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
google.org

First draft: July 2006
This draft: August 2007

Abstract: Diarrhea, particularly from water-related causes, kills almost two million children annually. We study the impact of source water quality improvements achieved via spring protection on diarrhea prevalence and other outcomes 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*. There are smaller gains in home water quality. Reported child diarrhea incidence falls by a marginally significant one fifth. Spring protection appears less cost effective than point of use water treatment in reducing diarrhea. Households greatly increase their use of protected springs, and these changes in household water source choices are used to derive revealed preference estimates of willingness to pay for improved water quality in a travel cost analysis. Households are willing to pay US$4.52-9.05 per year on average for protected spring water. Assuming the principal benefit of improved water quality is better child health implies that households are willing to pay US$0.83-1.67 to avoid one child diarrhea episode. Stated preference valuations for spring protection yield much higher willingness to pay estimates, sometimes by a factor of three, casting doubt on the reliability of stated preference methods to capture valuations for environmental amenities in a setting like ours.

---

* This research is supported by the Hewlett Foundation, USDA/Foreign Agricultural Service, International Child Support (ICS), Swedish International Development Agency, Finnish Fund for Local Cooperation in Kenya, google.org, and the Bill and Melinda Gates Foundation. We thank Alicia Bannon, Lorenzo Casaburi, Anne Healy, Jie Ma, Owen Ozier, Camille Pannu, Eric Van Dusen, Melanie Wasserman, Heidi Williams and especially Clair Null for excellent research assistance, and thank the field staff, especially Polycarp Waswa and Leonard Bukeke. Jack Colford, Alain de Janvry, Andrew Foster, Michael Greenstone, Michael Hanneman, Danson Irungu, Ethan Ligon, Steve Luby, Enrico Moretti, Kara Nelson, Sheila Olmstead, Judy Peterson, Rob Quick, Elisabeth Sadoulet, Sandra Spence, Ken Train, Chris Udry, and numerous seminar participants have provided helpful comments. Preliminary draft, comments welcome. All errors are our own.
-- Corresponding author: Edward Miguel (emiguel@econ.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 to a constructed public water point (standpipe, borehole with hand pump, protected spring, protected well or rainwater collection point) within one kilometer of their home.[1]

A central rationale for promoting safe drinking water is the persistently high level of water-related disease 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 two 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. Despite the call to arms in the MDGs, in fact it remains unclear whether investing in the water sector is the most effective way of reducing the diarrheal disease burden. Randomized trials have established that several other health interventions—increased breastfeeding, immunization, oral rehydration therapy (ORT), micronutrient supplementation—are both effective

[1] 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).

and cost-effective in preventing diarrhea (see Hill *et al*. 2004).[2] 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 in-home point-of-use (POU) treatment to reduce microbiological contamination is most cost-effective. While several studies from the 1980s find that source water quality interventions reduce improve child health outcomes, a recent strand of the academic literature has increasingly emphasized water quality at the time of consumption rather than at the collection point. 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, consistent with the MDGs.[3]

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 runoff, improving quality at an already existing 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). Protected springs are included in the standard World Health Organization definition of an "improved" water source and thus this is

[2] 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).
[3] The current study is one component of a larger project also examining point-of-use, water quantity, and health, which may provide guidance on whether there is scope for some readjustment of priorities in the rural water sector.

an example of the type of investment being made to fulfill the water-related MDG. 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.

Using a randomized impact evaluation approach, in which spring protection is phased-in across nearly 200 springs in a randomized order, we estimate impacts of spring protection 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 isolate the impact of a single treatment rather than a package of services. Second, we use a randomized design with a large sample size and several rounds of follow-up data and are able to take intra-cluster correlation into account, thus making it easier to distinguish the impact of water improvements from potentially confounding omitted variables and from background noise. 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, and are thus able to directly assess the extent to which source water quality improvements translate into household water gains. Fourth, we have data on household hygiene and sanitation at baseline and are able to evaluate the claim that source water quality improvements are most valuable in the presence of pre-existing access to improved sanitation and hygiene practices.

In the second part of the analysis, we explore how household behaviors – most importantly, the choice of water source – change in response to source water quality improvements when many households can choose from among several different water sources. We develop a formal economic model of the water source choice decision, in which households trade-off the distance walked to a water source against water quality. This framework highlights the importance of accounting for endogenous household sorting among water sources in the econometric analysis, and allows us to develop revealed preference estimates of average household willingness-to-pay for source water quality improvements using a conditional logit 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.

We find that spring protection very effectively improves the quality of water at the source, reducing fecal contamination by approximately three quarters. Spring protection is also partially effective at improving household water quality, reducing contamination by roughly one quarter. The incomplete pass through from spring-level water quality gains to the home is likely due in part to people obtaining water from multiple sources and in part to recontamination in transport and storage. There is little evidence that the limited home water gains are due to 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 better household water quality. While we do find some measurement error in our water quality measures, it is not of sufficient magnitude to explain the gap between source water quality and home water quality that we observe.

Diarrhea among young children in treatment households falls by a marginally significant 4.7 percentage points, or one-fifth, after up to thirty months of spring protection. Yet calculations using results from other recent studies suggest that spring protection is less cost effective than point of use water treatment in averting cases of childhood diarrhea. The reductions in child diarrhea that we observe do not translate into any detectable improvements in child anthropometrics.

We estimate willingness to pay (WTP) for improved source water by analyzing how 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, in a conditional logit discrete choice model. We find that households shift their water collection patterns quite dramatically in response to spring protection. In addition to allowing us to uncover a fundamental behavioral parameter in households' utility functions, this revealed preference 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, it provides guidance on the magnitude of feasible user-fees at water sources. These revealed preference results indicate that the average valuation of spring protection is on the order of US$4.52-9.05 per household per year.

We are also able to provide evidence on households' preferences for better child health by combining these spring protection WTP figures and the estimated reduction in child diarrhea episodes due to spring protection. Under the assumption that all household valuation of better quality water is due to improved child health, we estimate that households are willing to pay US$0.83-1.67 to avert a single child diarrhea episode. To the extent that households obtain other benefits from spring protection, this should be considered an upper bound. This value is very similar to the cost per case of diarrhea averted through spring protection, but falls below the cost usually found for "software" interventions like hygiene or handwashing education (Varley et al. 1998). However, these estimates of cost per case averted encompass only the costs of the software intervention borne by the public health sector, not the costs to households of changing their behavior; with these costs to the household factored in, the costs of software interventions would likely be even higher.

Finally, we contrast the revealed preference figures with those from two different stated preference methodologies, stated preference ranking of water sources and contingent valuation. Environmental economics has long been interested with comparing revealed preference and stated preference estimates of willingness to pay for amenities, however such data is rarely available in a single setting, and almost never in less developed countries (Carson et al. 1996). Both of these approaches generate much higher willingness to pay estimates than the revealed preference travel cost approach, in some cases by as much as three times. The large discrepancy casts doubt on the reliability of stated preference methods to capture household valuations for environmental amenities like cleaner water in settings like ours.

2 Related Literature

Two influential papers (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).[4] Esrey *et al.* (1991) attempt to separate the relative impacts of water supply, sanitation, and hygiene education interventions on diarrheal morbidity. 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 an investment in water quantity and water quality together. Using multivariate regression analysis of household infrastructure and diarrhea prevalence in 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 epidemiology (Fewtrell *et al*. 2005) reports that source water quality improvements, sanitation interventions, hygiene programs, and point-of-use water 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). Fewtrell *et al*. (2005) conclude that point-of-use water treatment may be more effective than source water quality interventions because of recontamination during transportation and storage. Similarly, Wright *et al*. (2004) analyze 57 studies that measured both source and in-home water quality, and conclude that improvements in source quality are often compromised by post-collection contamination. However, these evaluations of source water quality investments remain less methodologically rigorous than evaluations of point-

[4] 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). 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.

of-use water treatment.[5] Moreover, to our knowledge there is no study in which household water quality has been measured following exogenous changes 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 exploit experimental variation in source water quality to directly measure the extent to which source water quality impacts diarrhea incidence using a longitudinal household dataset.

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 and child health in a less developed country context using a conditional logit estimation 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 stated willingness to pay for piped water connections (World Bank Water Demand Research Team 1993).

The relative shortcomings of such stated preference and 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

---

[5] There are two prospective studies of source water quality interventions that suggest positive impacts 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 5 km. The published article does not mention if these villages were randomly selected. Following the intervention children between six months and five years of age experienced 25% fewer diarrhea episodes 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 effect of improved sanitation beyond that achieved by wells alone. 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 nutrition in Nigeria. 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 (less than 80% of desirable weight-for-height) among children under three years of age declined significantly in the intervention villages. Generalizing the results to other settings is hampered by their small sample sizes (each includes only five villages), and the fact that they evaluate interventions that improved both water quality and quantity simultaneously (by providing wells).

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, environmental economist have developed several alternative approaches to eliciting willingness to pay based on actual behavior. 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 (McFadden 1974, Phaneuf and Smith 2003). To our knowledge, our estimates below are the first such application of the travel cost approach to value improved drinking water quality in a less developed country context. Water choices in rural less developed country settings have been studied by Whittington, Mu, and Roche (1990) and Mu, Whittington, and Briscoe (1990), however neither accounts for the role of water quality in the source choice decision (they focus on distance and price) 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 (agricultural wages range from US$1-2 per day) and few households have access to improved water services. Both local law and custom require that private landowners allow public access to water sources on their land. Landowners therefore do not have incentives to improve a water source and recoup the cost of such an investment via the collection of user fees. There is no elected local government so spring protection is generally

undertaken by donors or the central government, often in conjunction with user groups set up to collect maintenance funds. However, collective action problems mean that investments in local public goods with positive returns often fail to occur.

Springs for this study were selected from a universe of local unprotected springs by a non-governmental development organization, International Child Support (ICS). The NGO first obtained lists of all local unprotected springs in the Busia and Butere-Mumias districts from Government of Kenya Ministry of Water offices. NGO field and technical staff then visited each site to determine which springs were suip.2(inr.5(rro   ontena7T SFwnps se7T .9(s)-1.easc0.lsit7 sprin)-7drsprin)-7ectmsit7 onthena7T

springs should not be related to treatment assignment: when the NGO was first informed that some sampled 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 in Table 1 if attention is restricted to the 184 springs where protection is viable (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 (using a computer random number generator) from among this spring user list for the household sample. 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 at least sometimes during subsequent household surveys, attesting to the validity of the method used to identify users. The few spring non-user households were nonetheless retained in the sample throughout the analysis.

The spring user households are largely representative of all households living near the springs. In a February 2007 census of all households living within approximately a 10 minute walk of seven sample springs, we found that 92% of these nearby households had been included on our original spring users list. The spring user list households may be less representative for households living farther than a 10 minutes walk away from sample springs.

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 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. 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 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 for repairing any defects during the first three months after protection, after which the protected spring was "handed over" to the community as their property.

A first 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 three to four months after the first round of spring protection, in April through August 2005. In this survey, water quality data was not collected at nine springs due to logistical issues. Surveys were administered and water quality data was collected at 1250 of the 1389 households with complete baseline survey data.

The second round of spring protection was performed in August-November 2005, and the second follow-up survey was collected in August-November 2006. In this survey, water quality data was collected at all springs but one, and there was a similarly low rate of sample attrition among households as in the first follow-up. The third follow-up survey round took place from January to March 2007. In total there are 1,354 households with baseline data and at least one survey follow-up round, and we consider these households in the analysis.

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 information on water source choices and 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.[6] Samples were 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

[6] At springs, the protocol is as follows: the cap of a 250 ml bottle is removed aseptically and not touched during sample collection. 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 at unprotected springs and sample bottles are filled from the water outflow pipe at protected springs. 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 reflect the conditions in the household's own water storage container and dipper.

present in fecal matter.[7, 8] Continuous, quantitative measures of fecal contamination are available after 18-24 hours of 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 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 this variability it to be appropriately modeled and accounted for statistically. We do not yet have such data and, to our knowledge, neither do any 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 attenuation bias 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 partially address this issue as well as the problem of omitted variable bias (confounding) more generally, as we discuss below.[9]

---

[7] 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.

[8] 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 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.

3.2.3 Household survey data

A household survey was administered to a representative sample of spring user households at all sample springs prior to the intervention, and again following each round of spring protection.[10] The target survey respondent was the mother of the youngest child living in the home compound (where the extended family often resides together) or another woman of child-bearing age, if the mother of the youngest child was not available. The respondent is asked about the health of all children under age five living in the compound, including recent diarrhea and dysentery incidence.

The household survey also gathered baseline information about hygiene behaviors and latrine use. Data on the frequency of water boiling, home water chlorination and water collection choices was collected. Respondents were also asked to give their opinion on ways to prevent diarrhea; they were not given options to choose from, and 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 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 two to three such correct preventative activities, with 47% volunteering either boiling water or practicing good hygiene at baseline.

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

[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 below, this procedure generated a sample of households that used the springs for varying amounts of water in practice.

[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 survey list were "solar water disinfection", "breastfeeding", and some variant of "use compost pit/keep compound clean".

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 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 for children under three years of age as the main health and nutrition outcomes.

3.3 Attrition

We successfully followed up 90% of the baseline household sample in the first follow-up survey round, 89% in the second follow-up survey, and 92% of the baseline sample in the third follow-up. We have data from all four survey rounds for 79.5% of baseline households and three survey rounds for an additional 14.5% of households in the baseline sample, thus 94% of baseline households were surveyed in at least two of the three follow-ups. Attrition is not significantly related to spring protection assignment: the coefficient estimate on the treatment indicators are only -0.09 (standard error 0.10) for first treatment group, and 0.02 (standard error 0.10) for the second treatment group in a regression of the attrition indicator on treatment assignment. Thus treatment households are no more likely to be lost across survey rounds than other households, and this result is robust to including further explanatory variables as controls (not shown).

The baseline characteristics of the households that we lose over time are typically statistically indistinguishable from those that remain in the sample. Economically better-off households, such as those with iron roofs, do not appear any more likely to be lost from the sample, nor are households with better baseline household water quality or hygiene knowledge. Overall, any 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, although results are very similar with the main analysis sample.

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" (between 100 and 1000) and "very poor quality" (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 (using a computer random number generator) created broadly comparable program groups.[14] The spring water in our sample is of moderate quality on average. Only about 5 to 6% of samples from unprotected

[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. [CITE]
[14] 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 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).

springs would meet the stringent U.S. EPA drinking water standards, while over a third of samples are poor or very poor quality.[15]

Summary statistics for household water quality are presented next (Table 1, Panel B). Home water is somewhat more likely to be of high quality prior to spring protection in the treatment group (and the difference between treatment and comparison group means is significant at 95% confidence), but there is no statistically significant difference in the proportion of samples where water is of moderate or poor quality.

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 0.52 (s.d. 2.64, n = 1389 households; results not shown). This likely occurs for at least two reasons: first, many households collect water from sources other than the sample springs and these may be less contaminated on average, and second, some households use point of use (POU) treatments to improve home water quality. Only a bit more than one half of the household sample gets all their drinking water from the local sample spring at baseline and overall respondents make about 70% of all their water collection trips to their sample spring. In a cross-sectional regression, households that collect all their drinking water from their sample spring have significantly more contaminated home water (not shown), consistent with the view that unprotected springs are a relatively contaminated source, although the extent of contamination is likely to vary by season.

Some households report taking additional measures to treat their home water. For instance, about 25% of households report boiling their drinking water at baseline.[16] We also collected data on chlorination in the first follow-up survey: 28% of households reported chlorinating their water at

---

[15] 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 fixed effects in all regression analysis.

[16] Solar disinfection is also occasionally practiced in this area, but we did not collect data on this at baseline.

least once in the last six months.[17] However, 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. Social desirability bias is a leading concern. One potential explanation for the low correlation 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.

Household water samples are also held for a shorter length of time than spring water samples, on average.[18] 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 significantly different than zero 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 evidence that the randomization was successful at creating balanced program groups. Average mother's education attainment is equivalent to 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 greater relative 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 as about 85% of households report having a latrine, and the average walking distance (one-way) to the closest local water source is approximately 10 minutes.

[17] 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.
[18] 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.

There are similarly no significant differences across treatment and comparison groups in terms of the respondents' diarrhea prevention knowledge score, water boiling behavior, or self-reported understanding of the links between water quality and diarrhea. There are also no differences between compound cleanliness and soap ownership. However, 90% of treatment group households and 93% of comparison households cover their drinking water containers and this difference is significant at 95% confidence. It is unclear what accounts for this difference, but we conclude that it is unlikely to be an important indicator of home water quality differences because there is no difference in *ln E. Coli* across the groups.

We report summary statistics for the subset of children under age three for whom we have both baseline and follow-up survey data in Table 1, Panel C. Children are comparable across treatment and comparison groups in terms of health and nutritional status at baseline. For example, a fairly high 21% of children in the comparison group had diarrhea in the past week at baseline, as did 23% in the treatment group. There are similarly no statistically significant differences in other non-diarrheal illnesses (e.g., fever, cough) or in breastfeeding across the two groups (results not reported).

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 – and when the dependent variable is an indicator variable (for high quality water, *E. coli* MPN < 1, for example), although results are similar using probit analysis in the latter case (not shown).

$$W_{it}^{SP} = \alpha_{dt} + \beta_1 T_{it} + X_i^{SP}{}'\beta_2 + (T_{it} * X_i^{SP})'\beta_3 + \varepsilon_{it}. \qquad (1)$$

W_{it}^{SP} is the water quality measure at spring i at time t ($t \in \{0, 1, 2, 3\}$ for the four survey rounds) and X_i^{SP} are baseline spring and community characteristics (e.g., initial level of spring water contamination). The variable T_{it} is a treatment indicator that takes on a value of one after spring

protection has occurred, and this is the case for treatment group 1 in all follow-up survey rounds, and for treatment group 2 in the second and third follow-up survey round. ε_{it} is the standard white noise disturbance term.[19] Randomized assignment implies that the coefficient estimate of β_1 is an unbiased estimate of the reduced-form ITT effect of spring protection. In some specifications we explore the possibility of differential effects as a function of spring-level baseline characteristics, captured in the vector of coefficients β_3. District-wave (season) fixed effects α_{dt} are also included in the regression analysis to control for any time-varying factors that could affect all treatment groups.

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, Panels A and C) and then for an indicator of whether water is high quality (*E. coli* < 1 MPN, Panels B and D), as the first step in tracing out the impacts of the intervention on water at springs and in homes, and ultimately on child health. The top two panels of the table report results for the first round of treatment springs (protected in early 2005) versus other springs, using the baseline data and the first follow-up survey, while the bottom two panels of the table report results comparing both treatment groups together (both those protected in early 2005 and in late 2005) to the other springs, using the baseline and the third follow-up surveys.

Spring protection dramatically reduces contamination of source water with the fecal indicator bacteria *E. Coli*. Using both rounds of data indicates that the average reduction in ln. *E. coli* is between 72-78% (Table 2, Panels A and C), with nearly identical results across treatment rounds..

[19] 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.

Figure 3 is a non-parametric representation of the data that shows some gains are experienced at nearly all treatment springs, with impacts not clearly a function of baseline contamination.

It is difficult to predict how the observed reductions in source water contamination translate into health outcomes, since the relationship between water quality and health is not necessarily log-linear. A more natural measure of improvement in drinking water is to focus on whether source water meets the stringent EPA drinking water standard, what we call "high quality" water. We find that spring protection does increase the probability of high quality source water, but that relatively few springs achieve this standard even after protection (Table 2, panels B and D). The first round of follow-up data indicates that protection increases the probability of meeting EPA/WHO standards by 9 percentage points (nearly significant at 90% confidence), while in the third follow-up round protection increases the probability of meeting the standard by 35 percentage points (significant at 99% confidence). Yet only 39% of sources meet the EPA/WHO standard after protection.

These estimated spring protection treatment effects on source water quality are robust to including controls for baseline contamination and district-wave (season) fixed effects (Table 3, regressions 1 and 2). Regression analysis also suggests that spring protection does not lead to a significantly greater percentage reduction in water contamination when initial contamination was highest (regressions 3 and 4). We also test 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 (regression 4), but these interaction terms are not statistically significant.

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

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

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 can change their behavior in response to source water quality changes. The two most immediate choices households face are the choice of a water source, and the choice of whether or not to use point-of-use technologies (e.g., boiling or chlorination). We discuss these in turn below, but focus mainly on the water source choice. The fact that households in our study area have access to multiple water sources, varying both in the quality dimension and walking distance from the home, allows us to value water quality using a travel cost approach (Freeman 2003).

Imagine first that 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 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 between the sample spring and alternative source is $D \equiv D^s - D^a$, which we call the "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. For now we assume that households are homogeneous along all dimensions except for the distance gap, but relax this below.

In choosing a water source, households trade off the cost (the distance to the source) versus the benefits (improved water quality, which affects health). The opportunity cost of time – per minute here – 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 CD_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 may be non-health benefits to getting water from a low contamination water source (for instance, the improved appearance or ease of collecting water at a protected source) that are also captured in *V*.

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.[20]

Household utility from a single water collection trip to source $j \in \{s, a\}$ can be represented as the linear function $U_j = V(W_j) - CD^j$. Household *i* chooses the sample spring over the alternative source if the benefits of higher water quality outweigh travel costs, namely when $\{V(W_s) - V(W_a)\} - CD_i \geq 0$. More generally, in a context with multiple alternative water sources like our empirical setting, the household chooses the source that maximizes utility over all options in the choice set.

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 $D_i \leq \{V(W_s) - V(W_a)\}/C \equiv D^*$. Households with distance gap up to some threshold level use spring water, while those farther away choose the alternative source. After spring protection, spring water quality improves relative to the alternative water source, and households choose spring water if $D_i \leq \{V(W_s^T) - V(W_a)\}/C \equiv D^{**}$, where $D^{**} > D^*$ since spring water is now less contaminated than before ($W_s^T < W_s$). So households living at a greater distance from the spring increasingly choose spring water.

Endogenous source choice has implications for the quality of household drinking water. For households that were spring water users in the pre-treatment period ($D_i \leq D^*$, corresponding to the households that used the sample spring at baseline, the "sole-source" users in our data), their home

[20] We do have data on the water quality of alternative local sources, but only for the third follow up survey round, and so cannot explicitly test this assumption.

water quality is unambiguously better after treatment since they still rely exclusively on the spring for drinking water and its quality improves after protection.

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 baseline multi-source users in our data. For these households, home drinking water quality could theoretically increase or decrease after protection.[21] To illustrate, imagine better water quality at the spring induces a household to switch from a distant but high quality alternative source (say, a borehole well) to the closer but relatively lower quality spring. This could be optimal because households are trading off water quality against collection time. In this case, even if the water quality chosen by the household deteriorates somewhat since they increasingly use the now-protected spring, the household is still made better off by spring protection in that household members benefit from time savings. The theoretical prediction on the change in home water quality for these multi-source users 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 straightforward to calculate households' valuation of the water improvements caused by spring protection in this model, focusing on those households on the margin between using the spring and the alternative source. After the water quality improvement at the spring ($W_s^T < W_s$) that yields 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 revealed preference measure of the value of improved water quality. This model can also be used to estimate valuation for avoided illness as a result of consuming better water when combined with estimated health impacts of the intervention, in our case on child diarrhea.

[21] 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 in our data with the largest distance gaps relied at least partially on the sample spring for drinking water at baseline. This is due to the initial selection of sample households as at least occasional "spring users" living near the spring.

Other factors can be added to increase realism and bring the model closer to the data. First, there may be more then two alternative sources, and the water contamination levels of each of these springs and alternative sources vary. Second, households make multiple trips and each trip is affected by un-modeled factors including the weather, the queue at the water source, the direction they are walking for another task (i.e., walking to the market for food) or individuals' mood on a given day. These factors enter the decision problem through the idiosyncratic error term. Incorporating this i.i.d. error term e_{jt}, which can conveniently be modeled as type I extreme value, the utility of a water collection trip to source *j* at time *t* is: $U_{jt} = V(W_{jt}) – CD^j + e_{jt}$, and the spring is chosen by household *i* for trip *t* if this maximizes household utility over all possible sources *j*. This yields the usual logit form for choice probabilities. In practice we estimate a conditional logit model (McFadden 1974).

Households also face the choice of whether or not to adopt a POU water technology, such as water boiling or chlorination prior to water consumption. Consider the case of chlorination for concreteness. There are several costs to adopting at-home chlorination, which we denote C_P. These include the purchase price, 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 tastes worse than 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 employ 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

[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 and point-of-use technologies could be complements and spring protection could actually boost demand for point-of-use chlorination technologies.

technologies is low in our study area and we do not see large 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 framework incorporates a role for hygiene practices and access to sanitation. Influential research argues that water quality improvements alone are insufficient in improving 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 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 and sanitation could also be endogenized). H_i can concretely 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.[23] Below we estimate this 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 in order to gauge the impact of spring protection on household behaviors – including water source choice, self-reported

[23] Spring protection and hygiene/sanitation could also be substitutes or complements even if the ϕ function is linearly separable, as long as *V* is convex.

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 who only used the sample spring at baseline) and multi-source users (those who also used other sources for water). The predictions are first that use of the protected spring should increase among initially multi-source user households while sole-source user households continue to exclusively use the spring, 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 are also interested in testing Esrey's (1996) hypothesis that sanitation and hygiene are complements with source water quality improvements.

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 district-wave (season) fixed effects. Regression error terms are clustered at the spring level in these household-level regressions. Households using the same spring at baseline are not independent units of study and their outcomes may be correlated. Not only do these households share a common water source, but they may be related by kinship ties, and may share the use of 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:

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

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. As before, we control for baseline treatment group assignment as well as district-time effects. The common spring-level error component is captured by v_i and e_{ijt} is a standard white noise error term.[24]

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 periods is the instrument for spring water quality. The first-stage regression equation is nearly identical to equation 1 above. The treatment assignment indicators and the time effects are included as explanatory variables in both the first and second stage regressions. This IV approach provides a conceptually attractive means of estimating the degree of water contamination between source and home, especially if the sole source user households almost exclusively use the sample spring for drinking water in all periods.

Unfortunately, this latter assumption does not hold consistently across our four years of panel data. In the first follow-up survey round, 74% of comparison group baseline sole source spring users remained sole source users, but by the third follow-up round two years later only 64% remained sole

[24] A point of use intervention providing in-home chlorination was launched before the third follow-up survey (2007) in a random subset of households. Due to possible impacts on household water quality and behaviors, the third follow-up survey data for this subset of households is excluded from the analysis. We plan to study the impact of this POU intervention, and its interactions with source improvements, in future research.

source users, and overall only one third of comparison group baseline sole source users remained sole source users in all follow-up rounds in which they are observed. There are several possible explanations for this churning in sole source user status over time, including changes in households' other water options over time (as other sources are improved or deteriorate), and changing water collection costs due to shifting household composition. Regardless of the cause of this churning in water choices, our baseline definition of sole and multi source user households becomes less empirically meaningful over time, and as a result, we explore the pass-through of source water quality improvements below focusing on only the first follow-up survey round, when the baseline source choice definitions are most relevant.

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 trips made to collect water from the household's primary source), 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 for those households not previously using it: treated households are 20 percentage points more likely to use their sample spring as a source of drinking water if they used other sources at baseline (Table 4, Panel A). Sole source users already used the springs and so make few changes in spring use as a result of treatment. There are similarly large impacts on the fraction of water collection trips made to the sample spring after protection for multi-source users. 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 18 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.

There were small statistically significant effects of spring protection on the average distance households walked to their main drinking water source (recall that the average length was about 8 minutes one-way or 16 minutes round-trip). There was no overall effect on the number of trips made to water sources in the past week. Similarly, there are no significant changes in most water transportation and storage behaviors, although some small shifts in self-reported water boiling at home (Table 4, Panel B). Households at treated springs are somewhat more likely to boil water (suggesting that this is complementary rather than a substitute for spring protection), but it is unclear to what extent this is the result of reporting bias. There is also no evidence of changes in self-reported diarrhea prevention knowledge nor in other household hygiene measures (Panel C).

Survey enumerators collected additional information on the physical conditions at the spring and the extent of maintenance activities, and find that protected springs have significantly "clearer" water, better fencing and drainage, and less fecal matter and brush in the vicinity (Table 4, Panel D). In contrast, there is no effect on the observed yield of water at the spring, confirming that spring protection allows us to isolate water quality effects in the analysis rather than water quantity effects.

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. We focus on the natural log of *E. Coli* as a measure of contamination, and look separately at treatment effects for the full sample of households (Table 5, Panels A and B), for sole-source spring users who get all their water from the local spring (Panels C and D) and for multi-source users who collect water from several locations (Panels E and F). As in Table 2, we present estimated treatment effects for two follow-up survey rounds (2005 and 2007). In both cases, the average impact of spring

protection on home water quality is far smaller than the impacts on source water quality. Using the 2007 data for the full sample of households, the average reduction in water contamination is only 19% (Table 5, Panel B), only about one quarter the 77% reduction at the spring level (Table 2, Panel C) and not statistically significantly different from zero.

One theoretically coherent 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 spring sources, although the random churning among sources that we document above is also a plausible explanation. Among the sole-source spring users, spring protection impacts on home water quality are substantial, including 39% reduction in average home water contamination using the 2005 data (Table 5, Panel C), while for multi-source users home water gains are essentially zero (Table 5, Panel E). By 2007 household water contamination reductions are nearly identical for sole source and multi-source user households at 17-21% (Panels D and F).

Similar results obtain for the full sample when baseline household characteristics are included as explanatory variables in a regression framework (Table 6). Once again, the overall effect of spring protection on home water quality is moderate (regression 1) with slightly larger reductions in contamination observed for the sole-source households than the multi-source users (regressions 2-3) though we cannot reject equal treatment effects for sole source and multi-source users in these specifications. The average reduction in *E. Coli* contamination is roughly 22%.

The analytical payoff from the multiple regression framework lies in allowing us to estimate treatment effects for households with different baseline characteristics. We find no evidence of differential treatment effects as a function of household sanitation, diarrhea prevention knowledge, or mother's education (Table 6, regression 3). 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 differential effects as a function

of pre-existing sanitation access or hygiene knowledge runs counter to claims common in the literature that source water quality improvements are most valuable when these complementary factors are also in place. Perhaps surprisingly, baseline mother's diarrhea prevention knowledge is also not significantly related to observed household water quality in any regression specification. One possible explanation is that these measures miss some important dimension of hygiene or sanitation access, but if so it is not immediately obvious what these are. Home water contamination reductions are significantly smaller for households that report boiling their water, as expected if that behavior is already removing the worst contamination, suggestion that boiling water is a substitute for spring protection. There is no evidence of positive water quality spillovers for springs within 1, 2, or 3 kilometers of protected springs (results not shown).

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 using the first follow-up survey round, when the baseline sole source user definition is most meaningful. 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 0.22 (Appendix Table 1, regression 1). With only these results in hand, a naïve conclusion would be that water recontamination in transport and storage prevents nearly 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. Even when attention is restricted to sole-source spring user households, and thus endogenous sorting is largely avoided, this framework leads to a similar estimated elasticity of 0.23 (regression 2).

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 tells a different story for the sole-source users: the elasticity estimate rises dramatically to 0.66 (statistically significant at 95% confidence, ;Appendix Table 1, regression 3). In this subsample of households

where endogenous source choice is mostly eliminated, nearly two thirds of the source water quality gains at the source generated by spring protection are thus translated into home water quality gains.[25]

Taken together, this analysis is strong evidence against the claim that recontamination renders source water quality improvements useless in this setting. We conclude that the impacts of spring protection on household water quality are large and statistically significant for those households that mainly use the same water source (the baseline sole-source user households). Our longitudinal household survey and water quality data, together with the experimental program design that generated exogenous variation in source quality, allow us to reach different conclusions than would be suggested by existing analyses using observational cross-sectional data.

7 Child health and nutrition impacts

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

$$Y_{ijt} = \alpha_i + \alpha_{dt} + \beta_1 T_{ijt} + X_{ij}'\beta_2 + (T_{ijt} * X_{ij})'\beta_3 + u_{ij} + \varepsilon_{ijt} \qquad (3)$$

where the main dependent variable we focus on is diarrhea in the past week. The coefficient estimate on the variable indicating treatment captures the spring protection effect, β_1. We include child fixed effects (α_i) and district-time effects (α_{dt}). We also consider time-varying treatment effects.

The moderate home household water quality gains that we estimate lead to marginally statistically significant reductions in diarrhea for children under age 3. Diarrhea incidence falls by 4.7 percentage points, on a comparison group average of 23% of children with diarrhea in the past week, so a drop of roughly one fifth (Table 7, regression 1). Effects are somewhat smaller in the third year of treatment (regression 2). In contrast, there are no statistically significant impacts on either child weight or BMI over the follow-up surveys and estimated impacts are essentially zero (not shown).

[25] Note that the instrumental variables regression cannot be interpreted in the same way for the multi-source users, precisely because these households respond to treatment by switching among sources with variable water quality.

These moderate impacts are consistent with the primary causes of diarrhea being water washed (arising because of insufficient water for washing and bathing) rather than water borne (transmitted via ingestion of contaminated water). Certainly, spring protection could only be expected to address waterborne illness here since we see empirically that there are no changes in the number of trips made to collect water as a result of treatment (Table 4, Panel A), and thus no increase in water quantity. The reduction in diarrhea prevalence among children under age two that could be expected from addressing waterborne illness with a point-of-use (POU) water treatment is about one sixth (confidence interval -34% to -4%) per 100 weeks (Crump et al. 2005), essentially the same reduction that we estimate. (Crump et al 2005 was also carried out in rural western Kenya.) This is the reduction observed in a cluster randomized control trial of a cheap and readily available POU water treatment product that resulted in 78% of treatment households with *E. coli* MPN <1, a far greater improvement than we observe as a result of spring protection (where 27% of sole-source user households in the treatment group have *E. coli* MPN <1 after protection). One interpretation is that, in this region, moderate improvements in water quality are as effective as larger improvements.

While source water quality improvements and point of use water treatment are roughly equally effective, this does not imply that they are equally cost-effective. We next compare the health benefits associated with spring protection with alternative reductions in diarrheal morbidity that might have been realized if the approximately $100,000 spent to protect our 100 treatment springs had instead been spent providing point-of-use treatment products to households with young children, using the results in Crump *et al*. (2005). A one month's supply of the product used for in-home treatment (WaterGuard) costs about 20 Kenyan Shillings (or $0.29).

We begin by noting that 23% of children under age three in our sample are reported having diarrhea in the past week at baseline. All other things equal, this implies 370,760 cases at households that use sample springs over the ten years that a spring might last and (370,760) * (0.047/0.23) = 75,764 cases averted as a result of the intervention. This implies per averted case of diarrhea of

US$1.32, although spring maintenance would somewhat raise this cost (the NGO is spending approximately US$55 per spring per year in maintenance). Spring protection is increasingly cost effective the higher the local population density, and thus the greater the number of households that benefit from the intervention. The above cost per case of diarrhea averted would fall by half, for instance, if the number of spring using households doubled from 31 to 62.

If WaterGuard were to be given to every household with children under age three in our sample (about 80% of homes) for ten years, this would cost $65,657 in current dollars with a discount rate of 5%. The cost per case averted with WaterGuard is about $1.04, and thus more cases of diarrhea could be averted by focusing on point-of-use water treatment products instead of spring protection. Crump et al (2005) report WaterGuard take-up on the order of 70%, and we implicitly assume the same take up, but Crump's study involved costly frequent field visits to promote use; ongoing research in a related project suggests that roughly 50% of households use WaterGuard when it is provided without frequent follow-ups. If only half rather than 70% of Kenyan households use WaterGuard (and there is no tendency for those most vulnerable to diarrhea to use it), this would raise the cost per case averted to US$1.04 * (0.7/0.5) = US$1.46.

The cost effectiveness of spring protection and WaterGuard use thus appear to be quite similar. However, this conclusion is quite sensitive to the pricing policy of the POU technology. This price is subsidized by the NGO marketing WaterGuard: it covers production and distribution costs but not marketing. It is worth noting that the cost of POU chlorination would be much lower if chlorine were distributed in bulk, without the marketing, packaging, and retail distribution costs that raise its local price. The local cost of a similar chlorine product, bleach for washing, distributed in much larger containers is only 27% of the cost of WaterGuard per unit of chlorine. If the cost of WaterGuard were reduced to 27% * US$1.46 = US$0.39, point of use treatment would be much more cost effective than spring protection at averting diarrhea cases. The decision to use a retail

model for distributing POU treatment, where the cost of packaging far exceeds the cost of the product, reduces the relative cost effectiveness of WaterGuard.

8 A revealed preference estimate of household valuation of cleaner water

We next use the data on household water source choices to recover a revealed preference measure of household valuation of the water quality gains generated by spring protection. We focus on a conditional logit model consistent with the linear random utility framework developed in section 6.1 above. Given a set of characteristics X_{ijt} for individual i and spring j

since collecting water is likely to be a job for relatively unskilled individuals in the household. (For instance, 11% of water collection trips in our sample are made by children under age 12.) Because limited time-income substitution possibilities are frequently encountered empirically (McKean, Johnson, and Walsh 1995), we focus on a range from 25 to 50% of this value of time.

There is potentially substantial heterogeneity in both household's valuation of spring protection as well as in their time costs. We allow the coefficient on these two terms to vary as a function of households' baseline characteristics, and in particular as a function of the number of children of different ages in the household and their health status, by including interactions between these characteristics and both the treatment indicator and the walking distance term.[26]

The conditional logit analysis yields a large, negative and statistically significant effect on the round-trip walking distance to water source (measured in minutes) term, at -0.058 (standard error 0.007, Table 8, regression 1) and a positive statistically significant effect on the treatment (protected) indicator term (0.56, standard error 0.21). Other terms in the regression indicate that streams and rivers are less preferred sources relative to the omitted category (non-program springs), while there is no clear preference among program springs, non-program springs, wells and boreholes.

One concern with the interpretation of this result is possible measurement error in the reported distance walking variable. The correlation across survey rounds in the reported walking distance to the sample spring is only 0.38, so there could be considerable attenuation bias. In addition to simple recall error, the variation in reported walking time may be to actual variation in travel time, depending on the weather, what else the individual is carrying to the source (e.g., carrying a baby, goods for market), and the respondent's health and energy level. To correct for classical measurement error in this term, we would inflate its coefficient to -0.058 / 0.38 = -0.153, and use this correction in valuation calculations below.[27]

[26] Future word will explicitly model this heterogeneity using a mixed logit approach (Train 2003).
[27] In future we will derive the measurement error correction inflator using Monte Carlo simulations.

The ratio of the two main coefficient estimates in this specification implies that one round trip to a protected spring compared to an unprotected spring is valued at (0.56)/(0.153) = 3.7 minutes of walking time. This is a moderately large effect: if household members' time is valued at 50% of the Kenyan average wage, or US$0.71 per day (US$0.0015 per minute), and households make our sample average of 32 water collection trips per week, 52 weeks per year, the total average value to households from protection is (3.7 minutes)*(US$0.0015/minute)*(32 trips/week)*(52 weeks/year) = US$9.05 per year (Table 9, Panel A). At the probably more realistic time value of $0.34 per day, household willingness to pay for spring protection is only US$4.52 per year. This places a range of values on the large increases in usage of the sample spring after protection shown in Table 4. Since there are roughly seven households members on average, this is a valuation of roughly US$1 per capita per year, a non-trivial welfare benefit for households in Kenya.

The two waves of spring protection allow us to assess whether households' valuation is changing through time, but valuations are nearly identical for those households who had had one additional year of spring protection (results not shown).

Spring protection costs about US$1,000 per spring and lasts for at least ten years, with maintenance costs of around US$55 per year, leading to a net present value of US$1480. Assuming households are willing to pay $4.52-$9.02 per year and that 31 households use the typical spring, the net present value of WTP for spring protection is US$1223 to 2447. Thus spring protection appears to be socially beneficial for most of the range of reasonable willingness to pay estimates, and the returns would be especially large if the number of households using a spring is large.

Combining the results from Tables 7 and 8 sheds light on the WTP to avert one episode of child diarrhea. The average number of averted diarrhea cases due to spring protection is (-0.047 cases / child-week) * (2.2 children / household) * (52 weeks / year) = -5.4 diarrhea cases per year. Using our spring protection WTP range of US$4.52-9.05 per year, this translates into US$0.83-1.67 per case of diarrhea averted under the assumption that all of spring protection's value works through

child health gains. These valuation figures lie below the estimated cost per diarrhea case averted calculated above, and below estimated costs per case of diarrhea averted with several other common interventions (Varley et al. 1998). Thus it is unclear if households would be willing to pay for spring protection and other water, hygiene and sanitation interventions if their benefits come mainly in the form of reduced child diarrhea, although if there are also benefits in terms of child mortality presumably the cost-benefit calculation would change.

To the extent that spring protection yields other benefits as well, this estimate could be seen as an upper bound on the willingness to pay to avoid diarrhea. However, while the non-health benefits of spring protection – in terms of water appearance, taste or ease of water collection – could theoretically contribute to the willingness to pay for spring protection, we find no evidence that other potential benefits have a significant effect on WTP. The inclusion of terms for measured *E. Coli* contamination at a subset of alternative water sources, as well as the household's perception of water quality at each source, reduces the coefficient estimate on the spring protection treatment indicator near zero (Table 8, regression 3). Thus nearly all the valuation appears to come from the water quality benefits rather than other amenities associated with spring protection.

Theoretically, households with young children should have both greater time costs of walking to get water (due to the demands of child care or carrying a small child to the source) and greater benefits of clean water, since the epidemiological evidence suggests the largest health gains of improved water quality are experienced by young children. Empirically, we find that households with more children under age 5 at baseline find additional walking distance to a source to be more costly, and this effect is especially strong for households who had young children with diarrhea at baseline: that effect is large and statistically significant at 95% confidence (Table 8, regression 2.) The interaction of 0-5 year old child diarrhea with the treatment indicator is large and positive, suggesting that households with sick children place greater value on clean water but the effect is not

statistically significant at traditional confidence levels (0.23, standard error 0.17). There is no evidence that the gender of household children affects household water source choices (not shown)

There are strong conceptual reasons to prefer revealed preference valuation estimates, like those above based on actual water source choices, over stated preference estimates when measuring use value. We next construct two distinct stated preference measures and compare them to the above revealed preference results.

The first approach is a stated preference ranking measure. Here, rather than relying on information on actual household trips to various local sources, we instead ask respondents to rank order these potential water source options. This is done sequentially in the survey, with the highest ranked source eliminated from the choice set at each subsequent round. These data are then used in a travel cost model nearly identical to what is used in the revealed preference estimates. The only aspect that is varied in the analysis is then using actual water trip data versus stated preference.

The stated preference ranking willingness to pay estimates are much higher than the revealed preference estimates. The magnitude of the coefficient estimate on distance walking falls to -0.034 while that on spring protection rises to 0.95 (Table 8, regression 4). Using the same attenuation bias correction as above, the WTP for one year of spring protection, using time valued at 50% (25%) of the average Kenyan wage, is US$26.19 (US$13.09). This is almost exactly three times greater than the value using the analogous revealed preference approach (Table 9, Panel B).

Comparing the analogous columns in Table 8 (regressions 2 and 5) also highlights interesting reporting patterns in the stated preference ranking case. The coefficient estimates on several sources people locally think of as “bad” or unclean (e.g., streams, rivers, lakes, ponds) are far more negative in the stated preference ranking regression than in the revealed preference case, while the coefficient estimate on spring protection is much more positive. Overall, this might give us pause about the validity of the stated preference ranking approach in our case, since social desirability bias appears to be playing a role in some respondents’ reported rankings.

The second stated preference method is contingent valuation. Here households in protected spring communities were asked how much they would be willing to pay to keep their spring protected. The contingent valuation questions were only asked of households in the treatment group, since they have a first-hand sense of what spring protection is worth. In the final wave of the survey, respondents were first asked if they would be willing to pay either 250 or 500 Kenya Shillings, followed by the question that emphasized the expenditure trade-off for their assigned amount (in other words, what goods they would be giving us by spending that much on spring protection), and then were asked if they would be willing to pay the next higher amount also with emphasis on the expenditure trade-off. This closed-end format, offering discrete value choices, has become standard in the contingent valuation literature (Bateman and Willis 1999).The survey question wording was:

> *"Now that you have seen the protected spring, suppose that somehow the spring had been 'split' so that there was free access to an unprotected spring and restricted access to a protected spring, both at the same site. Would you be willing to pay _____ Ksh [price for this household] for one year's access to the protected spring, assuming everyone else would also have to pay this amount too?"*[28]

The main finding is that nearly all households claim to be willing to pay up to US$7.14 for spring protection over one year, and the majority of households are willing to pay twice that amount (US$14.29) even after being walked through the expenditure trade-offs by the enumerator (Table 9, Panel C). The use of the expenditure trade-off prompt does reduce willing to pay substantially (by 10-13 percentage points), indicating how sensitive contingent valuation results are to survey framing.

If we assume spring protection valuations are normally distributed, the distribution that best fits the contingent valuation response data has mean willingness to pay of US$17.47 (standard

---

[28] The wording of the question emphasizing expenditure trade-offs was: *"So, just to be sure I understand, you would be willing to give up [say price from name list for this specific household] Ksh of purchases that you currently make in order to have access to the protected part of the spring. 250 Ksh per year is about 20 Ksh every month. That's a little bit less than a half-liter of kerosene or a quarter-kilo of sugar every month. For another reference, a school uniform costs about 500 Ksh. If you had to give up something you would otherwise spend money on, would you still be willing to pay _____ Ksh [price for this household] for access to the protected part of the spring?"* We thank Michael Hanemann for discussions on the phrasing and framing of these questions.

deviation US$12.76). This is considerably more than our best revealed preference estimates of US$4.52-9.05 per year, but lies within the range of stated preference ranking valuations discussed above (US$13.09-26.19). Thus both stated preference methods yield similar estimated valuation for improved source water quality. These findings suggest that stated preference methods exaggerate household willingness to pay for environmental amenities in a rural Kenyan setting, and that the more reliable revealed preference approach yields much more modest valuations.

9 Discussion and conclusion

Spring protection dramatically and quite cheaply improved source water quality in a rural African setting, reducing contamination by three quarters on average. Home water quality gains were considerably smaller. Child diarrhea fell by a marginally significant one fifth. However, source water protection appears to be somewhat less cost effective than point of use treatment, like chlorination at home, at reducing child diarrhea.

An 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, 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.

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. We find moderate household valuation for spring protection, on the order of US$4.52-9.05 per household annually. This translates

into a household willingness to pay US$0.83-1.67 per averted case of child diarrhea. In contrast, stated preference valuation approaches produce estimated WTP estimates much higher than these, and as much as three times higher in some cases.

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. Beyond spring protection, we are using randomized evaluations to investigate the role that point-of-use water technologies as well as water quantity increases can play in achieving safe drinking water, and in particular to determine if these approaches would be most effectively employed as complements or substitutes for source water improvements 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.

Bateman, I. and K. G. Willis, (1999), Valuing Environmental Preferences: The Theory and Practice of the Contingent Valuation Method in the US, EU and Developing Countries. Oxford University Press: New York..

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.

Carson, R.T., N.E. Flores, K. Martin, and J. Wright (1996), Contingent valuation and revealed preference methodologies: Comparing the estimates for quasi-public goods Land Economics,72(1): 80-99.

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.

McFadden, D. (1974), “Conditional logit analysis of qualitative choice behavior”, in P. Zarembka, ed., Frontiers in Econometrics, Academic Press, New York, pp. 105-142.

McKean, J. R., D. Johnson, and R. Walsh, (1995) "Valuing time in travel cost demand analysis: An empirical investigation," Land Economics, 71(1).

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.

Train, K. E. (2003). Discrete Choice Methods with Simulation. Cambridge: Cambridge.

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," WHO 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.

Wooldridge, Jeffrey M. (2002). *Econometric Analysis of Cross-section and Panel Data*, MIT Press, Cambridge.

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: 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.

Table 1: Baseline descriptive statistics (2004 survey)

	Treatment (protected) Mean (s.d.)	Obs.	Comparison Mean (s.d)	Obs.	Treatment – Comparison (s.e)
Panel A: Spring level data					
Ln. *E. coli* MPN (CFU/ 100 ml)	3.90	98	3.79	95	0.11
	(1.95)		(1.97)		(0.28)
Water is high quality (*E. coli* MPN ≤ 1)	0.05	98	0.06	95	-0.01
	(0.22)		(0.24)		(0.03)
Water is moderate quality (*E. coli* MPN 2-100)	0.61	98	0.60	95	0.01
	(0.49)		(0.49)		(0.07)
Water is high or moderate quality (*E. coli* MPN <100)	0.66	98	0.66	95	0.00
	(0.48)		(0.48)		(0.07)
Water is poor quality (*E. coli* MPN 100-1000)	0.23	98	0.26	95	-0.03
	(0.43)		(0.44)		(0.06)
Water is very poor quality (*E. coli* ≥ 1000)	0.10	98	0.07	95	0.03
	(0.30)		(0.26)		(0.04)
Latrine density (fraction of homes with latrines)	0.85	98	0.88	95	-0.02
	(0.16)		(0.15)		(0.02)
Average diarrhea prevention knowledge score	3.06	98	3.19	95	-0.13
	(0.87)		(1.17)		(0.15)
Iron roof density (fraction of compounds with iron roof)	0.70	98	0.68	95	0.03
	(0.21)		(0.23)		(0.03)
Panel B: Household summary statistics					
Ln. *E. coli* MPN (CFU/ 100 ml)	3.22	733	3.33	712	-0.11
	(2.22)		(2.13)		(0.14)
Water is high quality (*E. coli* MPN ≤ 1)	0.15	733	0.12	712	0.04
	(0.36)		(0.32)		(0.02)**
Water is moderate quality (*E. coli* MPN 2-100)	0.57	733	0.62	712	-0.05
	(0.49)		(0.49)		(0.03)
Water is high or moderate quality (*E. coli* MPN <100)	0.73	733	0.74	712	-0.01
	(0.45)		(0.44)		(0.03)
Water is poor quality (*E. coli* MPN 100-1000)	0.20	733	0.19	712	0.01
	(0.40)		(0.39)		(0.03)
Water is very poor quality (*E. coli* ≥ 1000)	0.07	733	0.08	712	-0.01
	(0.25)		(0.26)		(0.01)
Respondent years of education	5.73	731	5.75	717	-0.01
	(3.65)		(3.73)		(0.23)
Children under age 12 in the compound	4.03	736	3.93	719	0.10
	(2.48)		(2.46)		(0.14)
Iron roof indicator	0.70	735	0.68	717	0.03
	(0.46)		(0.47)		(0.03)
Walking distance to closest water source (minutes)	10.23	715	9.58	700	0.66
	(9.99)		(8.77)		(0.65)
Water collection trips per week by household	48.08	733	47.99	716	0.10
	(36.55)		(38.48)		(2.52)

	Treatment (protected) Mean (s.d.)	Obs.	Comparison Mean (s.d)	Obs.	Treatment – Comparison (s.e)
Ever collects drinking water at “assigned” spring indicator	0.82	661	0.80	668	0.02
	(0.38)		(0.40)		(0.03)
Multi source user (uses sources other than assigned spring)	0.45	732	0.44	715	0.00
	(0.50)		(0.50)		(0.04)
Fraction of respondent water trips to “assigned” spring	0.72	655	0.71	663	0.01
	(0.41)		(0.42)		(0.04)
Rates water at the spring “very clean” (rainy season)	0.33	736	0.33	719	0.00
	(0.47)		(0.47)		(0.04)
Rates water at the spring “very clean” (dry season)	0.74	736	0.74	719	-0.00
	(0.44)		(0.44)		(0.03)
Fraction of water trips by those under age 12	0.10	727	0.10	711	-0.00
	(0.20)		(0.20)		(0.01)
Water storage container in home was covered	0.90	673	0.93	656	-0.03
	(0.30)		(0.26)		(0.02)**
Yesterday's drinking water was boiled indicator	0.25	731	0.29	711	-0.03
	(0.43)		(0.45)		(0.02)
Respondent diarrhea prevention knowledge score	3.06	736	3.19	719	-0.13
	(2.14)		(2.26)		(0.15)
Respondent said dirty water causes diarrhea	0.68	736	0.67	719	0.01
	(0.47)		(0.47)		(0.03)
Household compound is clear of debris	0.52	734	0.54	716	-0.01
	(0.50)		(0.50)		(0.03)
Household has soap in the home	0.91	733	0.91	717	0.00
	(0.28)		(0.29)		(0.02)
Panel C: Child demographics and health					
Child age (years) §	1.70	1045	1.72	995	-0.02
	(0.95)		(0.97)		(0.04)
Child gender (=1 if male)	0.52	1045	0.50	995	0.01
	(0.50)		(0.50)		(0.02)
Child had diarrhea in past week indicator	0.23	997	0.20	961	0.03
	(0.42)		(0.40)		(0.02)
Child height (cm)	76.05	868	76.13	835	-0.08
	(11.68)		(12.16)		(0.57)
Child weight (kg)	9.97	862	10.02	810	-0.06
	(3.05)		(3.09)		(0.16)

Notes: The treatment springs were later protected (in 2005). 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 Panel C were reported to be under age 3 at baseline or have been born since then..

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

	Panel A: Dependent variable, Ln(Spring *E. coli* MPN) 2005 (round 1 post-treatment)			Panel B: Dependent variable, Spring water high quality (*E. coli* MPN ≤ 1) 2005 (round 1 post-treatment)		
	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.)	3.86	3.75	0.12	0.09	0.05	0.03
	(2.12)	(1.88)	(0.35)			(0.05)
After protection, mean (s.d.)	2.64	3.79	-1.15	0.18	0.05	0.12
	(2.19)	(1.85)	(0.36)***			(0.06)**
After – Before difference (s.e.)	-1.22	0.05	-1.27	0.09	0.00	0.09
	(0.44)***	(0.19)	(0.48)***	(0.07)	(0.03)	(0.07)
% Change in contamination	-70%	5%	-72%	9%	0%	9%

	Panel C: Dependent variable, Ln(Spring *E. coli* MPN) 2007 (round 3 post-treatment)			Panel D: Dependent variable, Spring water high quality (*E. coli* MPN ≤ 1) 2007 (round 3 post-treatment)		
	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.)	3.82	3.73	0.09	0.05	0.07	-0.01
	(1.95)	(1.94)	(0.29)			(0.04)
After protection, mean (s.d.)	1.82	3.26	-1.44	0.39	0.10	0.29
	(1.99)	(1.83)	(0.28)***			(0.06)***
After – Before difference (s.e.)	-2.00	-0.47	-1.53	0.33	0.03	0.30
	(0.23)***	(0.23)**	(0.33)***	(0.05)***	(0.04)	(0.06)***
% Change in contamination	-86%	-37%	-78%	39%	3%	35%

Notes: N=184 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)			
	(1)	(2)	(3)	(4)
Treatment (protected) indicator	-1.06	-1.07	-1.03	-1.09
	(0.27)***	(0.27)***	(0.23)***	(0.24)***
Baseline ln(Spring water *E. coli* MPN)		0.43	0.97	0.98
		(0.04)***	(0.04)***	(0.05)***
Baseline ln(Spring water *E. coli* MPN)			-0.17	-0.16
* Treatment indicator			(0.12)	(0.13)
Baseline latrine density				-0.20
				(0.61)
Baseline latrine density * Treatment indicator				0.84
				(1.75)
Baseline diarrhea prevention score				-0.04
				(0.07)
Baseline diarrhea prevention score				-0.31
*Treatment indicator				(0.24)
Baseline boiled water yesterday density				0.42
				(0.65)
Baseline boiled water yesterday density				0.88
*Treatment indicator				(1.53)
Baseline mother's years of education density				-0.04
				(0.04)
Baseline mother's years of education density				0.07
*Treatment indicator				(0.14)
Treatment group 1 (phased in early 2005)	-0.27	-0.34	-0.37	-0.30
	(0.30)	(0.20)*	(0.17)**	(0.20)
Treatment group 2 (phased in late 2005)	-0.22	-0.24	-0.27	-0.21
	(0.25)	(0.17)	(0.15)*	(0.18)
R^2	0.19	0.33	0.42	0.45
Observations	726	726	726	726
Mean (s.d.) of dependent variable	3.65 (1.95)	3.65 (1.95)	3.65 (1.95)	3.65 (1.95)

Notes: Estimated using OLS. Huber-White robust standard errors are presented (clustered at the spring level), significantly different than zero at * 90% ** 95% *** 99% confidence. There are 184 spring clusters with data for the four survey rounds (2004, 2005, 2006, 2007). 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.

Time fixed effects are included in all regressions but not reported. When interactions included, baseline variables are interacted with time indicators and treatment group indicators in addition to the treatment indicator. These coefficients not reported.

Baseline iron roof density and the interaction with the treatment indicator are included as additional control variables (not shown in the table).

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

Dependent variable	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.) comparison group in 2006, 2007 surveys, Full sample
Panel A: Water collection and source choice	(1)	(2)	(3)	(4)
Use assigned spring for drinking water indicator	0.09 (0.04)**	0.03 (0.04)	0.20 (0.06)***	0.63 (0.48)
Fraction of trips to assigned spring	0.07 (0.04)*	0.02 (0.04)	0.19 (0.06)***	0.56 (0.47)
Perceive water at assigned spring to be very clean (rainy season)	0.18 (0.03)***	0.17 (0.04)***	0.18 (0.04)***	0.12 (0.33)
Perceive water at assigned spring to be very clean (dry season)	0.09 (0.03)***	0.05 (0.03)*	0.13 (0.05)***	0.51 (0.50)
Self-reported distance to nearest water (min.)	-1.44 (0.45)***	-1.65 (0.51)***	-1.08 (0.79)	8.08 (7.61)
Calculated distance to assigned spring (km)	0.03 (0.03)	0.06 (0.05)	0.01 (0.01)	0.36 (2.52)
Trips made to get water (all uses, members, sources) past week	-2.48 (2.15)	-0.89 (2.39)	-4.52 (3.51)	31.78 (24.42)
Panel B: Water transportation and storage				
Fraction of water trips by those under age 12[(a)]	-0.00 (0.01)	0.00 (0.02)	-0.00 (0.02)	0.09 (0.19)
Water storage container in home covered indicator	-0.00 (0.01)	-0.01 (0.02)	0.01 (0.02)	0.98 (0.15)
Ever treated water with chlorine indicator[(b)]	0.03 (0.03)	0.04 (0.05)	0.02 (0.05)	0.45 (0.50)
Yesterday's drinking water boiled indicator[(c)]	0.04 (0.02)	0.05 (0.03)*	0.01 (0.03)	0.25 (0.44)
Panel C: Complementary sanitation and hygiene behaviors				
Diarrhea prevention knowledge score	0.10 (0.13)	0.19 (0.16)	-0.01 (0.17)	2.65 (2.50)
Respondent says drinking clean water is a way to prevent diarrhea	-0.03 (0.02)	-0.03 (0.03)	-0.03 (0.04)	0.50 (0.50)
Household compound is clear of debris indicator	0.03 (0.03)	0.02 (0.04)	0.04 (0.04)	0.71 (0.45)
Household has soap in the home indicator	-0.01 (0.02)	-0.02 (0.02)	0.01 (0.03)	0.89 (0.31)
Panel D: Spring amenities (recorded by enumerators)				
Spring has "clear" water	0.26 (0.07)***	-	-	0.71 (0.45)
Fence around spring	0.95 (0.03)***	-	-	0.00 (0.00)
Spring has "high" water yield	-0.05 (0.06)	-	-	0.73 (0.45)
Fecal matter around spring	-0.15 (0.06)**	-	-	0.27 (0.44)
Animals around spring	-0.03 (0.06)	-	-	0.12 (0.33)
Trees planted around spring	0.16 (0.05)***	-	-	0.03 (0.16)
Trench for spring water cleared in last month	0.29 (0.11)***	-	-	0.59 (0.49)
Vegetation near spring cleared in last month	0.17 (0.10)*	-	-	0.36 (0.48)
Reported spring maintenance quality (5=excellent, 1=poor)	0.53 (0.14)***	-	-	1.78 (0.92)

Notes: N=1354 households at 184 springs (full sample), 755 of whom are baseline sole source users. Each cell 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. Reported means of the dependent variables are in the comparison group 2006 and 2007 (rounds 2 & 3 post-treatment) surveys. Assigned spring is the spring that we believed households used at baseline based on spring user lists.
(a): Because of changes in survey design, responses to this question are not available for the third (2006) round of data collection.
(b): Because of changes in survey design, responses to this question are not available for the first (2004) round of data collection.
(c): Because of changes in survey design, responses to this question are not available for the fourth (2007) round of data collection.

Table 5: Spring protection household water quality impacts, difference-in-differences (Dependent variable: ln(*E. coli* MPN))

	Panel A: Full sample, 2005			Panel B: Full sample, 2007		
	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.)	3.24	3.26	-0.02	3.19	3.31	-0.12
	(2.27)	(2.14)	(0.17)	(2.22)	(2.13)	(0.15)
After protection, mean (s.d.)	3.09	3.33	-0.24	2.81	3.14	-0.33
	(2.16)	(2.10)	(0.18)	(2.19)	(2.07)	(0.17)**
After – Before difference (s.e.)	-0.15	0.07	-0.22	-0.38	-0.17	-0.21
	(0.20)	(0.11)	(0.23)	(0.14)***	(0.14)	(0.20)
% Change in contamination	-14%	7%	-20%	-32%	-16%	-19%
	Panel C: Sole Source Spring Users, 2005			**Panel D: Sole Source Users, 2007**		
	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.)	3.38	3.29	0.09	3.22	3.40	-0.14
	(2.32)	(2.11)	(0.22)	(2.22)	(2.11)	(0.19)
After protection, mean (s.d.)	3.07	3.46	-0.40	2.83	3.20	-0.37
	(2.10)	(2.04)	(0.21)*	(2.17)	(1.97)	(0.21)*
After – Before difference (s.e.)	0.05	0.18	-0.49	-0.39	-0.20	-0.19
	(0.12)	(0.13)	(0.28)*	(0.19)**	(0.17)	(0.25)
% Change in contamination	5%	20%	-39%	-32%	-18%	-17%
	Panel E: Multi-source Spring Users, 2005			**Panel F: Multi-source Spring Users, 2007**		
	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.)	3.05	3.22	-0.17	3.16	3.20	0.00
	(2.20)	(2.17)	(0.23)	(2.22)	(2.14)	(0.22)
After protection, mean (s.d.)	3.12	3.16	-0.04	2.79	3.07	-0.28
	(2.24)	(2.16)	(0.24)	(2.22)	(2.21)	(0.29)
After – Before difference (s.e.)	-0.02	-0.06	0.13	-0.37	-0.13	-0.24
	(0.13)	(0.15)	(0.30)	(0.22)*	(0.25)	(0.33)
% Change in contamination	-2%	-6%	14%	-31%	-12%	-21%

Notes: N = 1354 households at 184 springs, 755 of which are sole source users. Huber-White robust standard errors are presented (clustered at the 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)		
	(1)	(2)	(3)
Treatment (protected) indicator	-0.25	-0.28	-0.65
	(0.15)*	(0.19)	(0.27)**
Baseline ln(Spring water *E. coli* MPN)	0.07	0.08	0.08
	(0.02)***	(0.02)***	(0.03)***
Baseline multi-source user		-0.28	-0.26
		(0.17)*	(0.17)
Baseline multi-source user * Treatment indicator		0.06	0.08
		(0.25)	(0.26)
Baseline latrine density	-0.83	-0.85	-0.06
	(0.33)**	(0.32)***	(0.59)
Baseline latrine density * Treatment indicator			1.43
			(1.01)
Baseline diarrhea prevention score	-0.02	-0.03	-0.05
	(0.02)	(0.02)	(0.04)
Baseline diarrhea prevention score * Treatment indicator			-0.05
			(0.06)
Baseline boiled water yesterday indicator	0.17	0.16	0.28
	(0.08)**	(0.08)*	(0.16)*
Baseline boiled water yesterday indicator * Treatment indicator			0.50
			(0.28)*
Baseline mother's years of education	0.01	0.01	0.03
	(0.01)	(0.01)	(0.02)
Baseline mother's years of education * Treatment indicator			0.02
			(0.04)
Treatment group 1 (phased in early 2005)	-0.03	-0.15	-0.02
	(0.14)	(0.18)	(0.27)
Treatment group 2 (phased in late 2005)	-0.13	-0.17	-0.21
	(0.12)	(0.16)	(0.28)
R^2	0.03	0.04	0.05
Observations (spring clusters)	4341 (184)	4341 (184)	4341 (184)
Mean (s.d.) of dependent variable in comparison group	3.25 (2.15)	3.25 (2.15)	3.25 (2.15)

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, home has iron roof indicator, iron roof density within spring community. When differential treatment effects are reported in column 3, we also include interactions with all of these control variables and the treatment indicator (not shown in the table). Baseline spring water quality, latrine density, and diarrhea prevention score are de-meaned.
Time fixed effects included in all regressions but not reported. When interactions are included, baseline variables are interacted with time effects and treatment group indicators, in addition to interactions with treatment (protected) indicator. These coefficients not reported in the table.

Table 7: Child health outcomes for children under age three at baseline or born since 2004

	Dependent variable: Diarrhea in past week		Dependent variable: Weight (kg)		Dependent variable: Body mass index, BMI (kg/m^2)	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment (protected) indicator	-0.047*	-0.049*	0.027	0.054	0.105	0.243
	(0.025)	(0.026)	(0.080)	(0.087)	(0.181)	(0.271)
Treatment (protected) indicator, Year 2		0.003		-0.090		-0.683
		(0.021)		(0.096)		(0.860)
Treatment (protected) indicator, Year 3		0.057*		-0.206		2.62
		(0.034)		(0.183)		(2.75)
Child fixed effects?	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.68	0.68	0.93	0.93	0.53	0.54
Child-year observations	7950	7950	6348	6348	6275	6275
Mean (s.d.) of the dep. var. in comparison group	0.23	0.23	11.40	11.40	17.0	17.0
			(3.87)	(3.87)	(2.1)	(2.1)

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. Data from all four survey rounds (2004, 2005, 2006, 2007), sample restricted to children under age three at baseline (in 2004) and children born since 2004. Diarrhea defined as three or more “looser than normal” stools within 24 hours at any time in the past week.

Table 8: Conditional logit estimates of water source choice (2007 surveys)

	---------- Revealed Preference ----------			Stated Preference Ranking	
	(1)	(2)	(3)	(4)	(5)
Distance to water source (minutes walking)	-0.058***	-0.042***	-0.057***	-0.034***	-0.021
	(0.007)	(0.014)	(0.009)	(0.009)	(0.015)
Treatment (protected) indicator	0.56***	0.41	0.09	0.95***	1.43***
	(0.21)	(0.34)	(0.27)	(0.30)	(0.51)
ln (source water E. coli MPN)			-0.13***		
			(0.04)		
Water quality at source perceived to be above average			0.95***		
			(0.21)		
Distance to water source * Children aged 0-5 at baseline		-0.003			-0.010*
		(0.006)			(0.006)
Distance to water source *		-0.016**			0.007
Children aged 0-5 at baseline with diarrhea last week		(0.007)			(0.012)
Treatment indicator * Children aged 0-5 at baseline		0.10			0.07
		(0.13)			(0.16)
Treatment indicator *		0.23			-0.18
Children aged 0-5 at baseline with diarrhea last week		(0.17)			(0.17)
Source type: Borehole/piped	-0.14	-0.11		0.10	0.07
	(0.33)	(0.33)		(0.26)	(0.28)
Source type: Well	-0.20	-0.26		-0.47*	-0.56**
	(0.24)	(0.24)		(0.27)	(0.26)
Source type: Stream/river	-0.52**	-0.48*		-2.16***	-2.12***
	(0.27)	(0.27)		(0.40)	(0.42)
Source type: Lake/pond	-0.13	-0.15		-2.82***	-2.85***
	(0.48)	(0.48)		(0.76)	(0.84)
Source type: Assigned spring	0.92***	0.92***	0.96***	1.04***	1.08***
	(0.18)	(0.18)	(0.20)	(0.23)	(0.24)
Log pseudo-likelihood	-110.17	-103.50	-75.36	-115.18	-107.15
Number of households	453	429	334	491	464
Number of observations	53445	51006	29460	2151	2042

Notes: Disturbance terms are clustered by spring, from a conditional logit model (grouped by trip or choice situation). Significantly different than zero at * 90 ** 95 *** 99% confidence. In columns 1-4 each observation represents a unique household-water source pair in a given water collection trip. In columns 5-6, each observation represents a household-water source pair from a series of questions in which the respondent is asked to choose their favorite source from among alternatives. The data are from the final round of household surveys (2007). The dependent variable is an indicator equaling 1 if the household chose the source represented in that household-water source pair in that collection trip. The omitted water source category is “spring”. Assigned spring is the spring that we believed households used at baseline based on spring user lists. In columns 3 and 5, additional controls are included for children aged 5-12 at baseline, and the distance to water source term, directly and interacted with the treatment indicator term (not shown).

Table 9: Alternative valuations for one year of spring protection (2007 survey)

Panel A: Revealed preference valuation (from Table 8, column 1)	Value of one year of spring protection (US$)	Value of one diarrhea case averted (US$)	
Assume value of time is 50% Kenyan worker average wage	9.05	1.67	
Assume value of time is 25% Kenyan worker average wage	4.52	0.83	
Panel B: Stated preference ranking valuation (from Table 8, column 4)			
Assume value of time is 50% Kenyan worker average wage	26.19	4.83	
Assume value of time is 25% Kenyan worker average wage	13.09	2.41	
Panel C: Contingent Valuation	Full Round	Final Wave	Final Wave – emphasizing expenditure trade-offs
Proportion willing to pay this for spring protection:			
US$3.57 (250 Kenya Shillings)	0.94 (308)	0.90 (98)	0.80 (98)
US$7.14 (500 Kenya Shillings)	0.90 (316)	0.92 (106)	0.79 (204)
US$14.29 (1000 Kenya Shillings)	-	-	0.60 (204)

Notes: The results in Panels A and B all correct for attenuation bias in the coefficient estimate on distance walking to water source, assuming the standard classical measurement error formula (the correlation between reported distance walking to the sample spring across survey rounds is 0.38.)

Number of observations in parentheses in Panel C. The contingent valuation questions were only asked of households in the treatment group, since they have a first-hand sense of what spring protection is worth. In the final wave of the survey, respondents were first asked if they would be willing to pay either 250 or 500 Kenya Shillings, followed by the question that emphasized the expenditure trade-off for their assigned amount, and then were asked if they would be willing to pay the next higher amounts also with emphasis on the expenditure trade-off.

The wording of the basic survey question was: *Now that you have seen the protected spring, suppose that somehow the spring had been "split" so that there was free access to an unprotected spring and restricted access to a protected spring, both at the same site. Would you be willing to pay ______________ Ksh [say price from name list for this specific household] for one year's access to the protected spring, assuming everyone else would also have to pay this amount too? |_____| (1=Yes, 2=No, 99=DK)*

The wording of the survey question that emphasized the expenditure trade-off was: *So, just to be sure I understand, you would be willing to give up [say price from name list for this specific household] Ksh of purchases that you currently make in order to have access to the protected part of the spring. 250 Ksh per year is about 20 Ksh every month. That's a little bit less than a half-liter of kerosene or a quarter-kilo of sugar every month. For another reference, a school uniform costs about 500 Ksh. If you had to give up something you would otherwise spend money on, would you still be willing to pay ______________ Ksh [say price from name list for this specific household] for access to the protected part of the spring? |_____| (1=Yes, 2=No, 99=DK)*

Appendix Table 1: The elasticity of household water quality with respect to spring water quality

	Dependent variable: ln(Home water *E. coli* MPN)		
	Full sample	Sole-source users	Sole-source users
	OLS	OLS	IV
	(1)	(2)	(3)
ln (Spring water *E. coli* MPN)	0.22***	0.23***	0.66**
	(0.02)	(0.03)	(0.31)
Latrine density	-0.72**	-1.14**	-1.18*
	(0.35)	(0.51)	(0.64)
Diarrhea prevention knowledge score	-0.010	-0.046*	-0.046
	(0.021)	(0.028)	(0.032)
Baseline boiled water yesterday indicator	0.111	0.135	0.147
	(0.095)	(0.111)	(0.126)
Baseline mother's years of education			
District-wave (season) fixed effects	Yes	Yes	Yes
R^2	0.06	0.08	--
Observations (spring clusters)	3282 (174)	1803 (159)	1803 (159)
Mean (s.d.) of dep. var. in comparison group	3.09 (2.26)	3.22 (2.14)	3.22 (2.14)

Notes: 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 in columns 1-3 are: number of children in home compound, iron roof indictor and iron roof density in the spring community. Time and treatment group fixed effects are also included in columns 1-3. The instrumental variable in column 3 is the treatment (protection) indicator. The results are based on the baseline household survey (2004) and the first follow-up survey (2005), to ensure that the sole source user definition is relevant.

Figure 1: Map of study region

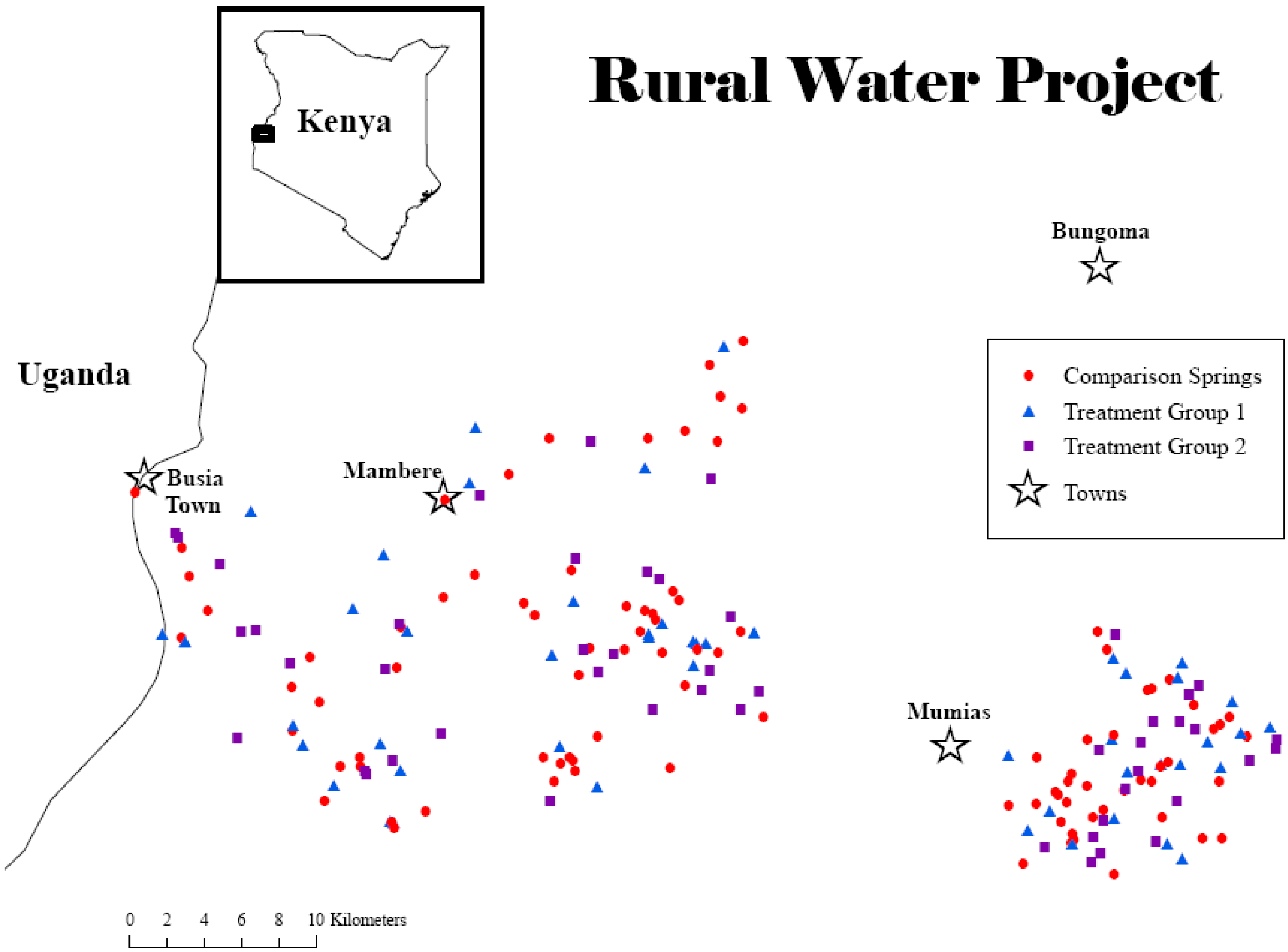


Figure 2: Timeline of Rural Water Project (RWP) Activities 2004-2007

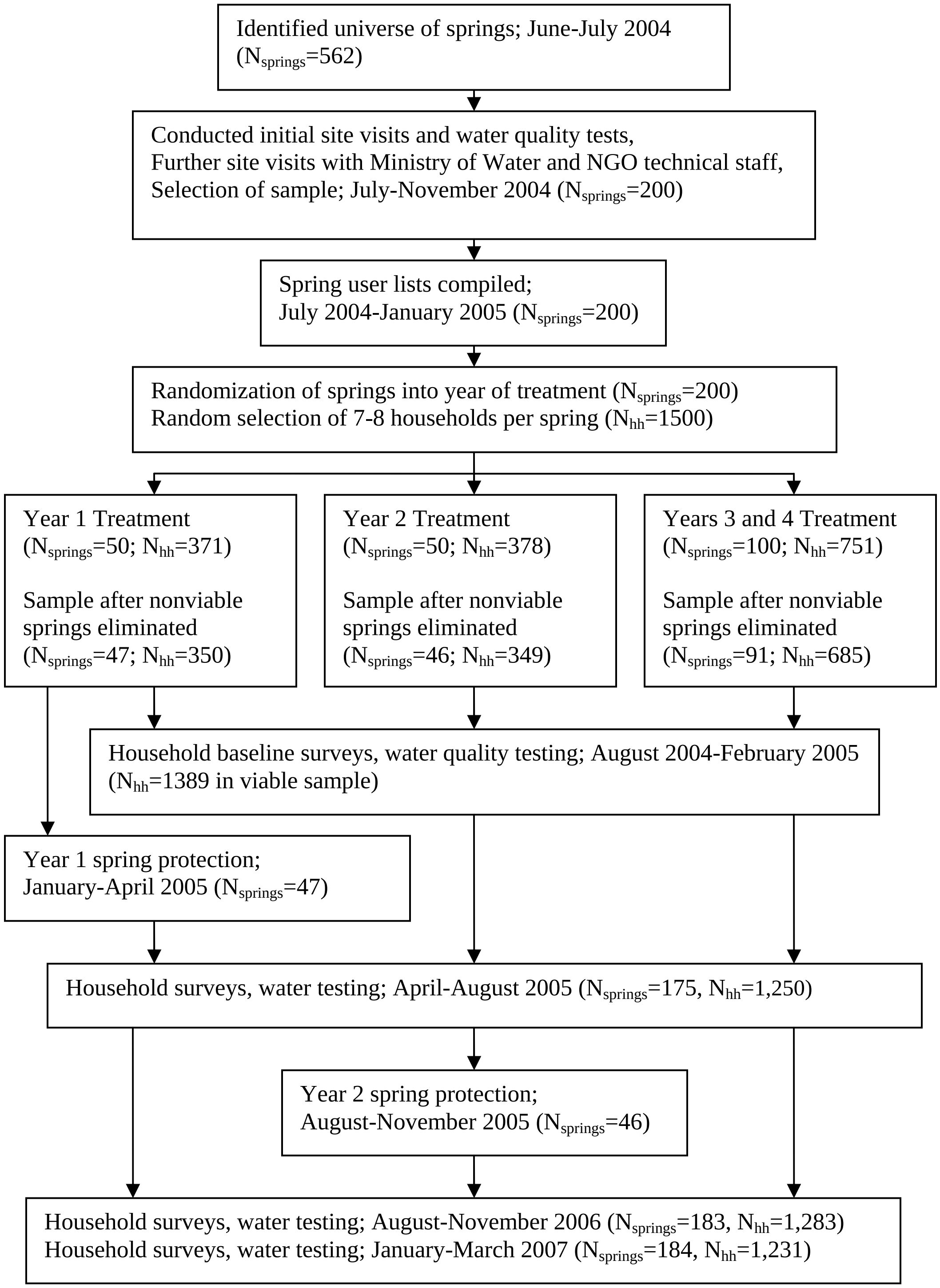
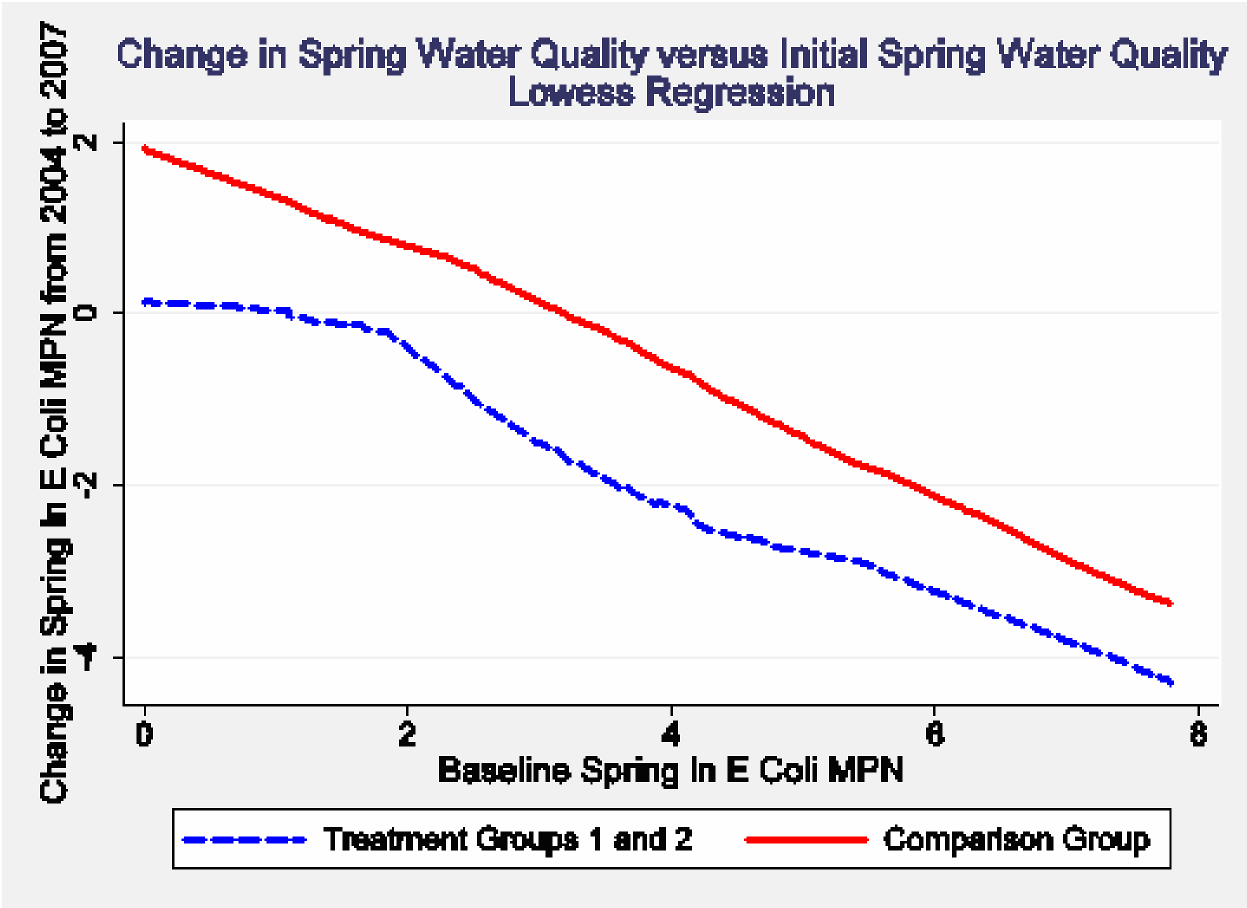


Figure 3: Change in water contamination from 2004 to 2007 versus baseline (2004) water contamination



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.