Spring Cleaning:
A Randomized Evaluation of Source Water Quality Improvement[*]

Michael Kremer
Harvard University,
Brookings Institution, and NBER

Jessica Leino
University of California, Berkeley

Edward Miguel
University of California, Berkeley
and NBER

Alix Peterson Zwane
University of California, Berkeley
and Giannini Foundation

First draft: July 2006
This draft: December 2006

Abstract: Water-related diseases, particularly diarrhea in young children, kill two million people annually. To address this problem, donors and governments often provide infrastructure such as communal standpipes, wells, and protected springs in rural areas, where piping water into homes is infeasible. We study the impact of source water quality improvements achieved via spring protection in rural Kenya using a randomized evaluation. Spring protection leads to large improvements in source water quality as measured by the fecal indicator bacteria *E. coli*. Water quality gains at the home are smaller on average, but this finding depends critically on households' water source choices. At households that only used the sample spring at baseline, 71% of the spring water quality benefits are translated into home water gains, suggesting that re-contamination in transport and storage may be less of a concern than is sometimes claimed. Consistent with this view, the home water quality gains from spring protection are no larger for households with better baseline sanitation or hygiene knowledge. Changes in household water source choices after spring protection are used to derive revealed preference estimates of the willingness to pay for improved water quality using a travel cost approach. The average willingness to pay for the moderate gains in home water quality due to spring protection is at least US$3.27 per household per year. We find no significant child health effects of spring protection.

---

[*] This research is supported by the Hewlett Foundation, USDA/Foreign Agricultural Service, International Child Support (ICS), Swedish International Development Agency, the Finnish Fund for Local Cooperation in Kenya, Google.org, and the Gates Foundation. We thank Alicia Bannon, Lorenzo Casaburi, Anne Healy, Jie Ma, Clair Null, Owen Ozier, Camille Pannu, and Heidi Williams for excellent research assistance and the field staff, especially Polycarp Waswa and Leonard Bukeke. Jack Colford, Alain de Janvry, Andrew Foster, Michael Greenstone, Danson Irungu, Ethan Ligon, Steve Luby, Enrico Moretti, Kara Nelson, Sheila Olmstead, Judy Peterson, Rob Quick, Elisabeth Sadoulet, Sandra Spence, Ken Train, and Polycarp Waswa provided helpful comments. Preliminary draft, comments welcome. All errors are our own.
Corresponding author: Alix Zwane (email: zwane@are.berkeley.edu).

1 Introduction

The sole quantitative environmental target in the United Nations Millennium Development Goals (MDGs) is the call to "reduce by half the proportion of people without sustainable access to safe drinking water" (General Assembly of the United Nations 2000). Meeting this goal will require providing over 900 million people in rural areas of less developed countries with either household water connections, which are often impractical because of dispersed settlement, or access within one kilometer to a constructed public water point (standpipe, borehole with hand pump, protected spring, protected well or rainwater collection point).

A central rationale for promoting safe drinking water is the persistently high level of water-related morbidity and mortality in less developed countries. The global health burden of diarrheal disease in particular is tremendous and falls disproportionately on young children. Diarrheal disease, the third leading cause of infant mortality following malaria and respiratory infections, kills approximately 2 million people annually and accounts for perhaps 20% of deaths among children under age five (Kosek *et al.* 2003). Diarrheal diseases are transmitted via the fecal-oral route, meaning that they are passed by drinking or handling microbiologically unsafe water that has been in contact with human or animal waste, or because of insufficient water for washing and bathing.

However, there remains active debate and little conclusive evidence regarding how best to tackle this scourge. To start, it remains unclear whether investing in the water sector is even the most effective way of reducing the diarrheal disease burden. Randomized trials have established that several health interventions—including increased breastfeeding, immunization, oral rehydration therapy (ORT), and micronutrient supplementation—are both effective and cost-effective in preventing diarrhea (see Hill *et al*. 2004).[1] Rotavirus kills about 600,000 children annually and although a vaccine exists few children receive it in the poorest countries.

Even within the environmental health sector, there is little consensus on the relative cost-effectiveness of different water, sanitation, and hygiene interventions when piping water into homes is impractical. For instance, there remains debate about whether improving water quality at the source, increasing the quantity of water available, or point-of-use (in home) treatment of microbial contaminants is most cost-effective. While several studies from the 1980s find that source water quality interventions reduce diarrheal morbidity, a recent strand of the academic literature has increasingly emphasized the importance of water quality at the point-of-use (POU). The efficacy of POU treatment has been convincingly demonstrated in several settings, but it is unclear whether most households are willing to use such treatments, and how much they are willing to pay for them. In the face of this ongoing debate, donor funding in the rural water sector continues to be overwhelmingly directed at source improvements. [2]

This paper evaluates the impact of source water quality improvements achieved via spring protection. Spring protection seals off the source of a naturally occurring spring and encases it in concrete so that water flows out from a pipe rather than seeping from the ground, where it is vulnerable to contamination from surface runoff, improving water quality at an already existing

---

[1] Exclusive breastfeeding of infants is widely accepted as a means of preventing diarrhea in infants up to six months of age and continued breastfeeding for older children is also protective (Raisler *et al.* 1999, Perera *et al.* 1999, WHO Collaborative Study Team 2000). Many public health experts believe vaccines have a valuable role to play in preventing at least two diarrheal diseases, rotavirus and cholera (Glass *et al.* 2004, WHO 2004). ORT appears to have been responsible for reductions in diarrheal mortality (Miller and Hirschorn 1995, Victora *et al.* 1996 and 2000). Micronutrient supplementation, including with zinc and vitamin A,has also been found to have positive impacts (Grotto 2003, ZICG 1999 and 2000, Black 1998, Ramakrishnan and Martorell 1998, Beaton *et al.* 1993).

[2] Currently about US$10 billion is spent annually to improve water and sanitation in less developed countries (United Nations 2003), through numerous initiatives, such as the US$1 billion European Union Water Facility. In rural Africa, these funds are overwhelmingly spent on providing community-level resources like water taps or shared wells (UN-Water/Africa 2006). Among the US$5.5 billion the World Bank invested in rural water and sanitation programs from 1978-2003, nearly all focused on improving source water supply and quality through interventions such as well-digging and spring protection, while 3% went to sanitation improvements, less than 1% on hygiene promotion, and only a small portion to household point-of-use (POU) interventions (Iyer *et al*. 2006).

source. This is a widely used technology in Sub-Saharan Africa (Mwami 1995, Lenehan and Martin 1997, UNEP 1998), though it is unsuitable for the most arid regions (UNEP 1998).[3]

Using a randomized impact evaluation approach, in which spring protection is phased-in across 200 springs in a randomized order, we estimate impacts on source water quality, household water quality, child health, and on household water collection choices and other health behaviors. Our approach differs from the existing literature on source water quality interventions in several ways. First, unlike many other studies, we are able to isolate the impact of a single treatment rather than a package of services. Second, we use a randomized design and a large sample size, and are able to take intra-cluster correlation into account. Third, rather than assuming or simulating ex post contamination between the source and the home, we have detailed longitudinal data on water quality at both points, where water contamination is measured by the fecal indicator bacteria *E. coli*. We are thus able to directly assess the extent to which source water quality improvements at springs translate into household water quality gains, and to evaluate the claim that source water quality improvements are most valuable in the presence of pre-existing household access to improved sanitation and good hygiene practices.

In the second component of the analysis, we use data on how household behaviors – most importantly, the choice of water source – change in response to source water quality improvements in a context where many households can choose from among multiple water sources. We develop a formal economic framework of water source choice, in which households trade-off the distance walked to a water source against water quality. This framework highlights the importance of understanding endogenous household sorting among water sources in the econometric analysis, and

---

[3] Spring protection is generally undertaken by outside donors or the central governments in this region because both law and custom require that private landowners allow public access to water sources on their land, leaving no private incentive to improve a water source in the hope recouping the investment via user fees. Due to limited powers of coercion, local authorities and community organizations also face difficulties overcoming free-rider rider problems in collecting contributions for infrastructure improvements. These collective action problems mean that even local public water investments with positive returns are often not undertaken.

allows us to develop revealed preference estimates of average household willingness-to-pay for source water quality improvements using a hedonic travel cost approach. To our knowledge this is the first such revealed preference estimate of household valuation for water quality improvements in a less developed country context.

In our first empirical result, we find that spring protection is very effective in improving the quality of water at the source. Among "sole-source users", those households that collected all of their drinking water from the sample spring at baseline, spring protection is also highly effective at improving household water quality. The high degree of pass through from source water quality gains to the home suggests that recontamination in transport and storage may be less of a concern in rural Africa than is sometimes claimed. The exogenous variation in source water quality caused by the spring protection program allows us to use an instrumental variables approach to address econometric concerns about both measurement error (in water contamination) and omitted variables that could bias estimates of the correlation between source water quality and home water quality.

However, among the "multi-source users" – those households that collected at least some of their water from sources other than the sample spring at baseline – estimated home water quality gains are much smaller. We find that these multi-source users dramatically shifted their water collection trips towards protected springs after the intervention. We show in the formal model of source water choice that this sorting may account for the lack of observed home water quality gains, to the extent that some alternative water sources are less contaminated than the sample springs. A possible implication is that programs improving all water sources in an area might have larger impacts on home water quality than programs that only improve a subset of sources. Alternatively, increasing the take-up of point-of-use water treatment may be desirable in contexts where improving all local water sources is costly.

We next estimate willingness to pay for improved source water by analyzing how multi-source user households change their choice of water source – and in particular, the distance they are willing to walk to collect water – in response to the improvements generated by spring protection. These are lower bound estimates on average household willingness to pay because households might be willing to walk even farther for improved source water than we observe. Taking the local unskilled agricultural labor wage as the most plausible opportunity cost of time for the adult women who collect most household water, we estimate that households are on average willing to pay at least about US$3.27 per year for the household water quality improvements generated by spring protection, or about one and a half days' wages. This is a modest figure, but the home water quality gains from spring protection are only moderate, and of course this remains a bound on the true valuation. The figure could have a range of uses by those interested in either source water quality improvements or point-of-use technologies in rural Africa; for example providing guidance on the magnitude of realistic user-fees at communal water sources.

In our data, there is no evidence that the limited home water quality gains we observe in the multi-source user sample are due to endogenous behavioral change or crowding out of other protective measures such as boiling drinking water or in-home chlorination. Nor does pre-program access to improved sanitation or hygiene knowledge appear to allow households to better translate source water quality improvements into household water quality gains, further evidence against the view that there is extensive recontamination of water during transport and storage.

Consistent with both the moderate home water quality gains and modest household willingness to pay for improved water we estimate, we find no evidence of large positive spring protection impacts on the average health or nutrition of young children in treatment households. Even for sole-source user households alone, a group that did experience significant home water quality gains, there are no statistically significant impacts on child health or nutrition, and this is true both for diarrhea as well as for child weight and height. This finding is consistent with an epidemiological model in which the primary causes of diarrhea are not waterborne. However, statistical power is a concern with these estimates, and ongoing project data collection will allow us to estimate child health and nutrition impacts more precisely.[4]

2 Related Literature

Reviews on the health impact of environmental health interventions to combat diarrheal diseases include Blum and Feachem 1983, Esrey *et al.* 1985, Esrey and Habicht 1986, Esrey *et al.* 1991, Rosen and Vincent 1999, and Fewtrell *et al.* 2005).[5] Two especially influential review articles (Esrey 1996, Esrey *et al.* 1991) are frequently cited as evidence for the relative importance of sanitation investments and hygiene education over the provision of improved water quality (*e.g*. USAID 1996, Vaz and Jha 2001, World Bank 2002). Esrey *et al.* (1991) attempt to separate the relative impacts of water supply, sanitation, and hygiene education interventions on diarrheal morbidity using observational data. They conclude that the median reduction in diarrheal morbidity from either sanitation supply or hygiene education provision is nearly twice the median reduction from an investment in water quality alone, or investments in water quantity and water quality together. Using multivariate regression analysis of household infrastructure status and diarrhea prevalence from several countries, Esrey (1996) reaches a similar conclusion: benefits of improved water quality occur only in the presence of improved sanitation, and only when a water source is present within the home (e.g., piped water). However, as a result of the observational nature of Esrey's (1996) data, these results are subject to omitted variable bias (confounding) of unknown magnitude.

More recent meta-analysis in the epidemiological literature (Fewtrell *et al*. 2005) reports that source water quality improvements, sanitation, and hygiene programs, along with point-of-use water

---

[4] The current study is one component of a larger project by the authors examining the relationships among water quality at the source, point-of-use, water quantity, and health, which taken together may provide guidance on whether there is scope for some readjustment of funding priorities in the rural water sector.

[5] As Briscoe (1984) and Okun (1988) emphasize, the welfare gains associated with infrastructure provision can extend far beyond mortality and morbidity impacts: for example, women's time may be freed from water transportation duties and thus other activities facilitated. We formalize this idea below.

treatment can all effectively reduce diarrhea, with point-of-use treatment the most effective of these interventions, in contrast to the conclusions in Esrey et al. (1991). In an argument related to Esrey et al (1991), Fewtrell *et al*. (2005) conclude that recontamination during transportation and storage likely make point-of-use water treatment more effective than source water quality interventions. Similarly, Wright *et al*. (2004) analyze 57 studies that measured both source and in-home water quality, and conclude that improvements in source water quality are often compromised by post-collection contamination. However, these evaluations of source water quality investments remain less methodologically rigorous than evaluations of point-of-use water treatment.[6] Moreover, to our knowledge there is no study in which household water quality has been measured following exogenous variation in source water quality, and no direct comparisons of the effectiveness of point-of-use water treatment and source water quality interventions in the same study setting have been made. In this paper, we directly measure the extent to which source water quality leads to home water quality gains using a longitudinal household dataset and exploiting experimental variation in source water quality. Future work will estimate effectiveness of point-of-use technologies in the same study area.

We are also able to contribute to a second literature by developing a novel revealed preference estimate of willingness to pay for improved water quality in a less developed country

---

[6] There are just two prospective studies of source water quality interventions that suggest positive impacts of these interventions on child health. Aziz et al. (1990) study the impact of an intervention in Bangladesh that simultaneously provided multiple interventions, including water pumps, hygiene education, and latrines, to two intervention villages (820 households), and compare them with three control villages (750 households), separated by about five kilometers. The published article does not mention if these villages were randomly selected. Following the intervention children between six months and five years of age in the intervention area experienced 25% fewer episodes of diarrhea than those in the comparison area. An almost identical reduction was observed after pumps had been installed but prior to the construction of latrines, which is consistent with a small treatment effect of improved sanitation beyond that achieved by well construction.

Huttly et al. (1987) study the impact of the provision of borehole wells with hand-pumps, pit latrines, and health education on dracunculiasis (guinea worm disease), diarrhea, and nutritional status in Nigeria in 1983-86. The study compared three intervention villages (850 households) and two comparison villages (420 households). Because of implementation difficulties, their results largely reflect the effect of the installation of wells with pumps. The prevalence of wasting (defined as less than 80% of desirable weight-for-height) among children under three years of age declined significantly in the intervention villages. Generalizing the positive results in these studies to other settings is hampered by their small village-level sample sizes (both include only five villages in total), and the fact that they evaluate interventions that improved both water quality and quantity simultaneously (by providing wells).

context using a travel cost approach. Understanding the determinants of household water demand was a research focus in the 1990s, and contingent valuation studies sponsored by the World Bank in several countries estimated both average stated willingness to pay for home piped water connections and the heterogeneity in this demand (World Bank Water Demand Research Team 1993).

However, the relative shortcomings of such stated preference contingent valuation approaches to measuring the use value of non-market goods are well-known (Diamond and Hausman 1994). Survey respondents in contingent valuation studies do not face a real budget constraint when telling survey enumerators their willingness to pay for hypothetical goods or services, and may strategically overstate their true valuation (to be polite, or in an attempt to influence a donor’s future investment decision) or understate it, to reduce the probability that they will be expected to pay if the service is later provided. Even in the absence of strategic motives, quick introspection during a survey can fail to reveal how one will actually behave when real trade-offs must be made.

In part to overcome these limitations, several alternative approaches to eliciting willingness to pay based on actual behavior have been developed in environmental economics. One such revealed preference approach is the travel cost method, in which time costs and other expenditures required to reach a site are used to estimate the willingness to pay for an amenity (Phaneuf and Smith 2003), for instance, to pay for access to recreational sites like state parks.

To our knowledge, the travel cost estimate we develop below is the first such application of the travel cost approach to value improved drinking water quality in a less developed country context. Water source choice in rural less developed country settings has been studied by Whittington, Mu, and Roche (1990) and Mu, Whittington, and Briscoe (1990), however neither of these papers accounts for the role of water quality in the source choice decision (they focus on quantity) and they explicitly rule out the use of multiple drinking water sources, which we find to be empirically important in our data.

3 Rural Water Project (RWP) overview and data

This section describes the intervention, randomization into treatment groups, and data collection.

3.1 Spring protection in western Kenya

Naturally occurring springs are an important source of drinking water in rural western Kenya. The region has land formations that allow the ground water to come to the surface regularly. The area of Kenya in which our study site is located is poor (the unskilled agricultural wage is roughly US$2 per day) and few households have access to improved water services. Like most water resources in rural Kenya, springs are often located on private land but landowners are expected (by both custom and law) to allow public access for the purpose of collecting water.

Springs were selected from a universe of local unprotected springs by a non-governmental development organization (NGO), International Child Support (ICS). The NGO first obtained lists of all local unprotected springs from Government of Kenya Ministry of Water offices. NGO field and technical staff then visited each site to determine which springs were suitable for protection. Springs known to be seasonally dry, in months when the water table is low, were eliminated, as were sites with upstream contaminants in the catchment area (e.g., latrines, graves).

From the remaining list of suitable springs, 200 were randomly selected (using a computer random number generator) to receive protection. These springs are in the Busia and Butere-Mumias districts of western Kenya (Figure 1), and Figure 2 summarizes the timing of the data collection and intervention.

The NGO planned for the water quality improvement intervention to be phased in over four years, due to their financial and administrative constraints. For the purposes of this paper, the springs protected in year 1 (2004/05) are called the treatment springs and those to be protected in later years are called comparison springs. Springs were first stratified on the basis of baseline water quality (this data is described below), distance from tarmac roads, numbers of known users, and geographic region, and then randomly assigned (using a computer random number generator) to determine the order of in which protection would occur. Table 1 presents the baseline summary statistics across the treatment and comparison groups in 2004-5.

Several springs were unexpectedly found to be unsuitable for protection after the baseline data collection and randomization had already occurred, when more detailed technical studies were undertaken. These springs, which are found in both the treatment and comparison groups, were dropped from the sample, leaving 184 springs in the viable sample. Identification of the seasonal springs should not be related to treatment assignment: when the NGO was first informed that some springs were seasonally dry, all 200 sample springs were re-visited to confirm their suitability for protection. Comparisons across the treatment and comparison groups are very similar to those presented in Table 1 if attention is restricted to the sample of 184 springs where protection is viable (results not shown).

A representative sample of households that regularly use each sample spring was also determined at baseline. Survey enumerators visited each spring to interview spring users, asking their names as well as the names and residential locations of other households that use the spring. Enumerators then also elicited information on which households are known to use the spring from a convenience sample of three to four households that lived very near the spring. Households that were listed at least twice among all interviewed subjects were designated as spring users. Seven to eight households per spring were then randomly selected (again using a computer random number generator) from among this spring user list for the household survey. The total number of household spring users varied fairly widely across springs, from eight to 59 with a mean of 31. Over 98% of this spring users sample was later found to actually use the spring during subsequent household surveys, attesting to the validity of the methods used to identify baseline spring users. The spring non-user households were nonetheless retained in the sample throughout the analysis.

Baseline water data was then collected at all 200 sample springs and a survey of local environmental contamination was completed at each spring (January-October 2004), including information on potential sources of contamination (e.g., latrines, graves), vegetation surrounding the spring, slope of the land, and spring maintenance conditions. Water quality in household drinking water storage containers was also tested, as was household survey data on demographic characteristics, health, anthropometrics, and water use choices. The survey is described in further detail below.

To address concerns about seasonal variation in water quality and health outcomes, all springs were randomly assigned (after being first stratified both geographically and by spring treatment group) to an activity "wave," and all data collection and spring protection activities were conducted by wave; there are three waves of springs. The regression analysis uses district-wave fixed effects throughout to control for any seasonal variation in local water quality and disease burden.

The NGO proceeded with community mobilization meetings after baseline data collection and assignment to program groups, and then contracted local masons to carry out spring protection at the treatment springs. The NGO held community meetings during which community permission was obtained for the project, and at which permission was received from the spring landowner to protect the spring (in the two cases where the landowner did not grant such permission, springs were retained in the sample, so results can be interpreted as intention-to-treat estimates). The NGO requested that each community raise a modest initial contribution of 10% of the cost of spring protection, collected mainly in the form of manual labor and construction materials (e.g., sand and bricks). The total cost of spring protection, including these supplies and estimated labor costs, ranges between US$830 and US$1070, depending on the type of construction, which is mainly a function of spring size and soil conditions. The spring was protected after the community raised the initial contribution, and this fundraising was successful at all treatment springs. A committee of spring users responsible for raising the community contribution and for maintaining the spring was also selected by community members attending the initial meeting. Construction quality was monitored by the NGO, and the mason was responsible repairing for any defects during the first three months after protection, after which the protected spring was "handed over" to the community as their property.

A follow-up round of water quality testing at the spring and in homes, spring environment surveys, and household surveys were completed in both treatment and comparison spring communities four to six months after spring protection, in April through August 2005.

3.2 Data collection procedures

The data collection strategy was designed to evaluate the impacts of spring protection on source water quality, home water quality, and child health (diarrhea incidence) and nutrition (anthropometrics). We also collected detailed information on the choice of water sources and other health behaviors.

3.2.1 Water quality data

Water samples were collected from both springs and households in sterile bottles by field staff trained in aseptic sampling techniques.[7] Samples are then packed in coolers with ice and transported to water testing laboratory sites for analysis that same day. The labs use Colilert, a method which provides an easy-to-use, error-resistant test for *E. coli*, an indicator bacteria that is present in fecal matter.[8] Continuous, quantitative measures of fecal contamination are available after 18-24 hours of

---

[7] At springs, the protocol is as follows: a 250 ml bottle's cap is removed aseptically and not touched by hands during the taking of samples. Samples are taken from the middle of standing water and the bottle is dragged through the water so that sample is taken from several locations. About one inch of space is left at the top of the bottle when full. The cap is replaced aseptically. In homes, the protocol is similar. Following informed consent procedures, respondents are asked to bring a sample from their main drinking water storage container (usually a ceramic pot). The water is poured into a sterile 250 ml. bottle using a household's own dipper (often a plastic cup) and resulting estimates of contamination thus reflect conditions in the household's own water storage container and dipper.

[8] The Colilert method has been accepted by the U.S. Environmental Protection Agency (EPA) for both drinking water and waste water analysis. This was one of the first uses of this method in Kenya. Our laboratory standard operating procedures were adapted from the EPA Colilert Quantitray 2000 Standard Operating Procedures. There is currently no consensus microbial indicator for tropical and subtropical climates (where bacteria may live longer in the environment) . However, it is common to use *E. coli* as a means of quantifying microbacteriological water contamination in semi-arid regions like our study site. The bacteria *E. coli* is not itself necessarily a pathogen, but testing for specific pathogens is costly and can be difficult. Dose-response functions for *E. coli* have been estimated

incubation. Quality control procedures used to ensure the validity of the water testing procedures included the use of weekly positive controls, negative controls and duplicate samples (blind to the analyst), as well as monthly inter-laboratory controls.

As we discuss below, there appears to be some significant mean reversion in the spring water quality measurements. This suggests that multiple samples from a given source should ideally be tested to estimate "field sampling variability" and allow for it to be appropriately modeled and accounted for statistically. We do not yet have such data and, to our knowledge, neither do the existing studies of water contamination between the source and home. Without such data, estimated correlations between spring and household water quality using cross-sectional observational data could suffer from some attenuation bias towards zero due to measurement error, leading the analyst to incorrectly conclude that there is more recontamination between water source and the home than there is in reality. The use of an instrumental variable (IV) approach, where source water quality is instrumented with assignment to spring protection, can address this issue as well as omitted variable bias (confounding) more generally. We find below that this econometric approach increases the estimated extent to which improved source water quality translates into home water quality gains.[9]

3.2.2 Household survey data

A household survey was administered to a representative sample of households using sample springs prior to the intervention, and again following the protection of the first group of treatment springs.[10] The target survey respondent was the mother of the youngest child living in the home compound

---

for gastroenteritis following swimming in fresh waters (Kay *et al*. 1994), but such functions may be highly-location specific because the particular pathogens present in fecal matter vary by location and over time.

[9] There are other potential sources of measurement error. First, Colilert generates a "most probable number" of *E. coli* coliform forming units per 100 ml in a given sample, with a known 95% confidence interval. Second, samples that are held for more than six hours prior to incubation may be vulnerable to some bacterial re-growth/death making the tested samples less representative of the original source.

[10] We identified households that were potential spring users by asking people who came to collect water at the springs to tell us the names of people that they thought used the springs. We also asked people living near the springs to provide such a list. If households were mentioned by two sources, we considered them spring users. A random sample of these people were then selected to be in our sample. As we discuss in greater detail, this procedure generated a sample of households that used the springs to varying degrees in practice.

(where the extended family often resides together), or another woman of child-bearing age if the mother of the youngest child was unavailable. The respondent is asked about the health status of all children under age five living in the compound, including recent diarrhea and dysentery (blood in stool) incidence.

The household survey instrument also gathered baseline information about hygiene behaviors and latrine use. Data on the frequency of water boiling, home water chlorination and water collection choices in the past week was collected. Respondents were also asked to give their opinion on ways to prevent diarrhea; they were not given options to choose from, but rather were prompted three times and their responses recorded. This information was then used to construct a "diarrhea prevention knowledge score" at baseline, namely, the number of correct responses provided by the respondent, from the choices: "boil drinking water", "eat clean/protected/washed food", "drink only clean water", "use latrine", "cook food fully", "do not eat spoiled food", "wash hands", "have good hygiene", "medication", "clean dishes/utensils" or "other valid response".[11] Survey respondents on average volunteered around seven such correct preventative activities, with 47% volunteering either boiling water and have/practicing good hygiene.

The definition of diarrhea asked of respondents in the survey is "three or more loose or watery stools in a 24 hour period," which has been used in related studies (see Aziz *et al.* 1990 and Huttly *et al.* 1987). The questionnaire does not attempt to differentiate between acute diarrhea (an episode lasting less than 14 days) and persistent diarrhea (more than 14 days), but differentiates between dysentery and diarrhea by asking whether blood was present in the stool.

Survey enumerators used a board and tape measure to measure the height of children older than two years of age, and digital bathroom-type scales for weight. The height of children under age

---

[11] We reviewed all responses other than those listed here and categorized them as valid or invalid. The major additional correct responses that were not included on the original list were "solar water disinfection", "breastfeeding", and some variant of "use compost pit/keep compound clean".

two was measured as their recumbent length using a pediatric measuring board, and enumerators used a digital infant scale to measure their weight.

We focus below on reported diarrhea in the past week as well as weight and height of children under three years of age as the main health and nutrition outcomes. To address concerns about measurement error in the health data (which may be due to imperfect matching of children across rounds of the household roster, for example), we define a "severe" outlier as an observation more than three times the interquartile range beyond the 25th or 75th percentile in our data for either of two nutritional measures, the body mass index (BMI) at baseline, or the change in BMI across the two household survey rounds. This eliminates 105 child-observations from the sample out of 1931 total observations for children under age three for whom we have data in both household survey rounds. We restrict our attention to children under age three at baseline because of concerns about increasingly misreported age data as children get older.

3.3 Attrition

We successfully followed up 90% of the baseline household sample between the initial and follow-up surveys rounds. Attrition is not significantly related to spring protection assignment: the coefficient estimate on the treatment indicator is only 0.01 (standard error 0.02) in a regression of the attrition indicator on treatment assignment, implying that treatment households are only one percentage point more likely to be lost across survey rounds, and this result is robust to including further explanatory variables as controls (regression not shown).

The baseline characteristics of the households lost over time are typically statistically indistinguishable from those that remain in the sample. Economically better-off households do not appear any more likely to be lost from the sample – iron roofing in the household compound, an asset ownership measure in our survey, is not significantly related to attrition – and the same weak relationship holds between attrition and both baseline household water quality and hygiene knowledge. One partial exception is that households lost from the sample are slightly (5 percentage points, standard error 3 percentage points) less likely to have soap in the home at baseline, but overall, sample attrition bias appears likely to be small.

4 Baseline descriptive statistics

Table 1 presents baseline summary statistics for springs (Panel A), households (Panel B) and children under age three (Panel C). For completeness, we report baseline statistics for all springs and households for which data was collected prior to randomization into treatment groups even if they are later not included in the regression analysis because the spring was later determined unsuitable for protection, though excluding these springs does not change the results.

The water quality measure, *E. coli* MPN CFU/100 ml, takes on values from 1 to 2419[12]. We categorize water samples with *E. coli* CFU/100 ml < 1 as "high quality" water. For reference, the U.S. EPA and WHO standard for clean drinking water is zero *E. coli* CFU/100 ml and the EPA standard for swimming/recreational waters is *E. coli* CFU/100 ml < 100. We call water between these two standards "moderate quality" water. We also create a category of "high or moderate quality" water (with *E. coli* CFU/100 ml < 100) because we rarely observe high quality samples in our data. This is not surprising as the water is neither in a sterile environment nor has residual chlorine as treated drinking water does. We divide the remaining values of *E. coli* CFU/100 ml > 100 into two categories, "poor quality" water (values between 100 and 1000) and "very poor quality" water (greater than 1000).[13]

There is no statistically significant difference between the water quality at treatment versus comparison springs at baseline (Table 1, Panel A), which implies that the randomization created broadly comparable program groups. The spring water in our sample is of moderate quality on

---

[12] In the laboratory test results, the E. *coli* MPN CFU can take values from <1 to >2419. We currently ignore the censoring of the data and treat values of <1 as equal to one and values of >2419 as equal to 2419.

[13] The value of 1000 *E. coli* CFU/100 ml was chosen as a threshold because observational studies suggest that diarrhea incidence can increase rapidly above this level in other less developed country contexts.

average. Only about 5 to 8% of samples from unprotected springs would be close to meeting the stringent EPA drinking water standards, while over a third of samples are poor or very poor quality.[14]

We also collected the holding times before sample incubation in a subsample. We only began collecting holding times several weeks into data collection, after identifying this as a potential concern, given the non-trivial travel times between field sites and the water testing laboratory. In practice, a substantial fraction of water samples were held for longer than six hours, the recommended holding time limit of the U.S. EPA, but we have also confirmed that baseline water quality measures are balanced across treatment and comparison groups when attention is restricted to those water samples that were incubated within six hours of collection, yielding the most reliable estimates (results not shown).

Summary statistics for household water quality are presented next (Table 1, Panel B). Home water is somewhat less likely to be of moderate quality prior to spring protection in the treatment group (and the difference between treatment and comparison group means is significant at 90% confidence), but there is no statistically significant difference in the proportion of samples where water is of high or moderate quality, and we focus on this variable in the analysis below.

At baseline, household water quality tends to be better than spring water quality on average. In the full sample, the average difference in log *E. coli* between spring and household water is 0.57 (s.e. 0.15, n = 1193 households; results not shown). This is likely occurring since many households collect water from sources other than the sample springs and these may be less contaminated in some cases. Only about one half of the household sample gets all their drinking water from the local sample spring and overall respondents make about two thirds of all drinking water collection trips to their sample spring.

---

[14] Previous research in Nigeria shows that unprotected spring water is generally of higher quality than water from ponds or rivers, but that it is vulnerable to spikes in contamination at the transition between rainy and dry seasons. Our data collection stretched over several months both at baseline and at follow-up (Figure 2), and data collection activities were stratified across geographic regions in data collection waves. To account for potential seasonal variation in water quality, we include seasonal (wave) fixed effects in all regression analysis.

Some households report taking additional measures to treat their home water. For instance, about one quarter of households report boiling their drinking water at baseline. We also collected data on chlorination in the follow-up survey: 28% of households reported chlorinating their water at least once in the last six months.[15] Yet the correlations between self reported household water boiling or chlorination with observed household water contamination are very low, raising questions about the accuracy of these self-reports. One potential explanation is that water is sometimes boiled or treated immediately before use (e.g., when making tea), and thus the water samples we tested could overstate contamination at the time of actual consumption. Twenty-seven percent of treatment group households and 21% of comparison households report boiling their water in the previous day, and this difference is significant at 95% confidence. Because of this difference we control for reported baseline water boiling in most regression specifications below. However, because there is no observed correlation between water boiling and observed home water quality in practice, this difference is unlikely to be a major estimation concern. There is no significant difference across treatment and comparison groups in the respondents' "diarrhea prevention knowledge" score.

Household water samples are also held for a shorter length of time than spring water samples, on average.[16] However, this does not explain the observed differences between household and spring water quality: the difference between mean spring and household water quality (measured by ln *E. coli* MPN) is nearly unchanged even when we restrict attention to those water samples held for less than six hours before incubation (the difference in means is 0.56, s.e. 0.08, n = 737).

There are few statistically significant differences in household, respondent and child characteristics across the treatment and comparison groups (Table 1, Panels B and C), further

---

[15] These chlorination levels are almost certainly higher than would usually be observed because the Government of Kenya distributed free chlorine tablets in part of our study region following a 2005 cholera outbreak. In future survey rounds, we will test for residual chlorine in home water for a more reliable measure of chlorine use. Solar disinfection is also occasionally practiced in this area, but we did not collect data on this at baseline.

[16] This is likely because spring water samples are often collected toward the beginning of a field day, while household water samples are collected throughout the day and are more likely to be collected at the end of the day.

evidence that the randomization was successful at creating balanced program groups. Average mother's education attainment is equivalent to slightly less than primary school completion, at about six years (primary school goes through grade 8 in Kenya). One-third of respondents do not have a building with an iron roof in their home compound, where in this area, iron roofing is an indicator of wealth. There are about four children under age 12 residing in each respondent's compound on average. Water and sanitation access is fairly high compared to many other rural settings in less developed countries. About 85% of households report having a latrine, and the average walking distance (one-way) to local water sources is only approximately 10 or 11 minutes.

Finally, we report summary statistics for the subset of children under age three for whom we have both baseline and follow-up survey data (Table 1, Panel C). Children are comparable across treatment and comparison groups in terms of height and weight, our preferred nutritional status measures. About 23% of children in the comparison group had diarrhea in the past 24 hours at baseline, as did 20% in the treatment group. There are similarly no significant differences in other non-diarrheal illnesses (e.g., fever, cough) or in breastfeeding across the groups (not shown).

5 Spring protection impacts on source water quality

5.1 Estimation strategy

Equation 1 illustrates an intention-to-treat (ITT) estimator using spring-level data. Linear regression is employed both when the outcome is continuous – such as the natural log of the *E. coli* MPN measure – and when the dependent variable is an indicator variable (such as for high or moderate quality water, *E. coli* MPN < 100, for example), although results are similar using probit analysis in the latter case (results not shown).

$$W_i^{SP} = \alpha + \beta_1 T_i + X_i^{SP\prime} \beta_2 + \varepsilon_i. \qquad (1)$$

W_i^{SP} is the water quality measure at spring i and X_i^{SP} are baseline spring and community characteristics (e.g., initial level of spring water contamination) and district-wave (season) fixed effects. The variable T_i is a treatment indicator, and ε_i is the standard white noise disturbance term.[17] Randomized assignment implies that the coefficient estimate of β_1 is an unbiased estimate of the reduced-form ITT effect of spring protection.

5.2 Spring water quality results

We report difference-in-differences estimates of the impact of spring protection on source water quality, first for the natural log of *E. coli* MPN (Table 2, Panel A) and then for an indicator of whether water is high or moderate quality (*E. coli* < 100 MPN, Panel B), as the first step in tracing out the impacts of the intervention on water at springs, in homes, and on child health. The average difference in log *E. coli* before and after protection is equivalent to a 74% reduction in contamination at the treatment springs. The difference over time in the comparison group springs is almost exactly zero, thus the difference-in-differences estimate of the impact of spring protection implies that source water contamination declined by 73% as a result of spring protection. Spring protection also increases the probability of high or moderate quality source water by 29 percentage points, from 60% to 89% (Panel B). Figure 3 is a non-parametric (lowess) representation of the data that shows some gains are experienced at nearly all treatment springs, with the spring most contaminated at baseline experiencing the largest average impacts.[18]

These estimated spring protection treatment effects on source water quality are robust to the inclusion of controls for baseline contamination levels and district-wave (season) fixed effects (Table 3). While point estimates suggest that spring protection leads to the greater percentage reduction in water contamination when initial contamination levels are highest, coefficient estimates on the

---

[17] Assignment to treatment may also be used as an instrumental variable for actual treatment (spring protection) status, to estimate an average treatment effect on the treated (TOT) using a two-stage procedure (Angrist, Imbens, and Rubin 1996). In practice, in only 10 springs (of 200) did assignment to treatment differ from actual treatment (because landowners declined to allow the NGO to protect a spring on their land or because the government independently protected springs that were in our comparison group, for example) and thus TOT regressions yield results very similar to the ITT estimates we focus on (not shown).

[18] The change in ln(*E. coli*) is actually positive for the springs with the highest baseline water quality, suggesting there is some mean reversion in spring water quality measures between survey rounds.

interaction terms are not statistically significant (Table 3, regressions 2-3 and 5-6). We also allow for differential treatment effects by baseline household survey respondent hygiene knowledge (the average among users of that spring) and as a function of average local sanitation (latrine) coverage at baseline, as well as by baseline household assets as proxied by iron roof density (regressions 3 and 6), but these interaction terms are not statistically significant.

6 Estimating home water quality impacts when water source choice is possible

We first develop a simple model of water source choice in the presence of travel costs and then derive implications for the estimation of home water quality impacts and for household valuation of water quality gains.

6.1 A travel cost of model of household water source choice

Estimating the impact of spring protection on water quality in the home is complicated by the possibility that households will change their behavior in response to source water quality changes. The two most immediate choices they face are in terms of the choice of a water source, and the choice of whether or not to employ point-of-use technologies (e.g., boiling or chlorination of water). We discuss each of these in turn below, but focus mainly on the choice of water source. The fact that households in our study area often have access to multiple water sources, varying in the quality dimension as well as in terms of distance from the home, allows us to value improvements in water quality using a travel cost approach (Freeman 2003).

Imagine households are located along a line between two water sources, the spring (denoted with letter s) and the alternative source (a), which could be a borehole well, a stream, or another non-sample spring. The round-trip distance (in minutes walking) from the home to the spring for the household is D^s, while the round-trip distance to the alternative source is D^a. The difference in walking times across the sample spring and alternative source is $D \equiv D^s - D^a$, which we call the walking "distance gap" between the two sources. The distance gap can take on positive or negative values, where negative values denote households that live closer to the sample spring than to the alternative source. The distance gap for a household i is denoted D_i. Households are homogeneous along all dimensions except for this distance gap, although we discuss relaxing this below.

In choosing a water source, households trade off the cost (the distance they need to walk to the source) versus the benefits (improved household water quality, which affects the health of household members). The opportunity cost of time – per minute in this case – is denoted $C > 0$. This is a function of the local market wage, and we assume this is constant across all households. Thus the extra cost household i bears to make one additional water trip to the spring (rather than to the alternative source) is C^*D_i, where again this cost can be positive or negative.

The water contamination level (measured as ln (*E. coli* MPN)) for water source $j, j \in \{s, a\}$, is denoted $W_j > 0$, where higher values denote more contamination and thus lower quality. The function relating water quality to household members' health is denoted $V(W_j)$, where $V' < 0$.

There are two time periods to consider, pre-treatment (pre-spring protection) and post-treatment. The water contamination level in the sample spring pre-treatment is denoted W_s and post-treatment is W_s^T (where "T" denotes treatment). Empirically, the experimental spring protection intervention led water contamination levels to fall, $W_s^T < W_s$. We assume the water contamination level in the alternative source, W_a, is constant over time.[19]

The utility of a household from a single water collection trip to source $j \in \{s, a\}$ is $U^j = V(W_j) - CD^j$. Household i chooses the sample spring over the alternative source if the benefits of higher water quality outweigh travel costs, namely when $\{V(W_s) - V(W_a)\} - CD_i \geq 0$.

Consider first the simplest case. In the pre-treatment period, household i chooses spring water if $\{V(W_s) - V(W_a)\} - CD_i \geq 0$, or equivalently if the distance gap is sufficiently small such that $D_i \leq \{V(W_s) - V(W_a)\}/C \equiv D^*$, where D^* can take on positive or negative values. Thus, in this model

---

[19] We do not yet have data on the water quality of alternative local sources, but are currently in the process of collecting this data and will incorporate these data into future analysis.

households with distance gap up to some threshold level use spring water, while those farther away choose the alternative source.

After protection, spring water quality improves relative to the alternative water source, and households choose spring water if $D \leq \{V(W_s^T) - V(W_a)\}/C \equiv D^{**}$, where $D^{**} > D^*$ since spring water is now less contaminated than before ($W_s^T < W_s$). Thus households living at a greater distance from the spring increasingly choose spring water.

Endogenous source choice has implications for the quality of drinking water chosen by households. For households that were spring water users in the pre-treatment period ($D_i \leq D^*$, corresponding to the baseline "sole-source" spring users in our data), their home water quality is unambiguously better after treatment since they still rely exclusively on the spring for drinking water and its quality has improved.

The story is more complicated for households that initially used the alternative source but switched to using the spring after treatment ($D_i \in (D^*, D^{**}]$), the group that corresponds most closely to the multi-source users in our data.[20] For these households, the quality of drinking water in the home could theoretically increase or decrease after treatment. To illustrate, imagine the case in which an improvement in water quality at the spring induces a household to switch from a distant but high quality alternative source (say, a new borehole well) to the closer but relatively lower water quality spring. This could be optimal because households are trading off water quality against time spent walking to collect water. In this case, even if the water quality chosen by the household deteriorates somewhat since they now increasingly use the protected spring, the household is still made better off by spring protection in the sense that household members benefit from time savings. The theoretical prediction on the change in home water quality for these multi-source user

---

[20] For households with an even larger distance gap, $D_i > D^{**}$, there is no change in home water quality since they continue to use the alternative water source just as before, and the alternative source's water contamination level does not change (by assumption). This is not an empirically relevant case for us since even households with large distance gaps rely at least partially on the sample spring for drinking water. This is due to the initial selection of sample households as at least occasional "sample spring users".

households remains ambiguous, in contrast to the sharp theoretical prediction of improved home water quality for the sole-source users who use the sample spring throughout.

It is conceptually straightforward to calculate households' valuation of the water quality improvement caused by spring protection in this simple model, focusing on those households on the margin between using the spring and using the alternative source. After the water quality improvement at the spring ($W_s^T < W_s$), yielding household utility benefits $\{V(W_s^T) - V(W_s)\}$, travel costs must increase by $C(D^{**} - D^{*})$ to restore households to indifference between using the two sources. The greater travel cost households are willing to incur is thus a direct revealed preference measure of the value of improved water quality.

A simple hedonic approach that provides a lower bound on the average household valuation of home water quality improvements is to regress the change in the travel costs incurred to collect source water on the change in the quality of home drinking water (before and after spring protection). This approach exploits both the longitudinal household survey and water quality data and the exogenous variation in water quality generated by the experiment. The coefficient estimate on the change in home water quality provides a lower bound on the average household valuation of cleaner water due to the fact that we do not observe the maximum distance households would be willing to walk for a given improvement in water quality, only the actual distance they actually walked.[21]

Other factors can be added to increase realism and bring the model closer to the data. First, there may be multiple alternative sources in the data, and the water contamination levels of each of these springs and alternative sources vary. Second, households make multiple trips to each spring, and each trip choice is affected by unmodeled factors including the weather, the queue at the water source, or the respondent's mood that day. These factors enter the decision problem through the error term. Incorporating this error term e_j, which can conveniently be assumed to follow a Type I Extreme

---

[21] In future work, when we have additional data on the water contamination of alternative local water sources, we plan to compute a revealed preference valuation of water quality improvements in a mixed logit framework.

Value distribution, the utility of a water collection trip to source *j* is: $U = V(W_j) – CD^j + e_j$, and the spring is chosen by household *i* for that trip if $\{V(W_s) – v(W_a)\} – CD_i + (e_{s,i} – e_{a,i}) \geq 0$. This yields the standard logit set-up.

Households also face the choice of whether or not to adopt a point-of-use water technology. Consider the case of chlorination, for concreteness. There are several dimensions of the cost of adopting chlorination in the home, which we denote $C_P > 0$. These include the purchase price of the chlorine, the time needed to purchase the chlorine and put it in the drinking water container, any psychic costs from learning how to use the product, or costs due to the fact that chlorinated water is not as tasty as untreated water. Offsetting these costs are benefits in the form of reduced water contamination. We model this as a reduction down to contamination level W_P. In this case chlorination and spring protection are substitutes (there are also scenarios under which they could be complements[22]), and thus improvements in water quality due to spring protection would, if anything, reduce point-of-use technology take-up.

The household chooses to use the point-of-use technology when the water quality gains of adoption outweigh the costs. Empirically, as we discuss below, the take-up of point-of-use technologies is very low in our study area, and we do not see substantial shifts in their use after spring protection. This is consistent with the view that the costs – pecuniary or otherwise – of point-of-use technologies are currently relatively large in our study area.

Another extension of the model incorporates a role for hygiene practices and access to sanitation. Influential research argues that water quality improvements alone are insufficient to improve health in the absence of complementary hygiene and sanitation investments that reduce recontamination in storage and transport (Esrey 1996). This can be incorporated into our framework

---

[22] For instance, if chlorination reduced water contamination by some fixed amount ΔW regardless of the starting contamination level, and the health benefits function $V(W)$ were convex and decreasing, then improved source water quality (a reduction in W) and point-of-use technologies could be complements and spring protection could actually boost demand for point-of-use chlorination technologies.

by making water quality from source j, W_j, a function of both protection ("treatment", $T_j \in \{0,1\}$) as well as the local hygiene and sanitation environment, denoted H_i, where improved hygiene and sanitation is associated with an increase in H_i. Imagine that this variable is fixed for household i (in a richer model investments in hygiene knowledge and sanitation could be endogenized along the lines of the point-of-use technology adoption discussed above). H_i can be thought of as the recontamination level of water from source j to the home.

The level of water recontamination in the absence of spring protection, in a setting with minimal hygiene and sanitation, is denoted W_j^*. Formally, let $W_j = W_j^* - \phi(T_j, H_i)$, where $\phi_1 > 0$ (spring protection reduces water contamination at the source) and $\phi_2 > 0$ (better hygiene and sanitation in household i reduces recontamination during transport and storage). The sign of the cross-partial derivate, ϕ_{12}, determines whether spring protection and hygiene/sanitation are substitutes or complements in reducing water contamination, and this is an empirical question.[23] Along these lines, in the analysis below we estimate the interaction effect of spring protection with measures of household hygiene knowledge and sanitation access.

6.2 Estimating spring protection impacts on water source choice and behavior

We estimate an equation analogous to equation 1 but using household level data to estimate the impact of spring protection on household behaviors – including water source choice, self-reported water boiling, self-reported water chlorination, diarrhea prevention knowledge, and number of trips made to collect water in the past week, a measure of water quantity used – as well as impacts on home water quality. Once again, econometric identification relies on the randomized program design.

We consider the theoretical predictions derived above by splitting the data into two subsamples, the baseline sole-source users (those households who only used the sample spring at

---

[23] Note that spring protection and hygiene/sanitation could still be substitutes or complements even if the ϕ function is linearly separable if V is convex.

baseline) and baseline multi-source users (those households who used at least one other source for water at baseline). The predictions are, first, that use of the protected spring should increase more among initially multi-source user households than among sole-source user households, and second, that home water quality improvements among sole-source user households should be at least as large as gains observed for multi-source user households.

We control for baseline household characteristics in some specifications including household sanitation access, the respondent's diarrhea prevention knowledge score, an indicator for whether a household has an iron roof (a proxy for wealth), the respondent's years of education, and the number of children under age 12 in the compound at baseline, in addition to the district-wave (season) fixed effects. Regression error terms are clustered at the spring level in the household regressions. Randomization of the intervention at the level of the spring community means that the households using the same spring at baseline are not independent units of study, and outcomes among these households may be correlated. Not only do these households share a common water source, but they may have kinship ties and share the same latrines and alternative water sources. This reduces the power of statistical tests relative to what would be possible if a source water quality intervention were randomized at the household level.

We also test the hypothesis that source water quality improvements are more valuable in the presence of improved household sanitation access and/better or hygiene knowledge (as argued by Esrey 1996) by interacting the spring protection treatment indicator variable with these variables. We also allow for differential treatment effects by self-reported water boiling at baseline, the leading point-of-use water treatment strategy in our study area. Households that boil their home water could reduce contamination levels, weakening the link between source and home water quality.

Finally, we also estimate the extent to which improvements in source water quality translate into to improved household water quality, where the equation of interest is:

$$W_{ijt}^{HH} = a + b_1\, W_{it}^{SP} + X_{ij}^{HH\,\prime} b_2 + v_i + e_{ijt}. \qquad (2)$$

The dependent variable is water quality (measured in units of ln (*E. coli* MPN)) in household *j* at spring *i* in time period *t*, and the independent variables are the analogous spring water quality measure and the vector of baseline household characteristics described above. The spring-level random effect is captured by *v*, and *e* is a standard white noise error term. Random assignment of springs to protection implies that we can avoid both omitted variable bias (confounding) and also reduce attenuation bias due to measurement error by estimating b_1 in an instrumental variables framework. In particular, assignment to spring protection treatment multiplied by an indicator variable for the "After treatment" time period is the instrument for spring water quality. The first-stage regression equation is nearly identical to equation 1 above, but data from both time periods (both *t*="Before treatment" and *t*="After treatment") are utilized, rather than simply the "After treatment" data as in equation 1. Both the treatment indicator variable and the "After treatment" period indicator variable are included as explanatory variables in both the first and second stage regressions in this case. This IV approach provides an analytically attractive means of estimating the degree of water contamination between source and home, especially among the sole source user households who almost exclusively use the sample spring for drinking water in both periods.

6.3 Household water choice and home water quality results

We first consider impacts on water collection and source choice (e.g., the number of water collection trips from each of the household's water sources), water transportation and storage behaviors (e.g., reported water boiling and water chlorination), and complementary sanitation and hygiene behaviors (e.g., diarrhea prevention knowledge score at follow-up). We report results for the full sample of households, and for sole-source users and multi-source users separately, given the interesting theoretical distinctions across these two groups of households.

The main behavioral change that resulted from spring protection is an increase in the use of the protected springs for drinking water, while other behavioral changes appear to be minor. Assignment to spring protection treatment is strongly positively correlated with use of the sample spring: treated households are 22 percentage points more likely to use their sample spring as a source of drinking water (92% overall, compared to 70% of households in the comparison group in the follow-up survey, Table 4, Panel A). As predicted by the water source choice model, effects are much more pronounced among the sample of multi-source users: in the treatment group they increase use of the spring by 36 percentage points, while the gain among the sole-users was very small (at only 7 percentage points), perhaps not surprising given that they only used the spring at baseline. There are similarly large impacts on the fraction of water collection trips made to the sample spring after protection. Underlying this increase in use of protected springs were increasingly positive perceptions about the quality of drinking water from protected springs: respondents at treated springs were 36 percentage points more likely to believe the water is very clean at the source during the rainy season, and these effects are similar for both sole-source and multi-source user households.[24]

There were no statistically significant effects of spring protection on the average distance households walked to their main drinking water source (the average length was about 10-11 minutes one-way or 20-22 minutes round-trip), nor on the number of trips made to water sources in the past week, and although there are some differences across sole-source and multi-source user households in the number of trips made, the differences are minor. Similarly, there are no significant changes in any water transportation and storage behaviors (Table 4, Panel B). Households at treated springs are not differentially likely to chlorinate or boil water, nor does the average age of water collectors change significantly. There is also no evidence of changes in self-reported diarrhea prevention knowledge nor in other hygiene measures (Panel C).

---

[24] It is worth noting that sole-source and multi-source user households are not statistically significantly different at baseline along the household or child health characteristics presented in Table 1 (results not shown), suggesting that they are broadly comparably along most relevant dimensions.

When we also examined the treatment effect on the fraction of household water collection trips made to the assigned spring, we find differential effects as a function of household walking distance to the spring: while multi-source users are still more likely overall to use more distant springs after protection, they are somewhat less likely to switch to the assigned spring when it is located farther away from their home, as predicted by the theory (results not reported).

We next turn to estimating the effect of spring protection on water quality in the home, reporting difference-in-differences estimates in a manner analogous to the spring-level analysis. The average impact of spring protection on home water quality is far smaller than the impacts at source: for the full sample of households, the average reduction in water contamination is only 28% (Table 5, Panel A), only about one third of the 73% reduction at the spring level (Table 2, Panel A). Estimated impacts on having water of high/moderate quality are similarly weak (Table 5, Panel B). In a regression in which water quality measures from both springs and homes is pooled, the hypothesis of equal spring protection treatment effects at the source and in homes is rejected at 99% confidence (results not reported).

As discussed above, one theoretically plausible explanation for limited observed home water quality gains is the possibility of endogenous sorting of households among water sources springs in reaction to spring protection, which would dampen observed gains at home if some households switch to using closer but lower quality sources. The data bears out these model predictions. Among the sole-source spring users, spring protection impacts on home water quality are substantial, including a 52% reduction in average home water contamination and a 13 percentage point increase likelihood of having high/moderate quality water at home (Table 5, Panels C and D), while for multi-source users home water gains are essentially zero (Table 5, Panels E and F). This result for sole-source users is our first empirical indication that recontamination of water in transport and storage is not a major factor reducing home water quality in our setting, since if that were the case, we would expect large drops in water quality between source and home even for the sole-source spring users who use the same water source (the sample spring) throughout.

Broadly similar results obtain for the full sample, and for the sole-source and multi-source users, when baseline household characteristics are included as explanatory variables in a regression framework, although the difference between sole-source and multi-source households is smaller in this case (Table 6). Once again, the overall effect of spring protection on home water quality is moderate (regressions 1-2), with apparently larger effects for the sole-source households (regressions 3-4) than the multi-source users (regressions 5-6), though we cannot reject equal treatment effects for sole source and multi-source users here (results not shown).

The payoff from the multiple regression framework lies in allowing us to estimate differential treatment effects across households with different baseline characteristics. The interaction between spring protection and the initial spring contamination level is large, negative and statistically significant for the sole-source user sample (Table 6, regression 4), implying that home water quality gains were largest when the spring was initially most contaminated. However, we find no evidence of differential treatment effects as a function of household sanitation, water boiling, diarrhea prevention knowledge, or distance to the water source. Households living in communities with greater latrine coverage do appear to have less contaminated water, but this is does not differentially affect the impact of the spring protection treatment. The fact that there are no robust differential treatment effects as a function of any pre-existing sanitation access or hygiene knowledge runs counter to claims common in the literature that source water quality improvements are only valuable when these complementary factors are also in place.

To more fully assess the extent of water recontamination in transit and storage, we next examine the relationship between spring water quality and home water quality. In the simplest linear regression of home water quality (in ln(*E. coli* MPN)) on spring water quality (in the same units), in a specification that effectively ignores the experimental project design, we estimate an elasticity of only 0.190 (Table 7, regression 1). With only these results in hand, a naïve conclusion would be that water recontamination in transport and storage prevents over 80% of source water quality improvements from reaching the home, and thus that source water quality improvements like spring protection are largely ineffective at improving home water quality. Yet an instrumental variable approach that exploits the experimental variation in source water quality and also addresses possible attenuation bias due to water quality measurement error begins to tell a different story: the estimate rises by sixty percent up to 0.332 (regression 2) in that case.

However, the interpretation of this coefficient estimate is still complicated by endogenous household sorting among water sources, as discussed above. In order to eliminate any such bias from the analysis to the greatest extent possible, we next focus on the sole-source spring user households, since they almost exclusively use the sample spring for drinking water both before and after spring protection. Focusing on this subsample again in an instrumental variable framework generates a much larger elasticity estimate of 0.707 (highly statistically significant at over 95% confidence, Table 7 regression 3). In other words, in this sample where endogenous water source choice is effectively eliminated, over 70% of the source water quality gains at the source generated by spring protection translate into home water quality gains, strong evidence that recontamination does not render source water improvements useless in this setting. In contrast, the analogous elasticity for multi-source users is only 0.075 (and is not statistically significant, regression 4), which we argue is likely to mainly reflect changes in their water source choice rather than only recontamination in transport and storage, since the latter would also presumably affect the sole-source users, but there is no evidence that is the case.

We conclude that the impact of spring protection on household water quality are large and statistically significant for those households that mainly use the same water source throughout (the sole-source user households). The richness of our longitudinal household survey and water quality data, together with the experimental program design that generated exogenous variation in source water quality, allow us to reach very different conclusions than would be suggested by existing analyses using observational cross-sectional data.

6.4 A revealed preference estimate of household valuation of cleaner water

We next use the data on household water source choices to estimate a revealed preference measure of household valuation of the water quality gains generated by spring protection. As discussed above, we focus on a simple specification that regresses the change in the price paid for water (in minutes walked) before versus after spring protection, on the change in observed home water quality over this same period. This yields a lower bound on average household valuation since we do not observe the maximum distance households would be willing to walk for the improved water. This is related to the hedonic approach in Chay and Greenstone (2005).

The variable that captures the distance household members walked for water is constructed as the (Fraction of drinking water collection trips to the sample spring) x (Round-trip walking distance to the sample spring) + (Fraction of drinking water collection trips to the alternative water source) x (Minimum round-trip walking distance to an alternative water source), based on household survey self-reports. To be conservative, when households report more than one alternative source we define the walking distance as the minimum distance among all alternative sources named. We focus on multi-source user households in the analysis rather than the sole-source users, since the change in this distance walked variable is equal to zero for nearly all sole-source users (since they are inframarginal, they do not change their water source choice patterns in response to spring protection). Restricting attention to only multi-source users with complete source water choice and home water quality data from both survey rounds leaves a sample of 256 households.

We regress the change in this walking distance – our measure of price paid – across survey rounds on the change in measured home water quality across rounds, conditional on month of survey controls (to capture any seasonal variation in water quality). The point estimate on the change in water quality in this specification is -0.217 (standard error 0.128, significantly different than zero at 90% confidence).[25] This implies that the average additional distance households are willing to walk (round-trip) for a 1.0 log reduction in home water contamination is (-1) x (2) x (-0.217) = 0.434 minutes per trip. For a reduction of 0.74 log points in home water contamination – the average home water gain among sole-source user households after spring protection (Table 5) – this implies a lower bound average willingness to walk an additional (0.434) x (0.74) = 0.322 minutes per water collection trip.

The final step is converting this relationship into a monetary valuation by putting a value on water collectors' time. In general, this could vary across households and across members of the same household, but since the majority of water trips in our data are made by adult women (fewer than 10 percent of trips are made by children under age 12), the unskilled adult agricultural wage is a natural benchmark, especially since nearly all of our households engage in subsistence agriculture. The daily local agricultural wage is approximately 150 Kenyan Shillings, or almost exactly US$2 at 2005 exchange rates. Assuming this is for eight hours of work per day, this is equivalent to US$0.0042 per minute of time working.

We compute the average value of improved water quality due to spring protection over the course of a year, where households make an average of 47 water collection trips per week. The lower bound on this valuation is (0.322 minutes walked per water collection trip) x (US$0.0042 per minute walked) x (47 trips/week x 52 weeks/year) = US$3.27 per household per year.

One important caveat is that the elasticity we estimate may only be valid over the range of water quality improvements that we observe in our data, and there may be substantial nonlinearities in valuation as water quality improves down to very low levels of contamination. This concern means that it may be less appropriate to use these numbers to inform policy about point-of-use treatment

---

[25] Measurement error in the water quality data is a possible concern (based on the results in table 7). We used simulation techniques to investigate whether the standard error on the Most Probable Number of CFUs as reported by our test is likely to generate considerable attenuation bias and we conclude that it does not (not shown).

(which when properly used results in water with zero *E. coli*) than for source water quality questions which do not generate such low levels of contamination.

We next compare these willingness to pay figures to the actual cost of spring protection to understand whether the estimated valuation would cover costs in the event that user fees could be collected. There are on average 31 households in each baseline spring user list. Assuming that each has the same valuation for the water quality improvement generated by spring protection, and that a spring lasts for at least 10 years before requiring further masonry work, then this yields a lower bound estimate of total community benefits of (US$3.27) x (31 spring users) x (10 years) = US$1013 per spring, ignoring discounting for simplicity. This is similar to the cost of initial construction for springs in this area, which ranges from US$830 and US$1070. Since our valuation estimates are lower bounds, we are unable to compute a return to a spring protection investment, but it appears unlikely the returns are negative. Separately, the observed levels of household valuation are also likely to exceed the annual costs of maintaining a spring. These costs – mostly labor for keeping storm water and drainage ditches clean and also fencing materials to protect the catchement area – are approximately $30 per year.

7 Child health and nutrition impacts

We estimate the impact of spring protection on child health outcomes using child-level data (usually reported by the child's mother) as well as anthropometric data collected by survey enumerators in the household survey in a difference-in-differences framework (equation 3):

$$Y_{ijt} = \alpha + \beta_1 T_i + \beta_2 After_t + \beta_3 After_t * T_i + X_{ij}'\beta_4 + u_i + \varepsilon_{it} \qquad (3)$$

where the dependent variables we focus on are diarrhea in the past week, child weight and child height. The coefficient estimate that captures the treatment effect is β_3, the coefficient on the *After * Treatment* term. We also include controls for child gender and age and district-wave (season fixed effects). When the dependent variable is discrete (e.g., diarrhea), we use probit estimation and present marginal effects at mean values. We consider separately treatment effects for children living in sole-source user households versus multi-source user households, given the different home water quality impacts across these two groups. The sample size falls considerably (by roughly one third) relative to the baseline sample since not all children were physically present to be measured on the day of the follow-up survey.

Despite the home household water quality gains that we estimate, particularly in the sole-source user households, there are no statistically significant estimated treatment impacts on any of the child health or nutrition measures (Table 8). The signs on the coefficient estimates on the key "After treatment * Treatment" terms provide suggestive evidence that diarrhea may have fallen and weight improved slightly, but these effects are not statistically significant at traditional confidence levels. The weak results hold even restricting attention to the sole-source users (in regressions 2, 4 and 6, the coefficient estimate on the "After treatment * Treatment" terms can be interpreted as effects for sole-source users).

The lack of a statistically significant impact on reported diarrhea may not be surprising, in light of the fact that we measured this outcome only at two moments in time and it is notoriously difficult to measure accurately. Yet with several alternative outcome measures (e.g., diarrhea plus fever or diarrhea in the past 24 hours), or alternative specification (e.g., with or without age controls, including children up to age five) we also do not find any indication of statistically significant treatment effects. Consolidated health gains are perhaps more likely to be captured in the anthropometric measures, but may not yet be apparent due to the short time period between spring protection and the follow-up survey, four to six months.

Statistical precision is a concern, given the limited sample of infants and young children (approximately 900), however we do have sufficient statistical power to reject large spring protection treatment impacts. The estimates in Table 8 allow us to reject reductions in diarrhea in the past week of greater than about 50% among sole source users. Among this same group, households who experienced quite large home water quality improvements, we can reject weight gains of greater than 0.49 kilograms (about four percent of baseline average weight) with 95% confidence, and height gains of only 0.55 centimeters (about one percent of baseline average height) with 95% confidence.

We will be collecting additional rounds of data in the future to improve the statistical precision of this child health analysis, and thus regard these estimated health impacts as preliminary.

8 Discussion and conclusion

We study spring protection, an intervention that dramatically and quite cheaply improved source water quality in a rural African setting, reducing contamination by 73% on average. We find that although, for a sample of sole source spring users, 71% of these source water quality gains appear to have been translated into improvements in home water quality, such water quality gains were not sufficiently large to generate substantial child health and nutrition improvements.

While measurement error and low statistical precision may explain the lack of statistically significant treatment effects for child diarrhea, these results are also consistent with what might be found if the primary causes of diarrhea were water washed (arising because of insufficient water for washing and bathing) rather than waterborne (transmitted via ingestion of contaminated water). Certainly, spring protection could only be expected to address waterborne illness since we find no change in water quantity collected (the number of trips made to collect water, Table 4). The maximum reduction in diarrhea prevalence that could be expected from addressing waterborne illness in this region is only about 25%, where this figure comes from a a cluster randomized control trial of POU water treatment that resulted in 82% of treatment households having *E. coli* MPN <1, a far greater water quality improvement than we observe as a result of spring protection (Crump et al. 2005). Thus, spring protection should have far smaller impacts on diarrhea incidence than this. Planned future data collection will give us sufficient statistical power to detect such effects.

Another possible interpretation common in the existing water literature is that source water quality improvements only translate into home water quality gains – and eventually child health gains – when there are good household hygiene practices and adequate local sanitation already in place. However, this alone does not appear to be sufficient to explain our result: we do not find any evidence that spring protection led to larger home water quality gains when hygiene knowledge or latrine coverage were better. Also, spring protection did not lead to any detectable changes in water collection, transport, or storage practices, or to changes in any other preventive health behaviors that we measured, although there were sharp changes in water source choices among some households. The fact that the water quality gains caused by spring protection did not largely dissipate during transport and storage for the sole source user households belies the conventional wisdom that recontamination renders source water quality investments alone ineffective.

We also estimate willingness to pay for improved source water by analyzing how households change their choice of water source – and in particular, the distance they are willing to walk for water – in response to the improvements generated by spring protection. These suggest that households are on average willing to pay at least US$3.27 per year for the water quality improvements generated by spring protection, or about one and a half days' wages. This figure could be used to provide guidance on the magnitude of feasible user-fees at communal water sources. This modest willingness to pay is also consistent with the small child health benefits we estimate

These findings are the first set of results from a larger research project by the authors whose goal is to shed light on how to best and most cost effectively provide safe drinking water in rural Africa. Future rounds of household data collection, as well as the protection of additional springs, will allow us to more precisely estimate the child health and nutrition impacts of spring protection. Beyond spring protection, we plan to use additional randomized evaluations to investigate the role that improvements in point-of-use water technologies – including chlorination, filters, and solar disinfections within the home – as well as water quantity increases can play in achieving safe drinking water, and in particular to determine whether these approaches would be most effectively employed as complements to or substitutes for source water investments like spring protection.

References

Angrist, J. *et al.* (1996), "Identification of causal effects using instrumental variables," Journal of the American Statistical Association 91(434), 444-472.

Aziz, K., B. Hoque, K. Hasan, M. Patwary, S. Huttly, M. Arman, and R. Feachem (1990), "Reduction in diarrhoeal diseases in children in rural Bangladesh by environmental and behavioural modifications," Transactions of the Royal Society of Tropical Medicine and Hygiene 84(3), 433-438.

Beaton, G.H. *et al.* (1993), "Effectiveness of vitamin A supplementation in the control of young child mortality in developing countries," Nutrition Policy Discussion Paper number 13 ACC/SCN: Geneva.

Black, R.E. (1998), "Therapeutic and preventive effects of zinc on serious childhood infections diseases in developing countries," American Journal of Clinical Nutrition 68(supplement), 476s-9s.

Blum, D. and R.G. Feacham (1983), "Measuring the impact of water supply and sanitation investments on diarrheal diseases: Problems of methodology," International Journal of Epidemiology 12(3), 357-65.

Chay, Kenneth, and Michael Greenstone. (2005). "Does Air Quality Matter? Evidence from the Housing Market", Journal of Political Economy, 113(2), 376-424.

Crump JA, Otieno PO, Slutsker L *et al.* (2005) "Household based treatment of drinking water with flocculant-disinfectant for preventing diarrhoea in areas with turbid source water in rural western Kenya: cluster randomised controlled trial," British Medical Journal, 331, 478–483.

Diamond, P.A and J. A. Hausman, (1994), "Contingent valuation: Is some number better than no number?" Journal of Economic Perspectives, 8(4),45-64

Esrey S.A. (1996), "Waste, water and well-being: A multicountry study," American Journal of Epidemiology 143(6), 608-22.

Esrey S.A. *et al.* (1991), "Effects of improved water supply and sanitation on ascariasis, diarrhea, dracunculiasis, hookworm infection, schistosomiasis and trachoma," Bulletin of the World Health Organization 60, 609-621.

Esrey S. A. *et al.* (1985), "Interventions for the control of diarrheal diseases among young children: Improving water supplies and excreta disposal facilities," Bulletin of the World Health Organization 63(4), 757-72.

Esrey S. A. and J.-P. Habicht (1986), "Epidemiologic evidence for health benefits from improved water and sanitation in developing countries," Epidemiology Review 8, 117-128.

Feacham, R.A. (1977), Water supplies for low-income communities: Resource allocation, planning and design for a crisis situation, in R. Feacham, M. McGarry, and D. Mara eds., “Waste, water, and health in hot climates,” John Wiley and Sons: London.

Fewtrell, L. *et al.* (2005), “Water, sanitation, and hygiene interventions to reduce diarrhoea in less developed countries: A systematic review and meta-analysis,” Lancet Infectious Diseases 5, 42-52.

Freeman, A. Myrick III. 2003. The Measurement of Environmental and Resource Values: Theory and Methods (Second Edition). Washington DC: Resources for the Future.

Glass, R. *et al.* (2004), Rotavirus vaccines, in C.A. de Quadros, ed., “Vaccines: Preventing Disease & Protecting Health,” Pan American Health Organization: Washington, D.C.

Grotto, I. *et al.* (2003), “Vitamin A supplementation and childhood morbidity from diarrhea and respiratory infections: A meta-analysis,” Journal of Pediatrics 142, 297-304.

Hill, Z. *et al.* (2004), Family and community practices that promote child survival, growth, and development: A review of the evidence, WHO: Geneva.

Huttly S.R.A. *et al.* (1987), “The epidemiology of acute diarrhea in a rural community in Imo State, Nigeria,” Transactions of the Royal Society of Tropical Medicine and Hygiene 81, 865-70.

Iyer, Param, Jennifer Davis, Elif Yavuz and Barbara Evans (2006), “Rural water supply, sanitation, and hygiene: A review of 25 years of World Bank lending (1978–2003),” World Bank Water Supply & Sanitation Working Notes, Note No. 10, July 2006.

Kosek, M. *et al.* (2003), “The global burden of diarrheal disease, as estimated from studies published between 1992 and 2000,” Bulletin of the World Health Organization 81, 197-204.

Lenehan, A. and J. Martin (1997), “Spring protection in southern KwaZulu Natal,” Paper presented at the twenty-third Water, Engineering, and Development Centre (WEDC) Conference on Affordable Water Supply and Sanitation, Durban.

Miller, P. and N. Hirschhorn (1995), “The effect of a national control of diarrheal diseases program on mortality: The case of Egypt,” Social Science Medicine 40(10), 1-30.

Mintz, Eric, Jamie Bartram, Peter Lochery and Martin Wegelin, (2001), “Not just a drop in the bucket: Expanding access to point-of-use water treatment systems,” American Journal of Public Health, October; 91(10): 1565–1570.

Mwami, J. (1995), “Spring protection- sustainable water supply,” Paper presented at the twenty-first Water, Engineering, and Development Centre (WEDC) Conference on Affordable Water Supply and Sanitation, Kampala.

Mu, X., D., Whittington, and J. Briscoe, (1990), “Modelling village water demand behavior: A discrete choice approach,” Water Resources Research 26(4), 521-529.

Pinfold J.V. (1990), “Faecal contamination of water and fingertip prints as a method for evaluating the effect of low-cost water supply and sanitation activities on faeco-oral disease transmission. A case study in rural north-east Thailand,” Epidemiology and Infection 105, 363–375.

Perera, B.J.C. *et al.* (1999), “The impact of breastfeeding practices on respiratory and diarrhoeal disease in infancy: A study from Sri Lanka,” Journal of Tropical Pediatrics 45, 115-8.

Phaneuf, D. and V..K. Smith (2003), Recreation demand models, in “Handbook of Environmental Economics,” K. G. Mäler and J. R. Vincent eds., Volume 2, pages 671-761 Elesevier: Amsterdam

Raisler, J. *et al.* (1999), “Breastfeeding and infant illness: A dose-response relationship?” American Journal of Public Health 89, 25-30.

Ramakrishnan, U. and R. Matorell (1998), “The role of vitamin A in reducing child mortality and morbidity and improving growth,” Salud Publica de Mexico 40(2), 189-198.

Rosen, S. and J.R. Vincent (1999), “Household water resources and rural productivity in Sub-Saharan Africa: A review of the evidence,” Harvard Institute for International Development Discussion Paper # 673.

Unicef (2006), “Meeting the MDG Drinking Water and Sanitation Target: A Mid-Term Assessment of Progress,” Available at: http://www.unicef.org/wes/mdgreport/index.php

UNEP (United Nations Environment Program) (1998), Sourcebook of alternative technologies for freshwater augmentation in Africa, Osaka: UNEP.

USAID (United States Agency for International Development) (1996), Environment, health and people, An update on USAID’s Environmental Health Project, USAID Environmental Health Project: Arlington.

Varley, R.C.G. *et al.* (1998), “A reassessment of the cost-effectiveness of water and sanitation interventions in programmes for controlling childhood diarrhea,” Bulletin of the World Health Organization 76(6), 617-631.

Vaz, L., and P. Jha (2001), “Note on the health impact of water and sanitation services,” World Health Organization Commission on Macroeconomics and Health Working Paper Series Paper #WG5: 23.

Victora, C.G. *et al.* (1996), “Falling diarrhoea rates in Northeastern Brazil: Did ORT pay a role?” Health Policy and Planning 11(2), 132-41.

Victora, C.G. *et al.* (2000), “Reducing deaths from diarrhea through oral rehydration therapy,” Bulletin of the World Health Organization 78(10), 1246-56.

UN-Water/Africa (2006), African Water Development Report, Economic Commission for Africa: Addis Ababa, Ethiopia.

Watson, T. (forthcoming), “Public health investments and the infant mortality gap: Evidence from federal sanitation interventions and hospitals on U.S. Indian Reservations,” Journal of Public Economics.

WHO (World Health Organization) (2000), Global water supply and sanitation assessment, WHO: Geneva.

WHO (World Health Organization) (2002a), Managing water in the home: Accelerated health gains from improved water supply, WHO: Geneva.

WHO (World Health Organization) Collaborative Study Team on the Role of Breastfeeding on the Prevention of Infant mortality (2000), “Effect of breastfeeding on infant and child mortality due to infectious diseases in less developed countries: A pooled analysis,” Lancet 355, 451-5.

Whittington, D., X. Mu, and R. Roche (1990) “Calculating the value of time spent collecting water: Some estimates for Ukunda, Kenya,” World Development 18(2), 269-280.

World Bank (1993), World development report: Investing in health, Oxford University Press: New York.

World Bank (2002), Water, sanitation, and hygiene at a glance, World Bank: Washington, DC.

World Bank (2003), Water - A priority for responsible growth and poverty reduction: An agenda for investment and policy change, World Bank: Washington, DC.

World Bank Water Demand Research Team (1993), “The demand for water in rural areas: Determinants and Policy Implications,” World Bank Research Observer 8(1), 47-70.

Wright, J., S. Gundry, and R. Conroy (2004), “Household drinking water in developing countries: A systematic review of microbiological contamination between source and point-of-use,” Tropical Medicine and International Health 9(1), 106-17.

ZICG (Zinc Investigators’ Collaborative Group) (1999), “Prevention of diarrhea and pneumonia by zinc supplementation in children in developing countries: Pooled analysis of randomized control trials,” Journal of Pediatrics 135(6).

ZICG (Zinc Investigators’ Collaborative Group) (2000), “Therapeutic effects of oral zinc in acute and persistent diarrhea in children in developing countries: Pooled analysis of randomized control trials,” American Journal of Clinical Nutrition 72:1516-22.

Figure 1: Map of study region

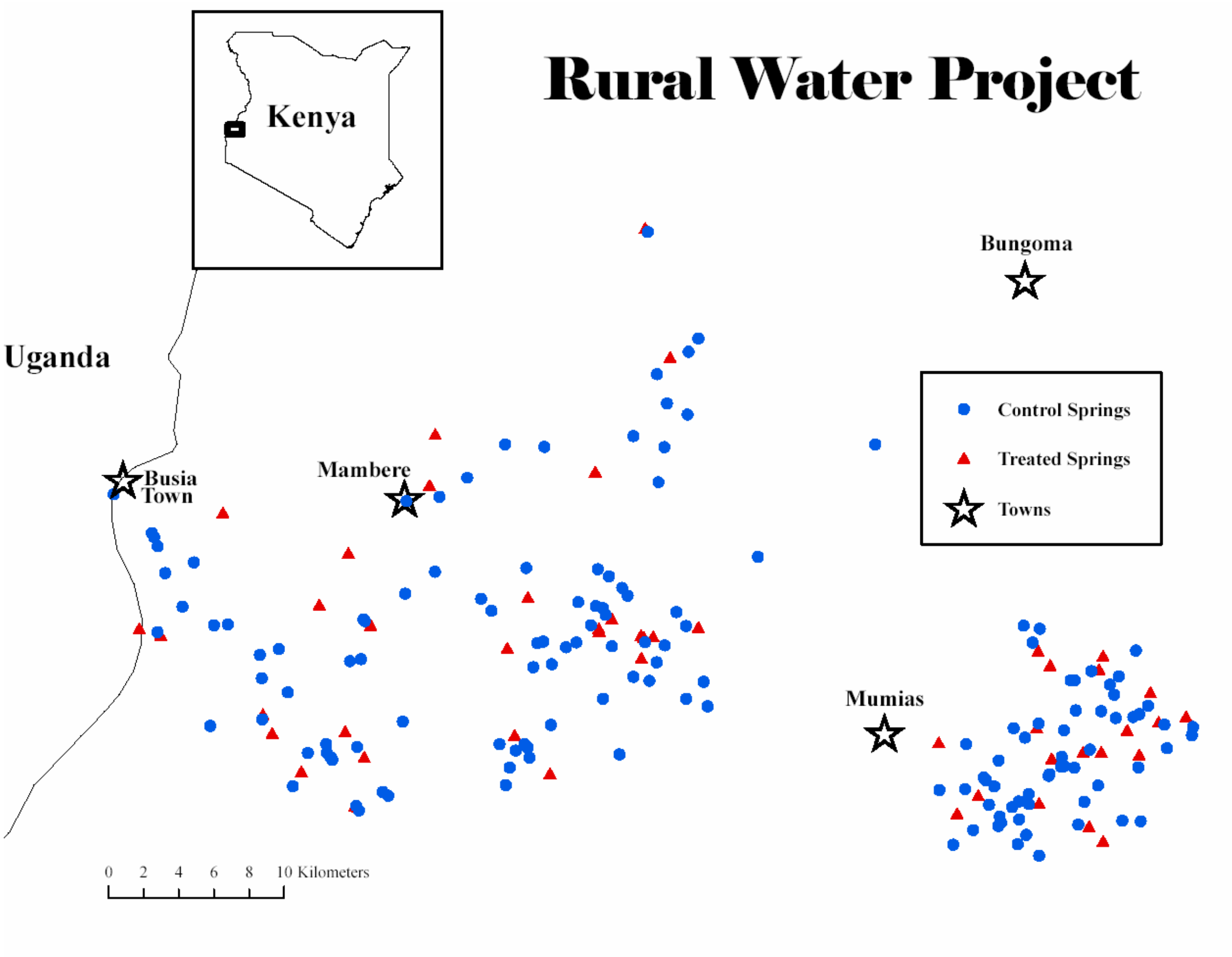


Figure 2: Timeline of Rural Water Project (RWP) Activities 2004-2005

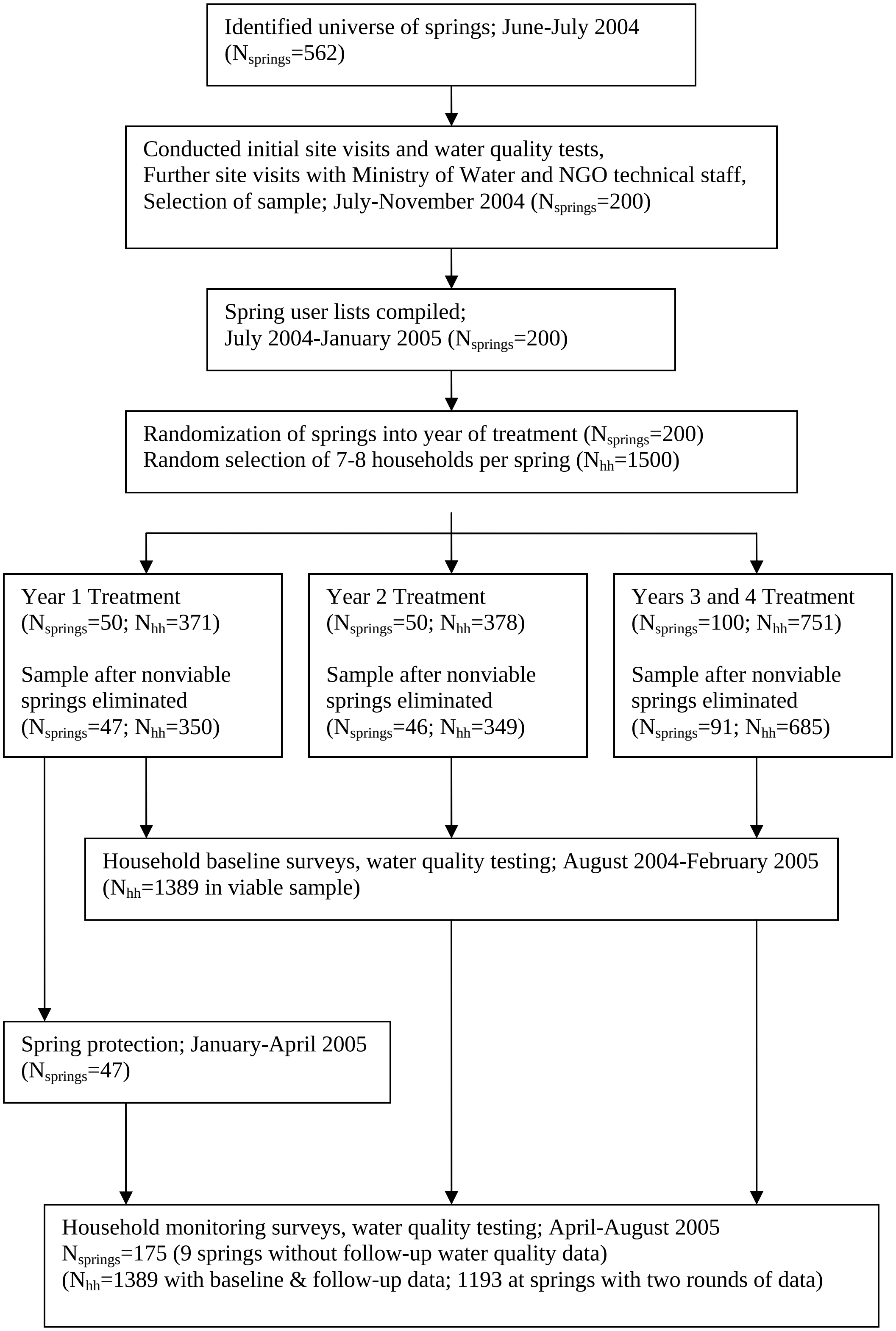
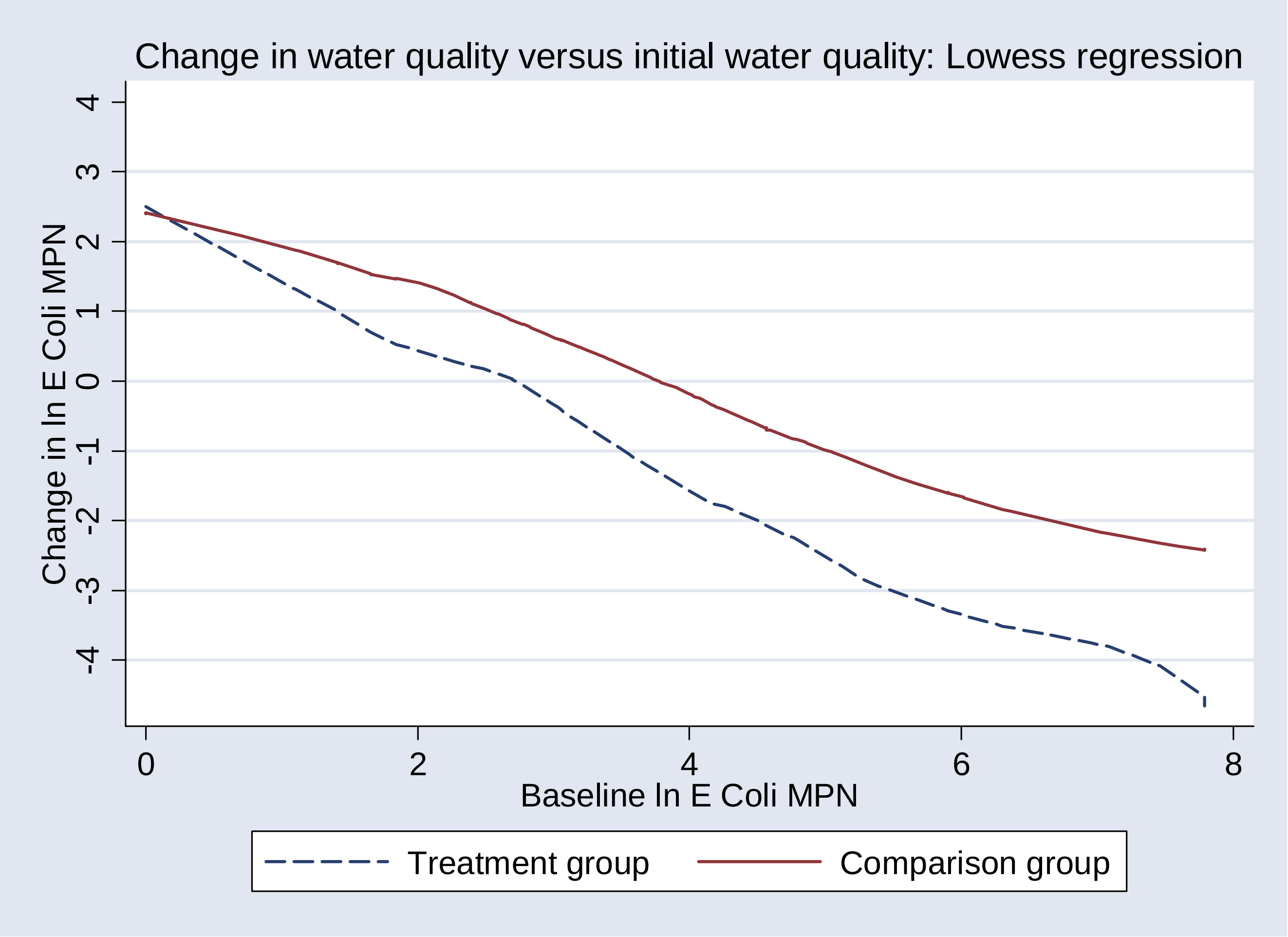


Figure 3: Change in water contamination versus baseline water contamination



Notes: To 10-90 range in Baseline ln (*E Coli* MPN) is [1.13, 6.31].
MPN stands for “most probable number” coliform forming units (CFU) per 100ml.

Table 1: Baseline descriptive statistics

	Treatment		Comparison		Treatment – Comparison
	Mean (s.d.)	Obs.	Mean (s.d)	Obs.	(s.e)
Panel A: Spring level data					
Ln. *E. coli* MPN (CFU/ 100 ml)	3.90	50	3.86	149	0.039
	(2.06)		(1.91)		(0.330)
Water is high quality (*E. coli* MPN<=1)	0.08	50	0.05	149	0.033
					(0.042)
Water is moderate quality (*E. coli* MPN 2-100)	0.56	50	0.62	149	-0.057
					(0.081)
Water is hi or mod. quality (*E. coli* MPN <100)	0.64	50	0.66	149	-0.024
					(0.079)
Water is poor quality (*E. coli* MPN 100-1000)	0.24	50	0.26	149	-0.022
					(0.071)
Water is very poor quality (*E. coli* >=1000)	0.12	50	0.07	149	0.046
					(0.051)
Latrine density (fraction of homes with latrines)	0.85	49	0.87	144	-0.020
	(0.16)		(0.15)		(0.026)
Average diarrhea prevention knowledge score	7.05	49	6.85	144	0.202
	(0.85)		(1.10)		(0.152)
Iron roof density (fraction of compounds with iron roof)	1.32	49	1.31	144	0.007
	(0.21)		(0.22)		(0.035)
Panel B: Household summary statistics					
Ln. *E. coli* MPN (CFU/ 100 ml)	3.28	374	3.32	1121	0.038
	(2.13)		(2.28)		(0.165)
Water is high quality (*E. coli* MPN<=1)	0.15	374	0.12	1121	(0.028
					(0.022)
Water is moderate quality (*E. coli* MPN 2-100)	0.56	374	0.61	1121	-0.057*
					(0.033)
Water is hi or mod. quality (*E. coli* MPN <100)	0.71	374	0.74	1121	-0.029
					(0.034)
Water is poor quality (*E. coli* MPN 100-1000)	0.20	374	0.19	1121	0.009
					(0.028)
Water is very poor quality (*E. coli* >=1000)	0.09	374	0.07	1121	0.020
					(0.017)
Respondent years of education	5.43	365	5.77	1083	-0.34
	(3.65)		(3.58)		(0.25)
No. of children under age 12 in the compound	3.80	362	3.96	1084	-0.16
	(2.48)		(2.58)		(0.17)
Iron roof indicator	0.66	378	0.66	1137	0.00
					(0.04)
Walking distance to main drinking water source (minutes)	11.12	357	10.10	1068	1.02
	(9.83)		(8.88)		(0.74)
No. of water collection trips per week by household	45.6	378	47.1	1137	-1.6
	(38.1)		(35.0)		(2.2)

	Treatment		Comparison		Treatment – Comparison
	Mean (s.d.)	Obs.	Mean (s.d)	Obs.	(s.e)
Ever uses “assigned” spring indicator	0.66	378	0.67	1137	-0.01
					(0.05)
Sole source user (uses only assigned spring)	0.55	316	0.51	988	0.04
					(0.05)
Fraction of respondent trips to “assigned” spring	0.69	341	0.72	1020	-0.02
					(0.02)
Rates spring water “very clean” (rainy season)	0.30	378	0.30	1137	-0.00
					(0.04)
Rates water at spring “very clean” (dry season)	0.69	378	0.68	1137	0.01
					(0.04)
Average age of water collector	25.90	337	25.76	1011	0.14
	(9.81)		(9.32)		(0.73)
Fraction of water trips by those under age 12	0.07	362	0.11	1076	0.03
					(0.04)
Water storage container in home was covered	0.79	378	0.80	1137	-0.01
					(0.03)
Yesterday's drinking water was boiled indicator	0.21	378	0.27	1137	-0.06**
					(0.03)
Respondent diarrhea prevention knowledge score	7.15	362	6.99	1075	0.16
	(2.09)		(2.27)		(0.15)
Respondent said clean water prevents diarrhea	0.67	363	0.67	1083	0.00
					(0.03)
Household compound is clear of debris	0.49	378	0.51	1137	-0.02
					(0.04)
Household has soap in the home	0.87	378	0.87	1137	0.00
					(0.03)
Panel C: Child demographics and health					
Child age[§]	1.71	487	1.71	1038	0.001
	(0.93)		(0.97)		(0.041)
Child gender (=1 if male)	0.51	487	0.51	1038	-0.007
					(0.03)
Child had diarrhea in past week indicator	0.20	419	0.23	1478	-0.028
					(0.025)
Child height (cm)	92.92	460	93.63	1038	-0.708
	(12.59)		(13.97)		(1.00)
Child weight (kg)	14.17	463	14.32	1038	-0.149
	(3.58)		(3.62)		(0.32)

Notes: In the final column, Huber-White robust standard errors are presented (clustered at spring level when using household level data), significantly different than zero at * 90% ** 95% *** 99% confidence. Standard deviations not presented for indicator variables.

Diarrhea is defined as three or more looser than normal stools per day.

Assigned spring is the spring that we believed households used at baseline, based on spring user lists.

Household survey respondent is the mother of the youngest child in the compound (or the next youngest woman available).

[§] All children in the sample in Panel C were reported to be under age 3 at baseline.

Table 2: Spring protection source water quality impacts, difference-in-differences

	Panel A: Dependent variable, Ln(Spring *E. coli* MPN)			Panel B: Dependent variable, Spring water high/moderate quality (*E. coli* <100)		
	Treatment, mean (s.d.)	Comparison, mean (s.d.)	T – C (s.e.)	Treatment, mean (s.d.)	Comparison, mean (s.d.)	T – C (s.e.)
Before protection, mean (s.d.)	3.97	3.80	0.16	0.60	0.68	-0.08
	(2.09)	(1.90)	(0.35)			(0.08)
After protection, mean (s.d.)	2.64	3.79	-1.15	0.89	0.68	0.21
	(2.18)	(1.85)	(0.36)***			(0.06)***
After – Before difference (s.e.)	-1.33	-0.01	-1.32	0.29	0.00	0.29
	(0.44)***	(0.20)	(0.48)***	(0.08)***	(0.05)	(0.10)***
% Change in contamination	-74%	-1%	-73%			

Notes: N=175 springs. Huber-White robust standard errors (clustered at the spring level) are presented, significantly different than zero at * 90% ** 95% *** 99% confidence. MPN stands for “most probable number” coliform forming units (CFU) per 100ml. Standard deviations not reported for indicator variables. Percent change in contamination calculated as (1-exp(After – Before difference) * 100.

Table 3: Spring protection source water quality impacts, regression specifications

	Dependent variable: ln(Spring water *E. coli* MPN)			Dependent variable: Spring Water is high/moderate quality (*E. coli* MPN <100)		
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment group indicator	-1.17***	-1.11***	-1.18***	0.22***	0.27**	0.27**
	(0.35)	(0.34)	(0.33)	(0.06)	(0.11)	(0.11)
Baseline ln(Spring water *E. coli* MPN)	0.19**	0.27***	0.26***			
	(0.08)	(0.09)	(0.09)			
Baseline ln(Spring water *E. coli* MPN) * Treatment indicator		-0.28	-0.26			
		(0.17)	(0.18)			
Baseline spring water is high/moderate quality (*E. coli* <100)				0.13*	0.15*	0.15*
				(0.07)	(0.09)	(0.09)
Baseline spring water high/moderate quality * Treatment					-0.07	-0.05
					(0.13)	(0.13)
Diarrhea prevention knowledge			0.09			0.01
			(0.13)			(0.03)
Diarrhea prevention knowledge * Treatment indicator			0.26			-0.07
			(0.39)			(0.06)
Latrine density			0.65			-0.01
			(1.08)			(0.27)
Latrine density * Treatment indicator			-1.45			0.55
			(3.00)			(0.55)
Iron roof density			0.19			-0.05
			(0.84)			(0.22)
Iron roof density * Treatment indicator			-1.05			-0.09
			(2.27)			(0.39)
District-wave (season) fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.20	0.21	0.23	0.14	0.15	0.17
Observations	175	175	175	175	175	175
Mean of dependent variable (s.d.)	3.79 (1.85)	3.79 (1.85)	3.79 (1.85)	0.68 (0.47)	0.68 (0.47)	0.68 (0.47)

Notes: Estimated using OLS. Huber-White robust standard errors are presented, significantly different than zero at * 90% ** 95% *** 99% confidence. MPN stands for “most probable number” coliform forming units (CFU) per 100ml.

Average diarrhea prevention knowledge calculated as average of demeaned sum of number of correct responses given to the open ended question “to your knowledge what can be done to prevent diarrhea?”

All variables that are interacted with the treatment indicator are de-meaned.

Table 4: Treatment effects on household water source choice and health behaviors

	Coefficient (s.e.) on treatment indicator Full sample	Coefficient (s.e.) on treatment indicator Sole source users	Coefficient (s.e.) on treatment indicator Multi-source users	Mean (s.d.) in comparison group in follow-up survey, Full sample
Dependent variable	(1)	(2)	(3)	(4)
Panel A: Water collection and source choice				
Use assigned spring for drinking water indicator	0.22 (0.05)***	0.07 (0.02)***	0.36 (0.07)***	0.70
Fraction of trips to assigned spring	0.11 (0.04)***	0.05 (0.03)*	0.18 (0.05)***	0.70
Perceive water at assigned spring to be very clean (rainy season)	0.36 (0.06)***	0.41 (0.08) ***	0.32 (0.07)***	0.19
Perceive water at assigned spring to be very clean (dry season)	0.15 (0.05)***	0.11 (0.05)**	0.18 (0.08)**	0.73
No. trips made to get water (all uses, members, sources) past week	-3.18 (3.10)	-3.88 (4.05)	-2.67 (4.66)	52.00 (33.06)
Panel B: Water transportation and storage				
Average age of water collector (years)	0.05 (0.67)	1.27 (0.93)	-1.04 (0.94)	25.2 (9.20)
Fraction of water trips by those under age 12	-0.01(0.05)	0.01(0.04)	-0.03(0.08)	0.08
Water storage container in home covered indicator	0.00 (0.03)	0.01 (0.04)	-0.01 (0.04)	0.85
Ever treat water with chlorine indicator	0.05 (0.05)	0.10 (0.06)	0.01 (0.06)	0.27
Yesterday’s drinking water boiled indicator	0.04 (0.03)	-0.02 (0.05)	0.09 (0.05)*	0.24
Panel C: Complementary sanitation and hygiene behaviors				
Diarrhea prevention knowledge score (demeaned)	-0.25 (0.20)	-0.34 (0.27)	-0.18 (0.28)	0.01 (1.93)
Respondent says drinking clean water is a way to prevent diarrhea	0.01 (0.04)	0.01 (0.06)	0.02 (0.05)	0.71
Household compound is clear of debris indicator	0.06 (0.05)	0.05 (0.07)	0.07 (0.07)	0.60
Household has soap in the home indicator	0.01 (0.02)	0.01 (0.04)	0.01 (0.04)	0.85

Notes: N=1136 households at 175 springs (full sample) 537 households are sole source users. Each row reports the differences-in-differences treatment effect estimate from a separate regression where dependent variable is reported in first column. Huber-White robust standard errors clustered at the spring level are presented, significantly different than zero at * 90% ** 95% *** 99% confidence. Standard deviations of the dependent variable not reported for indicator variables. Reported means of the dependent variables are in the comparison group follow-up (post-treatment) survey.
Assigned spring is the spring that we believed households used at baseline based on spring user lists.

Table 5: Spring protection household water quality impacts, difference-in-differences

	Panel A: Full sample, Dependent variable: ln(*E. coli* MPN)			Panel B: Full Sample, Dependent variable: Water is high/moderate quality (*E. coli* MPN <100)		
	Treatment, mean (s.d.)	Comparison, mean (s.d.)	T – C (s.e.)	Treatment, mean	Comparison, mean	T – C (s.e.)
Before protection, mean (s.d.)	3.33	3.26	0.08	0.70	0.74	-0.04
	(2.26)	(2.17)	(0.17)			(0.04)
After protection, mean (s.d.)	3.08	3.34	-0.25	0.77	0.72	0.05
	(2.16)	(2.11)	(0.17)			(0.03)
After – Before difference (s.e.)	-0.25	0.08	-0.33	0.07	-0.02	0.09
	(0.21)	(0.11)	(0.23)	(0.04)	(0.02)	(0.05)*
% Change in contamination	-22%	8%	-28%			
	Panel C: Sole Source Spring Users, Dependent variable: ln(*E. coli* MPN)			Panel D: Sole Source Spring Users, Dependent var.: Water is high/moderate quality (*E. coli* MPN<100)		
	Treatment, mean (s.d.)	Comparison, mean (s.d.)	T – C (s.e.)	Treatment, mean	Comparison, mean	T – C (s.e.)
Before protection, mean (s.d.)	3.66	3.25	0.40	0.63	0.73	-0.10
	(2.33)	(2.11)	(0.27)			(0.06)*
After protection, mean (s.d.)	3.14	3.47	-0.34	0.75	0.72	0.03
	(2.15)	(2.03)	(0.24)			(0.05)
After – Before difference (s.e.)	-0.52	0.22	-0.74	0.12	-0.01	0.13
	(0.30)*	(0.14)	(0.33)**	(0.06)*	(0.03)	(0.07)*
% Change in contamination	-41%	20%	-52%			
	Panel E: Multi-source Spring Users, Dependent variable: ln(*E. coli* MPN)			Panel F: Multi-source Spring Users, Dependent var.: Water is high/moderate quality (*E. coli* MPN<100)		
	Treatment, mean (s.d.)	Comparison, mean (s.d.)	T – C (s.e.)	Treatment, mean	Comparison, mean	T – C (s.e.)
Before protection, mean (s.d.)	3.06	3.22	-0.16	0.75	0.75	0.00
	(2.18)	(2.20)	(0.22)			(0.04)
After protection, mean (s.d.)	3.03	3.19	-0.15	0.78	0.72	0.06
	(2.18)	(2.16)	(0.22)			(0.04)
After – Before difference (s.e.)	-0.03	-0.03	0.00	0.03	-0.03	0.06
	(0.24)	(0.15)	(0.28)	(0.05)	(0.03)	(0.06)
% Change in contamination	-3%	-3%	0%			

Notes: N = 1191 households at 175 springs in Panels A and B, N = 559 households at 152 springs in Panels C and D, and N = 1262 households at 162 springs in Panels E and F. Huber-White robust standard errors are presented clustered at spring level, significantly different than zero at * 90% ** 95% *** 99% confidence. Standard deviation not reported for indicator variables. Percent change in contamination calculated as (1-exp(After – Before difference) * 100.

Table 6: Spring protection household water quality impacts, regression specifications

	Dependent variable: ln(Home water *E. coli* MPN)					
	Full sample		Sole Source Spring Users		Multi-source Users	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment indicator	-0.28	-0.447*	-0.326	-0.828**	-0.231	-0.051
	(0.17)	(0.257)	(0.233)	(0.352)	(0.214)	(0.397)
Baseline ln(Spring water *E. coli* MPN)	0.01	0.053	-0.057	0.028	0.076	0.097*
	(0.04)	(0.041)	(0.056)	(0.063)	(0.049)	(0.056)
Baseline ln(Spring water *E. coli* MPN) * Treatment indicator		-0.171**		-0.288***		-0.094
		(0.080)		(0.110)		(0.100)
Baseline diarrhea prevention score	-0.02	-0.031	0.017	-0.004	-0.061	-0.049
	(0.03)	(0.032)	(0.042)	(0.048)	(0.042)	(0.047)
Baseline diarrhea prevention score * Treatment Indicator		0.019		0.110		-0.083
		(0.078)		(0.103)		(0.102)
Baseline boil water yesterday	0.10	-0.073	0.283	0.128	-0.095	-0.305
	(0.14)	(0.160)	(0.180)	(0.213)	(0.214)	(0.244)
Baseline boil water yesterday * Treatment indicator		0.646**		0.489		0.830*
		(0.318)		(0.390)		(0.485)
Baseline latrine density	-0.99**	-1.075*	-0.553	-0.717	-1.419**	-1.524**
	(0.49)	(0.549)	(0.625)	(0.689)	(0.654)	(0.726)
Baseline latrine density * Treatment indicator		0.405		0.288		0.271
		(1.250)		(1.691)		(1.464)
Baseline distance to drinking water (min.)	0.00	0.009	-0.002	0.008	0.007	0.017
	(0.01)	(0.007)	(0.010)	(0.010)	(0.009)	(0.010)
Baseline distance to drinking water (min.) * Treatment indicator		-0.024		-0.031		-0.028
		(0.016)		(0.024)		(0.021)
Treatment effect point estimate at sample means		-0.30 (0.03)		-0.36 (0.06)		-0.25 (0.05)
Mean of dependent variable in comparison group (s.d.)	3.2 (2.13)	3.2 (2.13)	3.38 (2.08)	3.38 (2.08)	3.13 (2.16)	3.13 (2.16)
R^2	0.04	0.05	0.05	0.08	0.05	0.06
Observations (spring clusters)	1137 (175)	1137(175)	537(152)	537(152)	600(160)	600(160)

Notes: Estimated using OLS. Huber-White robust standard errors clustered at the spring level are presented, significantly different than zero at * 90% ** 95% *** 99% confidence. MPN stands for “most probable number” coliform forming units (CFU) per 100ml. Additional control variables included are: season fixed effects, number of children under 12 living in the home, mother’s years of education, home has iron roof indicator, iron roof density within spring community. When differential treatment effects reported (columns 2, 4, 6) we also include interactions with all of these control variables and the treatment indicator (not shown). Baseline spring water quality, latrine density, distance to water source, and diarrhea prevention score are de-meaned.

Table 7: The elasticity of household water quality with respect to spring water quality

	Dependent variable: ln(Home water *E. coli* MPN)			
	Full sample OLS	Full sample IV	Sole source users IV	Multi-source users IV
	(1)	(2)	(3)	(4)
ln (Spring water *E. coli* MPN)	0.190***	0.332**	0.707**	0.075
	(0.028)	(0.166)	(0.324)	(0.199)
Indicator for treatment group		0.074	0.208	-0.144
		(0.140)	(0.250)	(0.220)
Indicator for after treatment		0.098	0.214	-0.014
		(0.108)	(0.170)	(0.153)
Diarrhea prevention knowledge score	0.002	-0.002	0.041	-0.031
	(0.022)	(0.023)	(0.038)	(0.032)
Boiled water yesterday indicator	0.102	0.096	0.323*	-0.141
	(0.108)	(0.110)	(0.176)	(0.167)
Latrine density	-0.639	-0.750*	-1.304	-0.558
	(0.414)	(0.436)	(0.801)	(0.562)
Distance to drinking water (min.)	-0.001	0.000	0.006	0.004
	(0.005)	(0.005)	(0.011)	(0.007)
District-wave (season) fixed effects?	Yes	Yes	Yes	Yes
R^2	0.04	0.07	0.07	0.07
Observations (spring clusters)	2280 (175)	2280 (175)	1074 (152)	1200 (160)
Mean of dep. var. (s.d.)	3.2 (2.12)	3.2 (2.12)	3.38 (2.07)	3.13 (2.16)

Notes: Regressions estimated using OLS (column 1) and instrumental variables (columns 2-4). Huber-White robust standard errors clustered at the spring level are presented, significantly different than zero at * 90% ** 95% *** 99% confidence. MPN stands for “most probable number” coliform forming units (CFU) per 100ml. All continuous variables are demeaned. Diarrhea prevention knowledge calculated as sum of number of correct responses given to the open ended question “to your knowledge what can be done to prevent diarrhea. Additional controls included are: number of children in home compound, respondent’s years of education, iron roof indictor and iron roof density in the spring community. In the IV regressions, the instrument for spring water quality is the After indicator * Treatment indictor term.

Table 8: Child health outcomes for children under age three at baseline

	Dependent variable: Diarrhea in past week		Dependent variable: Weight (kg.)		Dependent variable: Height (cm.)	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment group indicator	-0.016	-0.083	0.111	0.134	0.830	0.705
	(0.034)	(0.040)**	(0.168)	(0.221)	(0.563)	(0.712)
After treatment indicator	0.050	0.062	0.293	0.314	0.768	1.402
	(0.024)**	(0.033)*	(0.102)***	(0.121)**	(0.351)**	(0.432)***
After treatment * Treatment	-0.005	-0.022	0.135	0.027	-0.598	-0.737
	(0.047)	(0.055)	(0.186)	(0.232)	(0.521)	(0.646)
Multi-source user		-0.049		0.115		1.238
		(0.036)		(0.158)		(0.536)**
Multi- source user * Treatment		-0.022		-0.045		-1.251
		(0.046)		(0.153)		(0.567)**
Multi- source user * After		0.147		-0.043		0.260
		(0.075)**		(0.301)		(1.016)
Multi- source user * Treatment * After		0.033		0.208		0.268
		(0.082)		(0.359)		(1.198)
R^2			0.62	0.63	0.69	0.70
Observations	1740	1740	1826	1826	1826	1826
Mean of the dep var in comp. group (s.d)	0.28	0.28	11.83	11.83	83.47	83.47
			(3.17)	(3.17)	(10.4)	(10.4)

Notes: Estimated using Probit (columns 1-2) and OLS (columns 3-6). When Probit used, marginal effects reported. Huber-White robust standard errors are presented, significantly different than zero at * 90% ** 95% *** 99% confidence. Sample restricted to children under age three at baseline. Diarrhea defined as three or more looser than normal stools in 24 hours. Severe outliers in terms of either body mass index (BMI) and in the change in BMI (across the two household survey rounds) are excluded; severe outliers are defined as observations more than three times the interquartile range away from the mean. Age and gender controls include but not reported. District-wave (season) fixed effects included.