dij}) + \sum_d (\phi_d \cdot N_{dij}) + \sum_d (\xi_d \cdot S_{dij}) + \sum_d (\eta_d \cdot M_{dij}) + \sum_l (\psi_l \cdot S^l_{ij}) + u_i + e_{ij} \quad (2)$$

Y_{ij} , T_{ik} , X_{ij} are all as before, and is the number of treated individuals receiving interventions of type l inside or outside the village within distance d from the individuals home. ($l \in W, S, H$, noting that unlike the village treatment arm indicator T , this excludes nutrition and does not separately control for WASH or WASH+ villages, which here would instead have the variables for W , S , and H all non-zero—this is an assumption that nutrient supplementation does not lead to cross-village externalities, and that water, sanitation, and hygiene externalities are additive and not somehow interactive or synergistic when delivered in combination). The terms are summed over different d , of which there will be four mutually exclusive concentric radii determined by the quartiles of the distribution of distances from each village center to the nearest other village center, divided by two (our best estimate at the radius of a village). N^d_{ij} is the number of unenrolled individuals as well as control subjects living within the same concentric radii, in order to roughly gauge the fraction of the total nearby population that is treated. S^d_{ij} and M^d_{ij} are the number of schools and markets within a given distance from a village. Lastly, S^l_{ij} is fraction of the given respondents children who attend school in a village assigned to treatment status l .

These cross-village externality measures are partly limited by the data that is feasible to collect in our situation. We will have information on our respondents, as well every other compound in their village. Based on our first block of enrollment, we expect to enroll roughly 60% of the villages in administrative areas where we work. We will have detailed population data (number of persons residing in each compound) on everyone in perhaps 75% of the villages in subdistricts where we work, as we exclude urban areas and areas that already have piped water or chlorine dispensers, and will have no data on the populations of these excluded villages. In villages we enroll, we expect to enroll households in roughly 8% of the compounds. In the entire area will enroll households in perhaps 3-5% of the compounds in the entire area. That is to say, with the exception of the water intervention, where most of a village is treated because everyone is encouraged to use the chlorine dispensers, and all compounds with a child under 5 will receive household chlorine bottles, only a small fraction of total individuals will be receiving our interventions, and we expect cross-village externalities to be small to non-existent.

However, if the coefficients are significantly different from zero, then we can still interpret the simple own-village estimate as a lower-bound on the estimate of the true effect of treatment with the weak assumption that cross-village externalities are positive. Positive externalities are a major reason for a government to subsidize a program or behavior that would otherwise be under-provided. The test for the presence of cross-village externalities is simply to look at the N^l , N , and S , S^l and M coefficients in equation 2. Due to randomization, under the null hypotheses of no cross-village externalities, these coefficients will be insignificant. We will run individual t- and joint F-tests for

these types of control groups using a simple regression along the lines of Equation 1. We do not expect major differences in health outcomes, though perhaps a few measures of uptake regarding compound cleanliness may be affected. If there are significant differences, we may include the passive control arm as an additional element of k .

significance. If they are not significant, we cannot completely rule out the idea that cross-village externalities exist in some form that is not captured by geographic distance, school, or market-based interactions, but we will at least have flexibly tested several of what we believe are the most reasonable avenues for these externalities, and found none. If these coefficients are significant, we can conclude that cross-village externalities exist. If cross-village externalities are found to exist, with the weak assumption that they are positive, comparisons between treatment and comparison arms using equation 1 will be lower bounds on the impact of a generalized program since we will be missing any contribution from cross-village externalities, and to the extent that the comparison groups are made better off than they would be in the absence of a program we will underestimate the difference between the treatment group and the comparison group.

2.3 Intention to Treat and Treatment on the Treated, Take-up

If, as expected, cross-village externalities are estimated to be not significantly different than zero by the process described above, we will also report TOT estimates at the village level. This will be done with two-stage least squares instrumental variable techniques described in Angrist et al. [1996]. The first stage will be to use village treatment assignment as an instrument for average village-level take-up, and the second stage will be to measure the effect of instrumented village take-up on average village outcomes. This is shown in the following two equations:

EQN 3 and 4:

$$W_i = \alpha + \beta \cdot T_i + X_i' \delta + e_i \quad (3)$$

$$Y_i = \alpha + \beta \cdot \hat{W}_i + X_i' \delta + e_i \quad (4)$$

Where W_i implies village-level take-up of a given intervention and the estimated values of take-up, instrumented with treatment status. Notice that the coefficients are not summed over different treatment coefficients in one regression—we will run separate regressions for each type of treatment take-up (W, S, H, and N), where each regression will compare the relevant treated arms to the control group. (For example, water take-up would regress the W, WASH, and WASH+ arm against the active control.) These regressions are done at the village level, because in the case of an individual, the exclusion restriction is unlikely to hold: assignment to a treatment village likely affects child health through within-village externalities, rather than only through one's own likelihood of receiving treatment. (If we didn't expect this to be the case, we would have less reason to cluster randomization at the village level.) At the village level, however, the exclusion restriction should hold if we find no cross-village externalities in the analysis described above: village assignment to treatment affects village child health only through the likelihood of the village receiving treatment. Village-level analyses are clearly expected to have less power than individual-level analyses due to the smaller sample size, but they may still be informative.

For all of the outcomes and control variables, village-level statistics will always be simple averages of the values of the study population in each village. For take-up, an index will be created for successful adoption of the interventions in each treatment arm (see details below), and this will be averaged across the study population in each village. Only the relevant take-up measure will be used to instrument for a given treatment (adoption of water chlorination technology for treatment in the water arm, for example.) Estimates and standard errors will be weighted by number of respondents to account for the averaging of all respondents within a village.

For water treatment, take-up will be measured exclusively by detection of 2 mg/L of free chlorine in a households stored water. For sanitation, take-up will be measured with three indicators:

observing no human feces throughout the compound, observing the latrine drop-hole covered, and observing no feces on the floor of the latrine. Hygiene uptake will be measured with two indicators: no visible dirt on the hands of the respondent and both of the tippy-taps are in working order stocked with both soap and water in them at the time of endline. Nutrition uptake will be measured by the fraction of empty LNS packets retrieved from the respondent. In the case of sanitation and hygiene, with multiple potentially useful indicators of uptake, the indicators will be averaged together (equally weighted) in one regression, and we will also test each indicator alone.

3 List of Variables and Regressions

3.1 Health Outcomes

Many of the most important outcomes of our study are health outcomes. These include both hard measures obtained by trained field enumerators and caregiver-reported measures. The following is a list of the outcomes WASH Benefits plans to analyze with the methods described above. All of the variables listed below in the primary, secondary, and tertiary health outcomes sections will be analyzed using each of the equations above (1-6) as well as unadjusted comparison of means.

3.1.1 Primary Health Outcomes

- Height-for-age Z-scores (HAZ) Anthropometry Since age and anthropometry are difficult to measure precisely, a note on their collection. Age will be collected by survey enumerators asking caregivers for child's date of birth, verified by clinic cards if at all possible, then comparing that to reported age in months and years, and enumerators stating their opinion on which of these is the more reliable source. Anthropometry of all children of the respondent under 36 months at baseline will be collected in a central location in each village by a set of dedicated field officers trained directly by PhD nutritionists.
- ASQ child development scores (motor, communication, personal/social) Exact measurements not yet chosen
- Caregiver-reported diarrhea and other illnesses Diarrhea and Illness Symptoms: We collect information using caregiver reports for the main study child, as well as any children under 36 months belong to the respondent, as well as any children under 36 months in the compound cared for by others (the latter referred to as secondary respondents). We collect information separately for the following time periods: today, yesterday, day before yesterday, and the last 7 days. We will place primary emphasis on the diarrhea results from all children under 36 months belonging to the respondent using one-week recall, although we will also report results from other disease symptoms, from only study children, from all children including secondary respondents (other mothers of children under 36 months residing in the same compound as the respondent), and using 2-day recall.
 - Fever
 - Diarrhea (colloquial definition: kuhara in Swahili)
 - Diarrhea (case definition: three or more defecation in 24 hours plus watery or soft stools)
 - Blood in stool

- Coughing
- Wheezing
- Congestion
- Bruising
- Toothaches

3.1.2 Secondary Health Outcomes

- Proportion of children stunted ($HAZ < -2$)
- Anthropometry:
- Enteropathy biomarker measurements (Subsample Only)
- Lactulose / mannitol dual sugar permeability test
- Antibody titers (Total IgG, IgG EndoCAb)

3.1.3 Tertiary Health Outcomes

- Weight-for-age Z-scores Childs weight (measured as mother and child's weight minus mother's weight):
- Proportion of children underweight ($WAZ < -2$)
- Proportion of children wasted ($WHZ < -2$)
- Head circumference Z-scores
- Mid-upper arm circumference Z-scores
- Parasitic infections at 2 year follow-up
- Soil transmitted helminths (Ascaris, Trichuris, Hookworm)
- Protozoans (Giardia, Cryptosporidium, E. histolytica)
- Verbal Communicative Development Inventory (CDI) at 1 year follow-up
- WHO Motor milestones at 1 year follow-up
- Appetite (Control + Nutrition arms)

3.2 Uptake

Uptake of our interventions is an interesting outcome measure in and of itself, in addition to its use in the instrumental variables strategy described above. So in addition to the main uptake measures described above for use in the instrumental variables strategy, we will also report several other measures of uptake. Though we will still estimate uptake using equations 1 and 2, we place more importance on the effect of a given type of treatment on the corollary uptake (i.e. hygiene treatment on hygiene uptake). A model where a respondent had a fixed amount of time to spend on all aspects of WASH would predict that those treated in the water arm might substitute away from washing their hands or improving the sanitation conditions of the compound, but a priori we assume it is more likely that those in water-treated arms will change their water-related behavior more than those in sanitation or hygiene arms.

However, in certain instances, our hardware is essentially required for a respondent's uptake to be registered. In these cases we will limit equations 1 and 2 to include only the combined treatment arms and the relevant single arm. For example, we expect that no one in our sample who is not in the N or WASH+ arms to eat the same LNS product that we provide. We also expect construction of tippy-taps to be extremely rare. Even if just a few individuals go out of there way to do so, (and thus we avoid perfect multi-collinearity) the regression will likely be based on so few data points as to be highly suspect. Thus we will limit the sample to those receiving the intervention, and compare combined arms to single arms. These instances are denoted with an asterisk below.

- Handwashing: Measurements of Take-up
 - Simple presence of tippy-taps *
 - Physical observation of Tippy-Taps (in working order, ground nearby worn-down or wet from frequent use.)*
 - Observed hand cleanliness
 - Use of soap (or soapy water) in handwashing demonstration
 - Cleaning both hands in handwashing demonstration
 - Of 5 critical times for handwashing, respondent correctly volunteers all times without prompting
- Sanitation: Measurement of Take-up and Adult Sanitation Behavior
 - Household has a toilet facility
 - Compound/household has child potty
 - Household has dedicated tool to clean up feces
 - Latrine drop hole covered
 - Stool visible on latrine floor
 - Feces disposal tool clean/accessible
 - Contact with feces/open defecation poses risk to health
 - Animal/Human feces observed in/around the house/compound
 - Last child feces disposed of appropriately (i.e., latrine used for defecation, or potty emptied into latrine)

- What was done with last older child feces
- Water: Measurement of Take-up
 - Free chlorine detected in household stored water
 - Primary or Secondary water source reported to have a chlorine dispenser
 - Current stored water treated with a chlorine dispenser
 - Current stored water treated with household bottled chlorine
 - E. coli count in household sample
- Nutrition:
 - Self-reported frequency of consumption of LNS*
 - Observed fraction of empty LNS packets collected*
 - Self-reported exclusive breastfeeding of study child up to age 6-months
 - Maximum age of study child when some breastfeeding still practiced
- Additional Questions
 - Study child has obtained all appropriate vaccines, according to clinic card
 - Household reports children sleeping under bednets
 - Household heard of or used ORS
 - Household seems certain of location of nearest health facility

3.3 Substitutes and Complements

In addition to hard measures of health outcomes and uptake of our interventions, we are interested in how other behaviors change after the interventions have been delivered. Treated persons may divert their effort away from practices they had been engaged in after our intervention makes another behavior easier. Some of these activities are listed above in the respective water, sanitation, or hygiene section, but are repeated here as well.

- Deworming medicine purchased
- Percentage of required vaccinations obtained
- Water boiled less due to availability of other treatment
- Water filtered with widely available VF Family LifeStraw less due to availability of other treatment
- Time spent collecting water
- Mosquito net usage
- Observed child cleanliness: child's hands
- Ambulatory children are wearing shoes
- Self-reported exclusive breastfeeding of study child up to age 6-months
- Maximum age of study child when some breastfeeding still practiced

3.4 Planned Control Variables

The X_{ij} variables are important to specify beforehand since in practice significance levels can change with different sets of included variables. We are planning to stratify randomization only on Location, a Kenyan administrative boundary that consists of widely varying average of 40 villages. Thus we will include a dummy variable for each Location. Our primary regression specifications will include other predetermined control variables that may decrease standard errors. Due to the randomized design, these should all be identically distributed across treatment arms, so controlling for these additional variables is not needed to prevent bias, but may increase the efficiency of our estimation. These extra control variables (the Xs in equations 1-6) will consist of :

- Location (Kenyan administrative region used for stratification)
- Field Officer
- Time between intervention delivery and survey (due to the nature of the roll-out, there may be some variation in this, but not much.)
- Month of survey—to account for dry/rainy season variation
- Age of child at time of intervention, by quintiles
- Gender of child
- Age of mother, by quintiles
- Mother’s education level (binary—had any secondary schooling)
- Mother’s Swahili literacy
- Mother’s English literacy
- Number of households within the study compound
- Total number of children under 36 months in the study compound
- Tin roof ownership

4 Planned Interactions and Distributional Analysis

In addition to determining the average treatment effect for our entire sample, we are interested in examining the heterogeneity of the treatment effects. We intend to do this primarily through sub-group analysis. As we are not deliberately powered to detect interaction effects, we are not expecting significant effects, but will investigate them nonetheless. If certain subgroup effects turn out significant, we will potentially investigate distributional effects using quantile regressions (see Angrist Mostly Harmless?).

The following is a list of the measures we intend to use for sub-group analysis. To avoid confusion it should be emphasized that all measurements described below are obtained at baseline, before the

intervention. (Some of these same measurements, obtained after the intervention, are of interest as outcomes.) Regressions will be of the following form:

$$Y_{ij} = \alpha + \sum_k (\beta_k \cdot T_{ik}) + \sum_k (\lambda_k \cdot T_{ik} \cdot Z_{ij}) + \rho Z_{ij} + X'_{ij} \delta + \sum_d \sum_l (\gamma_{dl} \cdot N_{dij}^l) + \sum_d (\phi_d \cdot N_{dij}) + \sum_d (\xi_d \cdot S_{dij}) + \sum_d (\eta_d \cdot M_{dij}) + \sum_l (\psi_l \cdot S_{ij}^l) + u_i + e_{ij} \quad (5)$$

Where the treatment and cross-village externality terms are the same as in equation 2, with the addition that Z is the variable with which we will interact all of our treatment variables. These variables are described below. In the case that tests of equation 2 reveal that cross-village externalities are insignificant, we will instead test interactions with the simpler equation 6.

$$Y_{ij} = \alpha + \sum_k (\beta_k \cdot T_{ik}) + \sum_k (\lambda_k \cdot T_{ik} \cdot Z_{ij}) + \rho Z_{ij} + X'_{ij} \delta + u_i + e_{ij} \quad (6)$$

Equations 5 and 6 allow for a Z variable to interact with all k treatments. While in some instances this might be sensible (e.g. age of the study child) we will not test every Z interacted with every treatment. For example, variables clearly related to the effectiveness of water treatment, such as type of water source, will only be tested for interactions with water treatment. This will be pre-specified for each variable below. In most cases the interactions are indicator variables, so no further discussion of functional form is necessary. In the case where the interaction variable is more complicated, details are provided.

The first group of these analyses are based on pre-existing levels of treatments. We expect that the benefit received from interventions will depend on the quantity of the interventions, of both the same and other types, that the household already had. In certain situations, there is unlikely to be any biological mechanism for further health improvements: if household water is already chlorinated, adding more chlorine is unlikely to lead to significant additional benefits. On the other hand, good WASH conditions at baseline may be an indicator that a respondent is diligent or motivated, and they may be likely to more consistently use our interventions, leading to greater health benefits.

Effect of interventions based on pre-existing levels of treatments at baseline.

- Initial water quality. We expect that households already treating their water sufficiently will not benefit from the water treatment arm, so we will interact baseline chlorination with all treatments using equation 5 or 6, since all treatments could conceivably be more effective in the presence of clean water.
- Effective levels of Free Chlorine detected in stored water
- Use of the VF Lifestraw: Respondent still has VF LifeStraw filter, reports frequent use, knows how to backflush mechanism, and filter appears recently used (moist or not dusty)
- Initial water source. Although we would like to install a chlorine dispenser at every water source in treated villages, this is not practical. We are excluding private sources, sources where landowners object to free use of their source, and sources that are dry too much of the

year to be of use. Thus some fraction of treated respondents will inescapably not have their sources treated with a chlorine dispenser. We will interact these characteristics with water treatment only.

- Baseline primary source was surface water. (Surface water may be more likely to have many possible collection points as opposed to one centralized source with a centrally located dispenser.)
- Primary Source did/did not receive a dispenser
- Secondary Source did/did not receive a dispenser
- Initial handwashing habits. It is possible that mothers who consistently wash their hands with soap at baseline may gain little from the hygiene intervention. But perhaps this handwashing is an indicator of diligence that would lead to high uptake of other interventions, so we will interact these with all treatment indicators.
- Both mother's and child's hands (fingerpads and palms) are observably clean at baseline
- Mother reports all five critical times for handwashing without prompting
- Initial sanitation conditions. Will interact with all treatment indicators.
- Household does not have a latrine at baseline, was assigned to receive one from IPA
- Latrine shared with other households/compounds
- Household already owns potty
- Household has a dedicated pooper-scooper tool (i.e., a tool used only for feces disposal)
- Beliefs regarding determinants of sanitation
- Initial compound cleanliness. Will interact with all treatment indicators.
- Number and type of livestock kept in compound (Quintiles of total livestock, separately presence non-presence of chickens)
- Feces from humans (separately for humans or livestock) visible anywhere in compound at baseline
- Initial nutrition conditions, from validated food insecurity scale.
- Child was breast fed colostrum immediately after birth
- Does child under 6 mo. consume anything other than breast milk?

Other prominent baseline characteristics that may interact with treatment. All of these will be interacted with all types of treatment.

- Age of study child. There is a significant amount of hypothesizing about a golden window of health opportunity at very young ages, so it will be worthwhile to test whether younger children benefit more.
- Gender. Parents may be more motivated to use the interventions to improve the health of certain children.
- Density of neighbors. (It may be that respondents nearer to their neighbors benefit less than others. Using our data from village census we will estimate the population density within one half kilometer, and separately interact two functional forms: quintiles of density using the study population, and a quadratic polynomial of population/sq. km.)
- Age and number of older siblings and those in compound. Small children may pick up more diseases from their older siblings bringing them home from school, or from poor sanitation practices of siblings too young to be attending school. We will test this interaction in two ways: with a quadratic polynomial of the number of older siblings in the compound, and with separate quadratic polynomials of children under school age (0-6) and school age (7-18).
- Mother's self-reported literacy. This will be interacted separately for both Swahili and English literacy.
- Mother's score on Silly Sentences. We will test a quadratic polynomial of the score on our silly sentences literacy and reading comprehension test, and by grouping mothers into quintiles.
- Maternal and Paternal level of schooling. We will look separately at the binary indicator of whether each parent self-reports as having any secondary schooling.
- Social Networks—mothers who visit frequently with other intervention recipients or promoters may take up the interventions at a higher rate, and mothers who are more socially active or confident may do the same.
- Respondent interacts with promoter multiple times per week prior to study.
- Distance of house from promoters house. We will test this separately with quintiles and with a quadratic polynomial of the distance.

References

- Joshua Angrist, Guido Imbens, and Donald Rubin. Identification of causal effects using instrumental variables. *Journal of the American statistical Association*, 91(434):444–455, 1996.
- Benjamin F. Arnold, Clair Null, Stephen P. Luby, Leanne Unicomb, Christine P. Stewart, Kathryn G. Dewey, Tahmeed Ahmed, Sania Ashraf, Garret Christensen, Thomas Clasen, Holly N. Dentz, Lia C. H. Fernald, Rashidul Haque, Alan E. Hubbard, Patricia Kariger, Elli Leontsini, Audrie Lin, Sammy M. Njenga, Amy J. Pickering, Pavani K. Ram, Fahmida Tofail, Peter Winch, and John M. Colford Jr. Cluster-randomized controlled trials of individual and combined water, sanitation, hygiene, and nutritional interventions in rural bangladesh and kenya: The wash benefits study design and rationale. *BMJ Open*, 3(8):e003476, 2013.
- Sarah Baird, Joan Hamory Hicks, Michael Kremer, and Edward Miguel. Worms at work: long-run impacts of child health gains. *Unpublished mimeo*. <http://elsa.berkeley.edu/~emiguel/pdfs/miguel.wormsatwork.pdf>, 2011.
- Katherine Casey, Rachel Glennerster, and Edward Miguel. Reshaping institutions: Evidence on aid impacts using a preanalysis plan. *The Quarterly Journal of Economics*, 127(4):1755–1812, 2012.
- Macartan Humphreys, Raul Sanchez de la Sierra, and Peter van der Windt. Fishing, commitment, and communication: A proposal for comprehensive nonbinding research registration. *Political Analysis*, 21(1):1–20, 2013.
- Edward Miguel and Michael Kremer. Worms: identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, 72(1):159–217, 2004.
- David Neumark. The employment effects of minimum wages: Evidence from a prespecified research design the employment effects of minimumwages. *Industrial Relations: A Journal of Economy and Society*, 40(1):121–144, 2001.