Spring Cleaning:
Rural Water Impacts, Valuation, and Institutions*

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: February 2008

Abstract: We study a randomly phased-in spring protection program to understand the health impacts and valuation of clean water, and the potential effects of alternative property rights institutions in rural Kenya. Spring protection leads to large improvements in source water quality as measured by the fecal indicator bacteria *E. coli*. This translates to moderate gains in home water quality and to a one quarter fall in reported child diarrhea incidence. Households increase their use of protected springs. Revealed preference estimates of household willingness to pay (WTP) for improved water quality derived from a travel cost analysis are only one-third of stated preference valuations for spring protection. An upper bound on willingness to pay per case of diarrhea averted is only US$0.86-1.72, considerably below figures used in health cost-effectiveness analyses. Simulations based on estimated preferences for cleaner water suggest a social planner would only protect springs with many nearby households. Springs in Kenya are common property resources, limiting private incentives to protect springs to improve water quality, yet strong private property rights would yield lower social welfare than the status quo. However, allowing landowners to charge households for protected spring water only if they continue to provide access to unprotected water is Pareto improving relative to the status quo.

---

* 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, Jeff Berens, Lorenzo Casaburi, Willa Friedman, Anne Healy, Jonas Hjort, Jie Ma, Owen Ozier, Camille Pannu, Changcheng Song, 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, Esther Duflo, Liran Einav, Andrew Foster, Michael Greenstone, Avner Greif, Michael Hanemann, Danson Irungu, Ethan Ligon, Steve Luby, Enrico Moretti, Kara Nelson, Aviv Nevo, Sheila Olmstead, Judy Peterson, Rob Quick, Mark Rosenzweig, Elisabeth Sadoulet, Sandra Spence, Duncan Thomas, Ken Train, Chris Udry, and numerous seminar participants have provided helpful comments. All errors are our own.
-- Corresponding author: Edward Miguel (emiguel@econ.berkeley.edu).

1. Introduction

This paper evaluates the impact of source water quality improvements achieved via spring protection, estimates the valuation that households place on these improvements, and assesses the welfare impacts of alternative water property rights institutions.[1] 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. In western Kenya, where our study site is located, about 90% of rural households have access to a spring (DHS 2003).

Using a randomized evaluation approach, in which protection is phased-in to springs over time, we first estimate spring protection impacts on source water quality, household water quality, and child health. The health benefits of cleaner water might seem obvious; diarrhea accounts for 20% of deaths of children under five each year (Bryce *et al.* 2005), and microbiologically unclean water can provide a route for the fecal-oral transmission of diarrhea. Yet many experts are skeptical about the impact of improved source water quality on health, arguing that sanitation and hygiene may be more important (Esrey 1996), that recontamination of water in transport and storage may vitiate many of the benefits of improved source water quality (Fewtrell *et al*. 2004), or that diarrhea is affected more by the quantity of water available for washing than by the quality of drinking water (Curtis, Carincross, and Yonli 2000). The randomized evaluation allows us to isolate the impact of a single intervention affecting the quality but not quantity of water, and 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.

We find that spring protection greatly improves water quality at the source, reducing fecal contamination by 66%. Spring protection is also moderately effective at improving household water quality, reducing contamination by 23%. The incomplete pass through from spring-level water gains

[1] The current study is one component of a larger project also examining point-of-use and water quantity interventions, which together may provide guidance on priorities in the rural water sector.

into the home is due both to households obtaining water from multiple water sources and to partial recontamination of water in transport and storage. The dampened home water gains are not due to crowding out of other water treatment measures (such as boiling water or chlorination) nor does improved sanitation coverage or hygiene knowledge appear to allow households to better translate source water quality gains into better household water.

Spring protection also improved child health: diarrhea among young children in treatment households falls by 4.6 percentage points, or one quarter on a base diarrhea prevalence of approximately 20 percent.

The second part of the paper contributes to the environmental economics literature on the valuation of environmental amenities. In our study area, most households choose from multiple local alternative water sources. The intervention we study generates exogenous variation in their relative desirability, and we explore how household water source choices and other behaviors respond to water quality improvements. A discrete choice model, in which households trade-off water quality versus walking distance to the source, generates revealed preference estimates of household willingness-to-pay (WTP) for better water quality.[2]

The average estimated WTP for spring protection is equivalent to 12.7 workdays, or approximately US$4.52-9.05 per household per year using wage rates of 25 to 50 percent of the Kenyan average. This suggests an upper bound of US$0.86-1.72 on households' willingness to pay to avert one child diarrhea episode. Under the stronger assumption that households' valuations are driven by reduced child diarrhea mortality, the value of averting one child diarrhea death ranges from

[2] Whittington, Mu, and Roche (1990) and Mu, Whittington, and Briscoe (1990) each study water source choice in rural African settings. 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. Yet the shortcomings of CV and other stated preference approaches to measuring the value of non-market goods are well-known (Diamond and Hausman 1994). Choe (1996) compares willingness to pay for reduced river and lake pollution in an urban Philippines setting, using both travel cost and contingent valuation methods, and finds that both are low and quite similar; Choe's sample consists of households with piped connections, limiting its generality to most rural areas. Two other papers have compared averting (or defensive) expenditure data to stated willingness to pay (Griffin *et al.* 1995 and Rosado *et al.* 2006 in India and Brazil, respectively), though neither exploits experimental variation in water quality as we do.

US$741 to US$1,482, substantially below the values typically used in health cost effectiveness analysis.

We contrast our revealed preference WTP for spring protection, which exploits experimental variation in water source characteristics, with two different stated preference methodologies – stated ranking of alternative water sources, and contingent valuation (see Carson *et al*. 1996, Whitehead 2006). Most WTP estimates rely on stated preference data, which is relatively cheap to collect, yet few stated preference estimates have been validated against reliable revealed preference benchmarks: revealed preference data is rarely available in less developed countries, and those studies that do exist are typically prone to omitted variable bias critiques. We find that stated preference approaches generate much higher WTP estimates than our revealed preference approach, by a factor of three, casting doubt on the reliability of stated preference methods in this setting.

Our third set of results simulates the impact of alternative policies and institutions in the water sector. Many have noted the possible tradeoff between property rights institutions that create efficient investment incentives, and those that achieve other objectives, such as insurance or equity. This debate has been particularly important in rural Africa where communal land tenure systems may distort land use decisions (as in Goldstein and Udry 2005), and where relatives' and community members' claims on others' earnings may inhibit investment (Platteau 2000). In many countries, both law and custom make water sources common property resources, leaving little private incentive to invest in water improvements. Some have argued that strengthening private property rights in water could lead to greater agricultural productivity (Feder et al. 1988), local public goods provision (Hoy and Jimenez 2006), and improved water quality in urban settings (Galiani et al. 2005).

Allowing land owners the right to charge for water would increase their incentives to invest in infrastructure that could deliver cleaner water, but positive water prices would also create static distortions. We show that a social planner would only protect springs used with a relatively large number of household users. We then conduct counterfactual policy simulations under alternative property rights institutions, where estimated household water demand is derived from the revealed preference valuations. We first find that a system of “pure” private property rights yields lower social welfare than the existing communal rights, since landowners charge high positive prices but protect few springs. Yet a more complex “conditional” private property rights system in which land owners can charge for protected spring water only if they also continue to provide free access to unprotected water is Pareto improving relative to the status quo.

Section 2 of this paper describes the intervention and data, and presents summary statistics. Section 3 discusses the impacts of spring protection on water quality and the child health results. Section 4 discusses the impact of spring protection on water source choice and behavior and presents estimates of willingness to pay for clean water. Section 5 discusses the provision of clean water under alternative institutions and policies, and the final section concludes.

2. Rural Water Project (RWP) overview and data

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

2. 1 Spring protection in western Kenya

Naturally occurring springs are an important source of drinking water in rural western Kenya, where the region’s topography frequently allows ground water to come to the surface. . Approximately 43 percent of rural western Kenyan households use springs for drinking water, and over 90 percent have access to springs (DHS 2003). Our respondents report that springs are the main source of water in this area: over 70% of all water collection trips are to springs (either unprotected or protected). The next most common source are wells (at 13%), followed by smaller numbers of water collection trips to boreholes (7%), rivers/streams (4%), lakes, ponds, and other sources. In all 81% of water collection trips are to sources the respondents used for drinking water in the last week. The area of Kenya in which our study site is located is poor – the daily agricultural wage ranges from US$1 - US$2 per day depending on the task – and few households have access to improved water services.

Property rights to land and other natural resources remain a combination of traditional customary law and formal legal statutes in Kenya (Mumma 2005). Custom requires that private landowners allow public access to water sources on their land, and under Kenyan law local authorities can, "where, in the opinion of the Authority the public interest would be best served" order owners of water resources to make water available "to any applicant so long as the water use of the owner of the works is not adversely affected."[3] In practice landowners were expected to make spring water available to neighbors for free. Landowners therefore have very weak incentives to improve a water source, since they are unable to recoup the costs of such an investment via the collection of user fees. There is no elected local government, and collective action problems mean that investments in valuable local public goods, including water points, often fail to occur.[4] When it occurs, spring protection is generally undertaken by outside donors or the central government, often in conjunction with user groups set up to collect maintenance funds, though they too lack authority to exclude free-riders.

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

---

[3] See the 9/28/2007 Kenyan Gazette (Supplement No. 92, Legislative Supplement No. 52, Legal Notice No. 171).
[4] See Miguel and Gugerty (2005) for an analysis of the determinants of local public good provision in rural Kenya.

graves). From the remaining 562 suitable springs, 200 were randomly selected (using a computer random number generator) to receive protection (see Figure 1).

The NGO planned for the water quality improvement intervention to be phased in over four years due to their financial and administrative constraints. Figure 2 summarizes the project timeline. Although all springs will eventually receive protection, for our analysis the springs protected in round 1 (January-April 2005) and round 2 (August-November 2005) are called the treatment springs and those protected later are the comparison group. To determine the order of treatment, springs were first stratified on the basis of geographic region and baseline water quality (this data is described below), and then randomly assigned (using a computer random number generator) to groups.

Several springs were unexpectedly found to be unsuitable for protection after the baseline data collection and randomization had already occurred, when more detailed technical studies were undertaken. These springs, which were found in both the treatment and comparison groups, were dropped from the sample, leaving 184 viable springs. Identification of the unsuitable springs is not 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.

A representative sample of households that regularly use each sample spring was also selected at baseline. Survey enumerators interviewed users at each spring, asking their names as well as the names of other household users. Enumerators elicited additional information on spring users from the three to four households located nearest to the spring. Households that were named at least twice among all interviewed subjects were designated as "spring users". The total number of household spring users varied widely, from eight to 59 with a mean of 31. Seven to eight households per spring were then selected (again using a computer random number generator) from this spring user list for the household sample used in this paper. In subsequent surveys, over 98% of this spring users sample was later found to actually use the spring at least sometimes, but the few baseline non-user households were nonetheless retained in the analysis.

The spring user list is also quite representative of all households living near sample springs. In a February 2007 census of all households living within roughly a 10 minute walk of seven sample springs, we found that 92% of these nearby households were included on the original spring users lists. Spring user list households are less representative, however, for households living more than 10 minutes away from sample springs.[5]

Baseline water data was then collected at all 200 sample springs and a survey of local environmental conditions carried out (January-October 2004). Water quality in household drinking water containers was also tested in local labs, and household data on demographic characteristics, health, anthropometrics, and water use choices collected, as described further below. To address concerns about seasonal variation in water quality and disease burden, all springs were stratified geographically and by treatment group and then randomly assigned to an activity "wave," and all project activities were conducted by wave.

The NGO proceeded with community mobilization meetings after baseline data collection, and then contracted local masons to carry out protection at treatment springs in 2005. Permission for protection was also received from the spring landowner in all but two cases. The NGO requested that each community raise a modest initial contribution of 10% of the project cost, mainly in the form of manual labor and construction materials, and this was successful at all springs. The total cost of protection, including these supplies and estimated labor costs, ranges between US$830 and US$1070, depending mainly on spring size and complexity. A committee of spring users responsible for maintenance was also selected by community members at the initial meeting.

A first follow-up round of water quality testing at the spring and in homes, spring environment surveys, and household surveys was completed three to four months after the first round of spring protection (April-August 2005). The second round of spring protection was performed in

---

[5] In ongoing field work we are collecting information on the proportion of households living 11 to 20 minutes from the source that were baseline users.

August-November 2005, and the second follow-up survey collected one year later (August-November 2006). The third follow-up survey round took place five months later, from January to March 2007. In total there are 184 springs and 1,354 households with baseline data and at least one survey follow-up round, and this is the main analysis sample. In only 10 springs (of 184) did treatment assignment differ from actual treatment (because landowners refused to allow protection or when the government independently protected comparison springs, for example); these springs are retained in the sample, allowing for an intention-to-treat analysis throughout.

2.2 Data collection procedures

Water quality data

Water samples were collected 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 laboratories for same day analysis. The labs use Colilert, a method which provides an easy-to-use, error-resistant test for *E. coli*, an indicator bacteria present in fecal matter.[7, 8] A continuous quantitative measure of fecal contamination is available after 18-24 hours of incubation. Quality control procedures used to ensure the validity of the water testing procedures included weekly positive and negative controls, and duplicate samples (blind to the analyst), as well as monthly inter-laboratory controls. As discussed

---

[6] At springs, the protocol is as follows: the cap of a 250 ml bottle is removed aseptically. Samples are taken from the middle of standing water and the bottle is dragged through the water so the sample is taken from several locations at unprotected springs, while bottles are filled from the water outflow pipe at protected springs. About one inch of space is left at the top of full bottles. The cap is replaced aseptically. In homes, 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).

[7] Our lab procedures were adapted from Environmental Protection Agency Colilert Quantitray 2000 Standard Operating Procedures.

[8] 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 water (Kay *et al*. 1994), but such functions are location-specific because fecal matter pathogens vary over space and time. In a district near our study site, a U.S. Centers for Disease Control project finds that the most common bacterial pathogens are Shigella and non-typhoidal Salmonella.

below, there appears to be mean reversion over time in water contamination, consistent with both some degree of measurement error and natural intertemporal variation.[9, 10]

Household survey data

The target household survey respondent was the mother of the youngest child living in the home compound (where extended families often reside together), or another woman with child care responsibilities if the mother of the youngest child was unavailable. The respondent is asked about the health of all children under age five living in the compound, including recent diarrhea episodes.

The household survey also gathered baseline information about hygiene behaviors and latrine use, as well as the frequency of water boiling, home water chlorination and water collection choices. Respondents were asked to give their opinion on methods to prevent diarrhea; they were not given options to choose from, but were prompted three times and their responses recorded. This information was used to construct a baseline "diarrhea prevention knowledge score", namely, the number of correct responses provided.[11] Respondents volunteered three correct preventative activities on average. There is moderate knowledge of water's role: 50% of respondents named avoiding contaminated water (or some variant of this answer) as a way to reduce diarrhea.

The definition of diarrhea 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

---

[9] There are several 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 an estimated 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 tested samples less representative of the original source. Third, sampling variation is an issue given the small size of the collection bottle (at 250 ml).

[10] 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). Extended holding time increases the noise in the *E. coli* estimate, but there is no clear direction of bias as bacteria both grow and die prior to incubation.

[11] The set of plausible answers include "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". We reviewed all responses other than those listed here and categorized them as valid or invalid.

questionnaire does not attempt to differentiate between acute diarrhea (an episode lasting less than 14 days) and persistent diarrhea (more than 14 days), but identifies dysentery by asking about blood in stool. Enumerators used a board and tape measure to measure the height of children older than two years of age, and digital scales for weight. The height of children under two was measured as their recumbent length using a measuring board, and a digital infant scale measured their weight.

2.3 Sample Attrition

We successfully interviewed 90% of the baseline household sample in the first follow-up survey round, 89% in the second follow-up, and 92% in the third. We have data from all four survey rounds for 79.5% of baseline households and for 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 estimated coefficient on the treatment indicator is -0.03 (p-value=0.7), and this result is robust to including further explanatory variables as controls (not shown).

The baseline characteristics of households lost over time are typically statistically indistinguishable from those that remain in the sample. Better-off households, like those with iron roofs, are not more likely to attrit, nor are households with better baseline household water quality or hygiene knowledge (not shown). Any sample attrition bias appears likely to be small.

2.4 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 statistics for all springs and households with baseline data (collected prior to randomization into treatment groups) even if they are dropped from the analysis because the spring was later found unsuitable for protection, although results are unchanged with the slightly smaller main sample (not shown).

The water quality measure, *E. coli* most probable number (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 < 126 (in geometric mean over at least five tests).[13] To be conservative, we consider water with counts between 1 and 100 "moderate quality". We rarely observe high quality samples in our data, which is not surprising as spring water is neither in a sterile environment nor has residual chlorine (as treated piped drinking water does). We divide the remaining values of *E. coli* CFU/100 ml > 100 somewhat arbitrarily into "poor quality" (between 100 and 1000) and "very poor quality" (greater than 1000).

There is no statistically significant difference between baseline water quality at treatment versus comparison springs (Table 1, Panel A), which implies that the randomization created comparable groups. Most spring water in our sample is of moderate quality, and only about 5-6% of samples from unprotected springs meet the stringent U.S. EPA drinking water standards, while over a third of samples are poor or very poor quality.[14] Household water is somewhat more likely to be high quality prior to spring protection in the treatment group (and the difference in means is significant at 95% confidence, though relatively small), but there is no statistically significant difference in the proportion of moderate or poor quality water samples (Panel B).

Household water quality is somewhat better than spring water quality on average at baseline: the average difference in log *E. coli* is 0.52 (s.d. 2.64; results not shown). This likely occurs for at least two reasons. First, many households collect water from sources other than the sample spring: only half of the household sample gets all their drinking water from their local sample spring at

---

[12] In the laboratory test results, the *E. coli* MPN CFU can take values from <1 to >2419. We ignore censoring and treat values of <1 as equal to one and values of >2419 as 2419. In practice, there are very few censored observations.
[13] The EPA website has details: http://www.epa.gov/waterscience/beaches/local/statrept.pdf (accessed 11/22/2007).
[14] Previous research in Nigeria indicates 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 dry and rainy seasons (Huttly *et al.* 1987). To account for such variation, we include seasonal controls in the analysis.

baseline, and overall nearly one third of water collection trips are to other sources. Second, some households use point-of-use (POU) water treatment at home. Nearly 25% of households report boiling their drinking water at baseline[15], However, the correlation between household water contamination and self-reported water boiling is low, raising the possibility of social desirability reporting bias. In the first follow-up (2005) survey, 28% of households reported chlorinating their water at least once in the last six months; these chlorination levels are higher than usually observed because the government distributed free chlorine tablets in part of our study region following a 2005 cholera outbreak.

Water quality tests were also collected at the two main alternative sources near each sample spring during the third follow-up survey (in 2007). Protected springs have the least contaminated water, with average ln *E. Coli* MPN/100 ml = 2.3. Unprotected springs and boreholes are the next best sources, with average contamination levels of 3.6 and 4.1, respectively. Shallow wells have higher average contamination at 5.2, followed by rivers/streams and lakes/ponds, which come in with contamination over 6. Local residents' perceptions of the relative water quality of these source types line up closely with the objective contamination measures: the proportion of respondents stating that a source is "very" or "somewhat clean" is highest for protected springs, the cleanest source, at 88%, followed by boreholes and unprotected springs (at 78% and 72%, respectively), and shallow wells (66%), while lakes/ponds (31%) and streams/rivers (14%) are widely viewed as unclean.

Most other household and child characteristics are similar across the treatment and comparison groups (Table 1, Panels B and C), further evidence that the randomization was successful. Average mother's education is six years, which is less than primary school completion. Approximately four children under age 12 reside in the average compound. Water and sanitation

[15] This is distinct from boiling water to make tea. It would be possible to drink only tea, and thus effectively drink only boiled water, but we do not find evidence of this coping strategy. Nearly all adults report drinking unboiled water on the day surveyed and, most importantly, young children are commonly given water to drink directly from the household storage container, not exclusively boiled water.

access is fairly high compared to many other 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 (the median one-way distance is 5 minutes). There are similarly no significant differences across the treatment and comparison groups in terms of the diarrhea prevention knowledge score, water boiling behavior, compound cleanliness, or presence of soap. However, 90% of treatment versus 93% of comparison households cover their drinking water containers and this difference is significant at 95% confidence.

Summary statistics for the subset of children under age three at baseline (Table 1, Panel C) indicates that children are comparable across treatment and comparison groups in terms of health and nutritional status. A fairly high 20% 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 (which is both curative and preventative for diarrhea) across the two groups (results not reported).

3. Spring protection impacts on water quality and health

This section discusses the estimation strategy and presents the impacts of spring protection on water quality and child health and nutrition.

3.1 Estimation strategy

Equation 1 illustrates an intention-to-treat (ITT) estimator using linear regression with spring data.

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

W_{it}^{SP} is the water quality measure for spring i at time t ($t \in \{0, 1, 2, 3\}$ for the four survey rounds) and T_{it} is a treatment indicator that takes on a value of one after spring protection assignment, (i.e. for treatment group 1 in all follow-up survey rounds and for treatment group 2 in the second and third follow-ups.) X_i^{SP} are baseline spring and community characteristics (e.g., baseline contamination) and ε_{it} is a white noise disturbance term which is allowed to be correlated across survey rounds for the same spring. Random assignment implies that β_1 is an unbiased estimate of the reduced-form ITT effect of spring protection. In some specifications we explore differential effects as a function of baseline characteristics, captured in the vector β_3. Survey round and wave fixed effects α_t are also included to control for any time-varying factors affecting all groups. Estimates of the average treatment effect on the treated (TOT) in a two-stage procedure (Angrist, Imbens, and Rubin 1996) are very similar to the ITT estimates since assignment differed from actual treatment for few springs.

3.2 Impact of treatment on spring water quality results

Spring protection dramatically reduces contamination of source water with fecal matter. The average reduction in ln *E. coli* across all four rounds of data, is -1.08, corresponding to roughly a 66% reduction (Table 2, regression 1). These estimated effects are robust to including controls for baseline contamination (regression 2). Protection does not lead to a significantly larger proportional reduction in water contamination where initial contamination was highest (regression 3). Figure 3 is a non-parametric representation of the data that shows some gains are experienced at nearly all treatment springs. The downward slope of these plots is consistent with mean reversion, likely reflecting measurement error in the water quality figures. The correlation in measured water contamination levels across survey rounds is moderate, at 0.29. There is no statistically significant evidence of differential treatment effects by baseline respondent hygiene knowledge (the average among local spring users), average local sanitation (latrine) coverage, or education (regression 4).

3.3 Home water quality impacts

Relying again on the randomized design, we estimate a regression analogous to equation 1 to estimate the impact of spring protection on home water quality, again measured in ln *E. coli* MPN. We control for baseline household characteristics in some specifications including sanitation access, respondent's diarrhea knowledge, water boiling (the leading POU water treatment strategy in our study area), an iron roof indicator, years of education, and the number of children under age 12 at baseline. We also allow for differential treatment effects as a function of these characteristics. Regression disturbance terms are clustered at the spring level in these regressions, since households using the same spring could have correlated outcomes: they share common water sources and the local sanitation environment, and may have kinship ties.[16]

For "sole source" households that were already only spring water users in the pre-treatment period, home water quality should be unambiguously better after treatment since they still rely mainly on the spring and its quality improves after protection. Interpretation is more complicated for baseline "multi-source" water users in our data, who were roughly on the margin between using the sample spring and some other source. For these households, improved spring water will be combined in the home with water of unknown quality from other sources, and this together with endogenous source choice could cause home drinking water quality to increase or decrease after protection. Theoretically better spring water quality could induce a household to switch from a distant but high quality alternative to the closer but lower quality spring. This could be optimal because households are trading off water quality versus collection time: even if household water deteriorates, the household could be made better off by spring protection due to time savings.

The average impact of spring protection on home water quality is smaller than the impacts on source water quality. The overall effect of spring protection on home water quality is moderate (Table 3, regression 1) with slightly larger reductions in contamination for sole-source households (regression 2) though we cannot reject equal treatment effects for sole-source and multi-source users. The average reduction in ln *E. coli* contamination is -0.26, or roughly 23%.

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

We again find no evidence of differential treatment effects as a function of household sanitation, diarrhea prevention knowledge, or mother's education (Table 3, regression 3). Households living in communities with greater latrine coverage do appear to have less contaminated water overall, but this does not differentially impact the spring protection effect. The absence of statistically significant differential spring protection effects as a function of pre-existing sanitation access or hygiene knowledge runs counter to claims that source water quality improvements are much more valuable when these factors are also in place, although the relatively large standard errors on these interaction terms argue for caution in interpretation. Perhaps surprisingly, neither mother's diarrhea prevention knowledge nor education is ever significantly related to observed household water quality.[17] Home water contamination reductions are somewhat smaller for households that report boiling their water, as expected if boiling is a substitute for spring protection (regression 3).

Diarrhea causing pathogens are infectious diseases, so improving one water source could potentially benefit neighbors. We test for spring protection externalities for both spring and household water quality by considering the effect of the number of nearby treated springs (located within 1, 3, or 6 kilometers), controlling for the total of local springs (protected or not). There is no evidence for externalities in spring water quality: the coefficient estimate on the number treated springs within 3 kilometers is small at -0.027 and not statistically significant (standard error 0.090), and similar results hold for springs at other distances (regressions not shown). There is some evidence for positive household water quality externalities – the coefficient on treatment springs located within 6 kilometers is -0.046 (standard error 0.021, not shown) – though this effect is consistent both with epidemiological externalities, and with some households switching to increasingly use protected sources, an important issue in our setting that we discuss further below.

[17] A direct measure of hygiene, respondents' fingertip fecal contamination, is also not significantly related to observed household water quality (results not shown).

3.4 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 2:

$$Y_{ijt} = \alpha_i + \alpha_t + \beta_1 T_{ijt} + X_{ij}'\beta_2 + (T_{ijt} * X_{ij})'\beta_3 + u_{ij} + \varepsilon_{ijt} \qquad (2)$$

where the main dependent variable is diarrhea in the past week. The coefficient estimate, β_1, on the treatment indicator T captures the spring protection effect. An advantage of this experimental design over existing studies, beyond the usual benefits of addressing omitted variable bias, is the ability to avoid measurement error in the key water quality explanatory variable (through use of the treatment indicator). We include child fixed effects (α_i), survey round and month fixed effects (α_t). We also explore heterogeneous treatment effects as a function of child and household characteristics, X_{ij}.

Spring protection leads to statistically significant reductions in diarrhea for children under age 3 at baseline or born since the baseline survey. In the simplest specification taking advantage of the experimental design, diarrhea incidence falls by 4.4 percentage points (standard error 1.2, Table 4, regression 1). In a probit specification including treatment group fixed effects and month of survey effects the impact is similar, at -5.2 percentage points (standard error 1.9, regression 2), and similarly in a linear specification with child fixed effects (-4.4 percentage points, regression 3). In our preferred specification with month and child fixed effects and child gender and age polynomial controls, the point estimate is -4.6 percentage points (standard error 2.3, p-value = 0.06, regression 4). On a comparison group average of 19% of children with diarrhea in the past week, this is a drop of one quarter. We conclude that the moderate reductions in household water contamination caused by spring protection were sufficient to significantly reduce diarrhea incidence.

While the estimated reduction in diarrhea remains negative for boys, the effects are driven mainly by reduced diarrhea among girls (Table 4, regression 5). For girls the estimated reduction is 8.1 percentage points, and this effect is significant at 95% confidence. This finding is surprising since baseline diarrhea rates are similar for boys and girls in our sample, and differential gender impacts are rarely found in the related epidemiology literature, and a solid explanation remains elusive.

Interactions with baseline local sanitation (latrine) coverage, diarrhea prevention knowledge, and education are not significant (regression 6), in line with the lack of additional water quality gains for these households. Effects are similar in the second and third years after protection, and also across baseline sole-source versus multi-source households (not shown). Spring protection effects do not differ significantly by month of year (rainy versus dry season), nor by child age up through age four years, nor in households with more young children (not shown).

Despite reduced diarrhea, there are no statistically significant impacts on either child weight or body mass index in the three follow-up surveys, and estimated impacts are close to zero overall for both girls and boys (Table 4, regressions 7-10). We do not find evidence of differential treatment effects at points along the anthropometric distribution using quantile regression (not shown).

There is suggestive evidence that spring protection produces a small reduction in diarrhea among children ages 5-12 as well – in a specification with child fixed effects and the full set of controls, the point estimate is -0.017 (standard error 0.012), on a base diarrhea rate of 0.05 – but the effect is not significant at traditional confidence levels, nor is there any evidence that spring protection improved anthropometrics or school attendance in this age group (regressions not shown). There is no evidence of diarrhea impacts among adults after spring protection (not shown).

We collected information on infant mortality from our household sample, and also from a somewhat larger sample of households with the assistance of local village elders who were asked to keep a diary of infant births and deaths in their communities. However, given the rarity of child death events and limited sample sizes, in neither sample is there sufficient statistical power to detect moderate infant mortality treatment effects at traditional confidence levels, although point estimates have the expected negative sign (estimated reduction -6.7 percent, results not shown).

4. Impacts on household behavior and implications for valuing clean water

This section discusses the impacts of spring protection on household behavior, then uses a travel cost model of household water source choice to estimate household willingness to pay (WTP) for cleaner water, and compares these revealed preference estimates to stated preference and contingent valuation approaches.

4. 1 Estimating spring protection impacts on behavior and water source choice

The main behavioral change that resulted from spring protection is an increase in the use of the protected springs for drinking water, while other changes appear to be minor. There were no significant changes in most water transportation and storage behaviors. There is a small shift in self-reported water boiling at home (Table 5, Panel A), though the effect is not statistically significant. Home chlorination of water take-up is limited in our study area and we do not see large shifts in use after spring protection, which is consistent with the view that adoption costs are currently relatively large for this population. There is also no evidence of changes in diarrhea knowledge or in a direct hygiene measure, fecal contamination on respondents' hands[18] (Panel B). Enumerators collected additional information on physical condition and maintenance, and find that protected springs have significantly "clearer" water, better fencing and drainage, and less fecal matter and brush in the vicinity (Panel C). In contrast, there is no effect on observed water yields, confirming that spring protection isolates source water quality impacts rather than affecting quantity.

In contrast, households substantially change their choice of water sources in response to spring protection. We split the data into two subsamples, sole-source users (those who only used the sample spring at baseline) and multi-source users (those who also used other water sources). Use of the protected spring should increase more among multi-source users than sole-source users, since the

---

[18] To measure fingertip contamination, respondents pressed their hands into KF Streptococcal media (agar plates), and the lab isolated *fecal streptococci* bacteria colonies.

latter group has little or no room to increase usage. Assignment to spring protection treatment is strongly positively correlated with use of the sample spring for those households not previously using it exclusively: treated households increase the fraction of water collection trips to their sample spring by 20 percentage points if they used other sources (multi-source users) at baseline (Table 5, Panel D). Underlying this increased 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" during the rainy season, with somewhat smaller effects in the dry season.

There was no effect on the total number of trips made to water sources in the past week. There were small but statistically significant effects of spring protection on the average distance households walked to their main drinking water source (average length was about 8 minutes one-way or 16 minutes round-trip), with an effect of roughly one minute. A possible explanation is that collection times lines at springs are slightly shorter after protection and that respondents mistakenly assign these time savings to reported walking times, although we cannot rule out the possibility that respondents mistakenly report shorter walking times to more frequently used sources.

[Conceptually, the sole source users could be a useful sample for estimating the degree of pass through of source water quality gains to the home, if these households almost exclusively use the sample spring for drinking water in all periods. Random assignment of springs to protection implies that we could avoid both omitted variable bias and also reduce attenuation bias (due to measurement error in water quality) by estimating this correlation in an instrumental variables framework in the sole-source users sample, with assignment to spring protection as the instrument for spring water quality. Unfortunately, water use patterns are not static across the four years of data: in the first follow-up survey round, 74% of comparison group baseline sole source spring users remained sole source users, but only one third remained sole source users in all three follow-up rounds. This "churning" in water use could be due to changes in households' other water options over time (as other sources are improved or deteriorate), or variation in water collection costs due to evolving household composition. Regardless of the cause, baseline sole- and multi-source user status becomes less meaningful over time, making it infeasible to reliably estimate pass-through.[19]

4. 2 A travel cost model of household water source choice

Suppose that in choosing a water source, households trade off the cost versus the benefits. The opportunity cost of time per minute, $C > 0$, is a function of the local market wage, and we assume for simplicity here that it is constant across households. Thus the cost household *i* bears to make an additional water trip to source *j* is CD_i, where D_i is the round trip distance to the source. The water contamination level for water source *j* is denoted $W_j > 0$, where higher values denote more contamination. The function relating water quality to health is $V(W_j)$, $V' < 0$. Any non-health benefits to using a low contamination source (for instance, the ease of collecting water at a protected spring) are also captured in *V*. We found above that spring protection ("T") causes contamination to fall sharply, so $W^T{}_j < W_j$ for spring *j*.

Household *i*'s indirect utility from a single water collection trip to source *j* can be represented as $U_{ij} = V(W_j) – CD_{ij}$. Household *i* chooses source *j* over an alternative source *k* if the benefits of its water quality outweigh any additional travel costs, namely when $\{V(W_j) – V(W_k)\} – C(D_{ij} – D_{ik}) \geq 0$. 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 its choice set.

[19] We explored the pass-through of source water quality improvements to the home using only the first follow-up survey round, when the baseline definitions are most relevant. In a linear regression of home water quality on spring water quality, in a specification that ignores the experimental design, we estimate an elasticity of 0.22 (Appendix Table 1, regression 1). A naïve conclusion would be that water recontamination in transport and storage prevents nearly 80% of source water quality improvements from reaching the home. Even when attention is restricted to baseline sole-source spring users, and thus endogenous sorting partially avoided, the estimated elasticity is only 0.23 (regression 2). An IV approach that exploits the experimental variation in source quality and addresses attenuation bias tells a different story for sole-source users: the elasticity rises dramatically to 0.66 (significant at 95% confidence; regression 3), so nearly two thirds of gains at the source are translated into home water quality gains. This is some evidence against the claim that recontamination renders source water quality improvements useless.

Focusing on those households on the margin between using the sample spring and an alternative source conceptually allows one to estimate the value households place on spring protection. Spring water quality improvements yield potential utility benefits of $V(W_j^T) - V(W_j)$, and travel costs would have to increase by the same amount to restore households to indifference. The additional travel cost households are willing to incur is a revealed preference measure of their willingness to pay (WTP) for improved water quality.

[Other factors bring the model closer to the data. Most importantly, households make multiple water collection trips and each trip is affected by unobserved factors, including the weather, whose turn it is to collect water, the expected queue, the direction an individual is walking for another task (i.e., to the market) or one's mood that day. Incorporating an i.i.d. error term e_{ijt} modeled as type I extreme value, the indirect utility of a water collection trip to source j at time t is:

$$U_{ijt} = V(W_{jt}) - CD_{ijt} + e_{ijt} \qquad (3).$$

More generally, given a set of characteristics X_{ijt} for individual i and spring j at time t, where controls include both the protection status of the sample spring and the walking time to each potential local water source, as above, the probability household i chooses source j from among all water alternatives $h \in H$ at time t can be represented by the conditional logit formula (McFadden 1974):

$$P(y_{ijt} \mid X) = \frac{\exp(X_{ijt}{}'\beta)}{\sum_h \exp(X_{iht}{}'\beta)} \qquad (4).$$

There is potentially heterogeneity in households' valuation of spring protection as well as in time costs. We allow the coefficient on these two terms to vary as a function of the number of children in the household and their health status, and household sanitation, hygiene knowledge, and education, by including interactions between these characteristics and the treatment indicator (and sometimes also the walking distance term).

We also explicitly estimate this heterogeneity using a mixed logit model (Train 2003). Mixed logit allows for random coefficients on water source characteristics (e.g., spring protection), in the indirect utility function. Simulation techniques are used since there is typically no closed-form solution. We estimate choice probabilities as:

$$P(y_{ijt} \mid X) = \int_{\beta} \frac{\exp(X_{ijt}{}'\beta)}{\sum_{h} \exp(X_{iht}{}'\beta)} f(\beta) d\beta \qquad (5).$$

where y, X and β are defined as above, and $f(\cdot)$ is the mixing distribution, which we take to be the normal distribution for the coefficient on spring protection in our application. Bayesian numerical methods allow us to maximize the log-likelihood to estimate the mean and standard deviation of β.

We use data from the third follow-up survey round, which explicitly asked respondents about the universe of all water sources they could potentially choose and the number of trips they made to each in the last week. The median respondent used two water sources in the last week, and 60% of respondents named additional alternative sources beyond those sources they actually used.

The ratio of the coefficient estimate on the treatment (spring protection) indicator variable to the coefficient estimate on the walking time to a source delivers the value of spring protection in terms of minutes spent walking.

4.3 Estimating households' willingness to pay for cleaner water

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 6, 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 are only minor differences in tastes for program (sample) springs, non-program springs, wells and boreholes. The distance to the closest water source is only weakly correlated with a range of household characteristics, including the distance to the second closest source (not shown), alleviating some concerns about omitted variables in the estimation of how walking distance affects choice.

One issue 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 moderate at 0.38, so attenuation bias could be important. In addition to simple recall error, the variation in reported walking time may be due to actual variation in travel time, depending on the weather, what else the individual is carrying to the source (e.g., a baby, goods for market), and the respondent's health or energy level that day. To approximately correct for classical measurement error in this term, we inflate its coefficient to -0.058 / 0.38 = -0.153 and use this correction in valuation calculations below.[20]

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. Over the course of a year, using the average number of trips per week to sample springs, this is equivalent to 12.7 work days, a reasonably large effect and one independent of the assumed value of time.

If spring protection yields other non-health benefits as well, these estimates could be upper bounds on the willingness to pay to avoid diarrhea cases or deaths. However, while the non-health benefits of spring protection – in terms of water appearance, taste or ease of water collection – could theoretically contribute to willingness to pay, we find no evidence that these have a significant effect on WTP in practice. The inclusion of terms for measured *E. Coli* contamination available 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 to near zero (Table 6, regression 2).

[20] The attenuation bias correction estimated in a Monte Carlo simulation is similar, at roughly 0.3 (not shown).

Valuations are nearly identical for households who had one additional year of experience with spring protection (results not shown), so households' valuation does not appear to change with greater exposure to the protected spring.

By placing a value on individuals' time, we can also estimate the willingness to pay for spring protection in monetary terms. Per capita income in Kenya is US$530 (World Development Indicators 2005), so with a labor share of 70% this translates into average annual worker earnings of US$371, or US$1.42 per work day; this falls in the middle of the range of agricultural labor wages in the study area. Even if people could perfectly substitute time for income at the margin, which is unlikely, this is almost certainly too high a value of time in our rural sample, which is relatively poor by Kenyan standards, and also since collecting water is a task for relatively unskilled household members, mainly young adult women and adolescents (or even younger children; 11% of water collection trips are made by children under age 12). Because limited time-income substitution possibilities are frequently encountered empirically (McKean, Johnson, and Walsh 1995), we follow other authors in focusing on a lower range of values, here 25 to 50% of the average wage as a starting point (Train 1999), although we develop an alternative approach to valuing households' time below.

If the water collector's time is valued at 50% of the Kenyan average wage, or US$0.71 per work day (US$0.0015 per minute), and households make our sample average of 32 water collection trips per week to the sample spring (over two thirds of total water collection trips), 52 weeks per year, the total average value to these households from protection is (3.7 minutes)*(US$0.0015/minute)*(32 trips/week)*(52 weeks/year) = US$9.05 per year (Table 7, Panel A). At the arguably more realistic time value of 25% of the wage ($0.35 per work day), household willingness to pay for spring protection is only US$4.52 per year.

Combining the results from Tables 4 and 6 yields a bound on the willingness to pay to avert child diarrhea. The average number of averted diarrhea cases due to spring protection is (-0.046 cases / child-week) * (2.2 children age 5 and under / household) * (52 weeks / year) = -5.3 diarrhea cases per household-year. Using our spring protection household WTP range of 12.7 work days, this translates into 2.4 work days per case of diarrhea averted. In monetary terms, this is US$0.86-1.72 per case, under the assumption that all of spring protection's value works through child health gains.

A different assumption, namely, that household WTP is driven entirely by reduced child mortality risk, allows us to estimate the value these households place on their children's lives using our revealed preference methodology. There are approximately 5.69 deaths per 1000 children under age five each year in Sub-Saharan Africa (Lopez *et al.* 2006, Table 3B.7), and roughly 4.9 annual diarrhea episodes per African child under age five, based on the findings in Kirkwood (1991), who reviews 100 longitudinal studies of diarrheal disease in 33 African countries. If each diarrhea episode averted (by better quality water) reduced mortality risk by an equal amount, then this translates into 1.16 deaths from diarrhea averted for each 1000 diarrhea cases eliminated. The value of averting one child diarrhea death among these households thus ranges from US$741 to US$1,482. However, this is a lower bound on the value of averting one child diarrhea death to the extent that the diarrhea cases tend to be relatively low mortality risk cases.

Theoretically, households with young children should have both greater time costs of walking to collect water (due to the demands of child care and difficulty carrying a small child) and also greater benefits of clean water, since the epidemiological evidence suggests that young children experience the largest health gains. Empirically, households with more children under age five 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 99% confidence (Table 6, regression 3). The coefficient on the interaction between the treatment indicator and an indicator for households with child under five who had diarrhea in the past week is positive, suggesting that households with sick children also place somewhat greater value on cleaner water, though on net the two effects nearly cancel out.

A lack of knowledge about the link between bad water and child health does not appear to be driving the results. Household valuation for spring protection is not significantly higher as a function of diarrhea prevention knowledge (Table 6, regression 4), nor does household knowledge of the link between contaminated water and diarrheal disease in particular affect the value placed on spring protection (not shown). There is some evidence it rises with latrine ownership and with mother's years of schooling, although the magnitude of the latter interaction effect is small (regression 4). Neither asset ownership nor child gender significantly affects the taste for spring protection (not shown); the latter is perhaps surprising given that health gains appear concentrated among girls.

Using the mixed logit approach, we find only moderate dispersion of spring protection valuations in our sample of households: the mean coefficient estimate on the spring protection indicator remains unchanged at 0.56, with a standard deviation of 0.11 (standard error 0.04, Table 6, regression 5), assuming normally distributed valuations. The mean and standard deviation of household willingness to pay for spring protection based on these revealed preference mixed logit results are presented in Table 7, Panel A, in terms of days worked and for a variety of time values.

Comparison of Revealed and Stated Preference Water Valuations

This subsection compares our to the revealed preference spring protection valuations to different stated preference approaches—stated ranking and contingent valuation. The stated ranking approach asks respondents to rank order their potential water source options rather than relying on information on actual household water trips. This ranking is done sequentially in the survey, with the highest ranked source eliminated from the choice set at each subsequent question. These data are then used in a travel cost discrete choice analysis nearly identical to the revealed preference regressions above.

Estimated stated preference ranking WTP for spring protection is much higher than the revealed preference estimates. The magnitude of the coefficient estimate on distance walking falls to -0.033 while that on spring protection rises to 0.96 (Table 6, regression 6). Using the same attenuation bias correction as above, this is almost exactly three times greater than the revealed preference value (Table 7, Panel B), and this ratio is independent of our assumption about the value of time. The WTP for one year of spring protection is 38.3 work days, and using time valued at 50% (25%) of the average Kenyan wage becomes US$26.19 (US$13.09).

Comparing the analogous columns in Table 6 (regressions 1 and 6) also highlights interesting reporting patterns in the stated preference ranking case. The coefficient estimates on several unimproved sources many Kenyans generally 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 more positive. This suggests social desirability bias may be playing a role in some respondents' stated preference rankings.

We also find considerably larger dispersion in stated preference valuations for spring protection than in the analogous revealed preference mixed logit estimates, with standard deviations three times larger (Table 7, Panel B). One plausible explanation for the dispersion is that many respondents fail to introspect carefully in this hypothetical exercise, and thus their resulting answers are far "noisier" than in the revealed preference case, where they face real (time) budget constraints.

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 first-hand experience with spring protection. In the final wave of the survey, respondents were first asked if they would be willing to pay either 250 or 500 Kenyan Shillings, followed by a question that emphasized the expenditure trade-off for their assigned amount (in other words, what goods they would be giving up 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, is standard in the contingent valuation literature (Bateman and Willis 1999).The 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 for one year's access to the protected spring, assuming everyone else would also have to pay this amount too?[21]

The main finding is that nearly all households claim to be willing to pay US$7.14 for one year of spring protection, and the majority of households are willing to pay twice that (US$14.29) even after being walked through the expenditure trade-offs by the enumerator (Table 7, Panel C). The use of the expenditure trade-off prompt reduces willing to pay substantially (by 11-14 percentage points), indicating that the CV results are sensitive to survey question framing. Valuations are also sensitive to the starting value assigned to households: those respondents randomly chosen to be asked whether they valued a year of spring protection at 500 Kenyan Shillings have mean willingness to pay that is twice as high ($23.91) as those respondents first asked about a value of 250 Kenyan Shillings ($12.62), perhaps because the proposed starting value implicitly contains some information for respondents about what a "reasonable" valuation should be.

If we assume spring protection valuations are normally distributed, the distribution that best fits the CV response data (using maximum likelihood) has mean willingness to pay of US$17.64 (standard deviation US$13.09, Table 7, Panel C). This is considerably more than our best revealed preference estimates of US$4.52-9.05 per year, but lies within the range of plausible stated preference ranking valuations discussed above (US$13.09-26.19).

We can estimate water collector's time by assuming that the mean WTP for spring protection is equated under both stated preference approaches (stated ranking and contingent valuation).[22] This implies an exact average value of time, and yields a mean revealed preference WTP of US$5.79

---

[21] 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.
[22] We thank Enrico Moretti for this suggestion.

(with standard deviation US$3.05) for one year of protection. The estimated WTP distributions for the three valuation approaches under this assumption (Figure 4) indicate that stated preference methods exaggerate household willingness to pay for environmental amenities in a rural Kenyan setting, and that the revealed preference approach yields more modest and less variable valuations.

5. Alternative policies and institutions for providing clean water

5. 1 The Social Return to Spring Protection

Current Kenyan law and custom prevent landowners from charging local spring users for water. Perhaps partially as a result of these weak private property rights, virtually no springs are privately protected in our study area, despite the identified health gains. In this section we use the valuation estimates derived above to determine the socially optimal level of spring protection in this region, before estimating welfare under alternative policies and institutions.

We begin by considering whether a social planner would opt to protect the springs in our sample, given observed willingness to pay. 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 discounted net present cost of US$1,480 (with a 5% annual discount).

If households are willing to pay $5.79 per year and, as in our data, 31 households use the typical spring, the net present value of WTP for spring protection is US$1386. Thus for a spring with a typical number of baseline users, the social rate of return to spring protection appears negative. Yet the number of households using a spring grows following protection, and it is appropriate to consider utility gains for these new users when calculating the social rate of return, as well. We conducted a census at sample springs and found a 47% increase in the number of spring user households. This boosts the average number of users to 46 households per spring and the benefits to approximately

US$2,037, in which case the social rate of return becomes positive. In even more densely populated rural areas, or in towns and cities, average social returns would be even larger.

However, real-world public provision would be costlier due to tax distortions. Under the assumption that the deadweight loss from tax distortions is roughly 0.3 of the revenue raised (Ballard, et al 1985), the total social cost per spring rises to US$1,924, bringing the social return on this investment close to zero.

The relevant decision for some policymakers may not be spring protection versus doing nothing, but rather investing in spring protection versus other interventions to improve water quality and ultimately child health. We next compare the diarrhea reduction cost-effectiveness of spring protection versus a point-of-use (POU) water treatment, at-home chlorination, also introduced in our study sample. The reduction in diarrhea incidence among children under age three from introducing POU treatment was about 45%, double the effect from spring protection.[23] This is an intention to treat (ITT) estimate, and POU take-up was about 54%, so treatment effects on the treated were extremely large for POU users. The obvious implication is that larger reductions in water contamination (from POU treatment) are more effective at reducing diarrhea than moderate water improvements such as those achieved from spring protection.[24]

These estimates allow us to compare the child health benefits associated with spring protection versus the reductions in diarrheal morbidity that could have been realized had the approximately US$148,000 spent to protect the 100 treatment springs in our sample and maintain them for ten years, instead been spent providing POU products to households with young children. In

[23] A fuller presentation of these POU results is contained in ongoing research (Kremer, Miguel, Null, and Zwane 2007). In that work we also estimate if POU technologies are most effectively employed as complements or substitutes for source water improvements like spring protection.
[24] In a near-by setting but using a different and more intensive method of measuring diarrhea, Crump *et al.* (2005) estimate that the effect of point of use water treatment is a reduction in diarrheal prevalence per 100 weeks of about 17 percent, similar to the reduction we estimate for source water quality improvements. This impact estimate is for children under two and in a setting where high take up led to 78% of treatment households with *E. coli* MPN <1.

our study area, a one month's supply of the in-home chlorination product (called WaterGuard locally) can be purchased for roughly $0.29 (20 Kenyan Shillings).

We begin by noting that 19% of comparison children in our sample (who were under age three at baseline or born since then) are reported having diarrhea in the past week across all survey rounds. All other things equal, and assuming constant numbers of children per household for the average of 46 spring using households post-protection, this implies 234,208 cases at households that use sample springs over a spring's ten year lifespan, and (234,208) * (0.046/0.19) = 56,703 cases averted as a result of spring protection. The cost per diarrhea case averted due to spring protection works out to US$2.61.

We can also calculate the number of Disability Adjusted Life Years (DALYs) averted by spring protection, using the standard WHO approach.[25] This calculation is a function of the average length of a diarrhea episode, and the number of deaths that occur per 1000 diarrhea cases. An ongoing high-frequency data collection effort, in which we collect morbidity diaries, indicates that average diarrhea episode length in our sample is 6.5 days. With 1.16 deaths per 1000 cases of diarrhea (as discussed above), spring protection averts 2,146 DALYs, a cost of US$69 per DALY averted. This is an underestimate of spring protection's benefits if there are benefits for people over age five, although above we did not find any statistically significant health gains in that group.

If the chlorine POU product were given to every household with eligible children in our sample (about 80% of homes) for ten years, this would cost $96,516 in current dollars using a time discount rate of 5%. The cost per case of diarrhea averted with WaterGuard is thus about $0.92, and so nearly three times as many cases could be averted by focusing on point-of-use treatment products instead of spring protection. Of course, this assumes that a 54% take-up rate could be sustained over ten years and that the delivery mechanism could perfectly target the households with young children.

[25] For more information on the DALY concept, see: http://www.who.int/healthinfo/boddaly/en/index.html.

Still, any long-run POU take-up rate above about 19% would result in in-home chlorine treatment being more cost-effective than spring protection at reducing diarrhea.[26]

The greater the number of households that benefit from the intervention, the more cost-effective spring protection is. Thus, as population density increases source water quality investments compete more attractively with POU treatment. Assuming the current retail price of WaterGuard, spring protection becomes more cost-effective than POU water treatment when there are at least 127 household users per spring. While very high for our rural study area, this usage level is certainly possible at standpipes in urban and peri-urban areas. Improving source water quality is more likely to be an appropriate policy in areas with denser population and fewer alternative sources, leading to many users per spring.

5.2 Privatization Policy Simulations

Privatization – allowing land owners to charge households for access to spring water on their land – is a widely discussed policy option in the rural water sector. Any resulting profits would provide an incentive for land owners to invest in spring protection. A downside would be the static distortion introduced by pricing spring water above its marginal cost of zero, as well as possibly adverse distributional consequences.

We simulate the following game. In t = 0, the property rights regime is chosen. Three cases are: the "status quo," which is characterized by spring water prices of zero and no spring protection; the "social planner" solution; and "privatization." In the latter case, landowners decide whether or not to protect their spring (in t = 1), and then set a non-negative price per unit of water collected (t =

[26] Conclusions about the cost-effectiveness of the POU technology are sensitive to its price. The current price of WaterGuard in local markets in Kenya is heavily subsidized by the organization marketing the product: the price covers production and distribution costs but marketing costs are fully subsidized. Yet POU chlorination costs would be much lower if chlorine were distributed in bulk, without the marketing, packaging, and retail distribution costs needed to sell in Kenyan shops. The local cost of a similar chlorine product also available in retail markets but distributed in much larger containers, bleach for washing, is only 27% of the cost of WaterGuard per unit of chlorine. If the cost of POU chlorine were reduced to this price, the cost per diarrhea case averted would be $0.25 and point-of-use treatment would be ten times more cost effective than spring protection at averting diarrhea cases.

2). Finally, in t = 3, households make their water collection choices, where these are governed by the revealed preference coefficients estimated in the mixed logit specification (Table 6, regression 5).

Given local demand, spring owners maximize profits, which are equivalent to revenues minus any spring protection construction and maintenance costs (since the marginal cost of a unit of spring water is zero). In practice, we find this optimal price level through a grid search over a wide range of non-negative prices. Spring owners optimize over a ten-year period (the same time horizon used in the social planner's problem above). In the status quo case, households make their water consumption choices given a water price of zero and no spring protection.

We allow spring resources to be privatized but not other source types. This is a natural policy in our setting: it is difficult to restrict access to natural rivers, streams, ponds and lakes, and many wells – especially boreholes – are sunk on public property, including school grounds or market centers, where there is no private landowner. We do not allow spring owners to engage in price discrimination since easy resale in rural settings makes it unlikely to succeed, and we also abstract from possible collusion among spring owners. Pricing is linear in the amount of water collected.

There is no spring protection in the status quo case, and we normalize status quo household utility to zero to facilitate later comparisons (Table 8). We present household utility and social welfare converted into U.S. dollar values, on a per spring basis throughout.

In the social planner solution, 26% of springs (48 springs in all) are protected, and these tend to be springs with largest number of household users. The net social gain is US$156 per spring in the planner case. When we consider the more realistic case of public provision taking into account tax distortions, only 11% of springs are protected, and slightly less than half of those the social planner would optimally protect. The social gain per spring falls to US$126 here.

We next focus on the simplest privatization case, what we call "pure" privatization, where there are no restrictions on landowners' behavior.[27] Land owners are assumed to understand the water source choice situation facing households, including the distance to each of the household's options, the disutility of walking time, and the average willingness to pay for spring protection (from the results in Table 6). Land owners have some market power, given the limited alternative water choices available to households. In this "pure" case, we find that only 3% of spring owners find it profit-maximizing to protect their spring, given local water demand conditions (Table 8). A particularly interesting subsample to consider are the 48 springs the social planner would protect. Yet among these springs only six, or 13%, choose spring protection.

In this analysis, the privatization of spring water resources thus appears only moderately effective at reaching socially efficient levels of water infrastructure investment, and has adverse distributional consequences. There are large rents for land owners – the net present value of profits per landowner is US$2642 – who are now able to charge positive prices for water, and losses for the vast majority of consumers. All households have lower utility in the pure privatization case than under the status quo, and the annual price charged per household in these springs reaches US$29, equivalent to several weeks' wages. All households are now paying for access to spring water but few realize health benefits since the vast majority of springs remain unprotected. Even those users of newly protected springs lose as land owners extract consumer surplus through positive prices.

Private property rights can be structured in alternative ways. Our first "conditional" policy would only permit spring owners to charge positive prices if they invest in spring protection. This sort of property right is common in rural Africa, including in Ghana where actively farming a plot is critical to securing property rights (Goldstein and Udry 2005). The simulation of this policy yields

[27] We currently only consider demand among baseline spring user households, but will incorporate demand from new household users post-protection into future versions. We also assume that each sample spring owner makes pricing and protection decisions without considering the decisions being made by other spring owners; we will also extend the analysis to consider strategic competition among landowners in future work.

much higher rates of investment in water infrastructure: now 72% of spring owners find it profit maximizing to invest in spring protection. Protection rates appear too high: most land owners choose to protect their spring even when it is not socially optimal, because only through protection can spring owners secure their property rights and capture rents (by extracting consumer surplus). Yet because households' willingness to pay for spring protection is fairly modest, this increase in protection does not leave most communities better off overall.

While this modified privatization policy still leaves most households worse off than the status quo, it is marginally better for household welfare than pure privatization: only 73% of households experience utility losses from this modified form of privatization, and total household utility is somewhat higher. However, net social welfare remains negative and far worse than the status quo.

A variant on this conditional privatization scheme is Pareto improving and moves these communities closer to socially optimal investment levels, surpassing the status quo in terms of social welfare. Consider a technological innovation that allows some spring water to be protected, while some fraction of the spring water continues to flow freely and unprotected. This is simple to achieve in rural Kenya, simply requiring land owners to allow some water to flow away from the protected spring and pool elsewhere, where it is exposed to the environment and becomes a pool of unprotected spring water. The third private property rights institution we consider would allow land owners to charge for the protected spring water, but not for the unprotected water.

The simulation indicates that, at the springs the planner would protect, 79% of land owners choose to protect their spring in this case, while no land owners protect springs the planner would not protect. While this property rights reform does not achieve the socially optimal level of protection, it does incentivize land owners to perform some socially beneficial spring protection while at the same time limiting both the distortions due to above marginal cost pricing and the welfare costs to consumers. The availability of free unprotected spring water shields households from the utility losses they experience under the pure privatization or conditional privatization case, and no households are worse off than under the status quo. Land owners earn small positive profits.

Taken together, this "conditional" privatization case with access to free unprotected water delivers social welfare gains similar to those realized under the public provision case. Private provision may even have some advantages over public provision, especially with regards to ongoing incentives for spring maintenance over time, but there are also potential drawbacks. In practice, it may be expensive for land owners to monitor the amount of water that households collect, and the transaction costs involved in water sales – which have not been considered in our simulations – would reduce the benefits from privatization. If the transactions costs of selling water exceed land owner profits, this privatization case will yield social welfare no higher than the status quo. Another concern with privatization in rural Kenya is the likely local resistance to allowing land owners to charge for water, which would run against strongly held traditional social norms.

6. Discussion and conclusion

Spring protection dramatically improved source water quality in a rural African setting, reducing contamination by two thirds on average and home water contamination by nearly one quarter. Child diarrhea fell by roughly one quarter, although source water protection appears less cost-effective than a point-of-use treatment at reducing child diarrhea. In contrast to common interpretations in the existing public health literature, we do not find 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, water quantity used, or to changes in any other preventive health behaviors that we measured. However, there were sharp changes in water source choices among some households.

By capitalizing on the observed changes in water source choice, we develop revealed preference estimates of willingness to pay for improved water quality. Because of the experimental research design, these travel cost estimates are not subject to typical econometric concerns, and can be used to validate the reliability of stated preference WTP estimates. We find moderate household valuation for spring protection, on the order of 12.7 work days, or US$4.52-9.05 per household annually. This translates into a household willingness to pay approximately 2.4 work days US$0.86-1.72 per averted child diarrhea case. In contrast, stated preference valuation approaches produce estimated WTP estimates higher than these by as much as three times, and with much greater dispersion in reported valuations, which raises questions about the reliability of such methods in settings like ours.

We can compare the valuation of these health benefits to the costs of spring protection, and find that the social returns to spring protection are near zero or even negative at the average population density level observed in our study and many other rural African settings.

The modest household valuations for clean water also imply that stronger private property rights to rural water resources need may not achieve socially optimal levels of investment: under several plausible privatization reform policies, the majority of household consumers are worse off relative to the status quo, as they now often have to pay for unprotected spring water that they previously collected for free. These counterfactual policy simulations highlight the possible negative unintended consequences of well-intentioned property rights reforms in the rural water sector. However, a “conditional” privatization reform that guarantees households continued access to free unprotected water is a Pareto improvement over the status quo.

The constellation of findings in this paper can be interpreted as calling into question many current policies in the rural water sector. The high-profile U.N. Millennium Development Goals call for access to safe drinking water (General Assembly of the United Nations 2000), and indeed in our study we find that improvements in water quality, especially the relatively large gains that come from at-home chlorination relative to source improvements, do improve child health and can be extremely cost effective.

However, in practice, progress in the rural water sector is measured in terms of improved water source coverage rates, not water quality, perhaps because measuring investments in springs, wells and stand-pipes is cheaper than measuring actual household water quality. The spring protection intervention that we study counts towards these targets, but we find it has a negative social rate of return for many communities in our sample. Were targets reframed in terms of water quality alone, source water investments of this sort might be reprioritized (and downgraded) relative to point-of-use water treatment. The source water investments that make up the bulk of spending in the rural water sector today do not appear to be the most cost-effective approach to reducing diarrheal mortality in settings like ours.

References

Angrist, J., G. Imbens, and D. Rubin (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.

Ballard, Charles L., John B. Shoven, and John Whalley. (1985). "General equilibrium computations of the marginal welfare costs of taxes in the United States", *American Economic Review*, 75(1), 128-138.

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 #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.

Bryce, J. *et al.* (2005), "WHO estimates of the causes of death in children," Lancet 365: 1147-52.

Cameron, T.A. (1992), "Combining contingent valuation and travel cost data for the valuation of nonmarket goods," Land Economics, 68(3):302-317.

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.

Choe, K. (1996) "The economic benefits of surface water quality improvements in developing countries: A case study of Davaeo, Philippines" Land Economics 72(4): 519-537

Crump J.A., Otieno P.O., 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.

Curtis, V., S. Cairncross, and R. Yonli. (2000), "Domestic Hygiene and diarrhea—Pinpointing the problem." Tropical Medicine and International Health 5(1):22–32.

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

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. (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.

Feder, G., T. Onchan, Y. Chalamwong, and C. Hongladarom, (1988), Land Policies and Farm Productivity in Thailand. Johns Hopkins University Press, for the World Bank: Baltimore.

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

Galiani, S., P. Gertler and E. Schargrodsky, (2005), "Water for Life: The Impact of the Privatization of Water Services on Child Mortality," Journal of Political Economy, 113:83-120.

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.

Goldstein, Markus, and Christopher Udry. (2005). "The Profits of Power: Land Rights and Agricultural Investment in Ghana", unpublished manuscript World Bank and Yale University.

Griffin, C. et al (1995) “Contingent valuation and actual behavior: predicting connections to new water systems in the state of Kerala, India” World Bank Economic Review 9(3): 373-395

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.

Hoy, M. and Jimenez, E., (2006), "Property Rights and the Urban Environment: Local Public Goods in Indonesian Cities," Working Paper 2006-9, University of Guelph, Department of Economics.

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.

J. C. Whitehead, (2006), “A comparison of contingent valuation method and random utility model estimates of the value of avoiding reductions in King Mackerel bag limits," Applied Economics, 38(15):1725-1735.

Kirkwood, B.R. (1991). “Diarrhea”, pp. 134-157, in R.G. Feacham and D.T. Jamison, eds., *Disease and Mortality in Sub-Saharan Africa*. New York: Oxford University Press for World Bank.

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.

Lopez, A.D., et al., eds. (2006). *Global Burden of Disease and Risk Factors*, Washington DC: World Bank and Oxford University Press.

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).

Miguel, E., and M. K. Gugerty. (2005). “Ethnic diversity, social sanctions, and public goods in Kenya”, Journal of Public Economics, 89(11-12), 2325-2368.

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.

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

Mumma, Albert. (2005). “Kenya’s new water law: an analysis of the implications for the rural poor”, paper presented at the *Workshop on African Water Laws: Plural Legislative Frameworks for Rural Water Management in Africa*, 26-28 January 2005, Johannesburg, South Africa.

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.

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

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.

Platteau, Jean-Phillipe. (2000). *Institutions, Social Norms, and Economic Development*. Harwood Academic Publishers.

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

Rosado, Maria et al (2006), “Combining averting behavior and contingent valuation data: an application to drinking water treatment in Brazil,” Environment and Development Economics, 11:729-746

Toubkiss, J (2006), Costing MDG Target 10 on water supply and sanitation: Comparative analysis, obstacles and recommendations. Paris: World Water Council & World Water Forum. Available: http://www.worldwatercouncil.org/index.php?id=3

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

Train, K., “Mixed Logit Models of Recreation Demand,” in C. Kling and J. Herriges, eds., *Valuing the Environment Using Recreation Demand Models*, Elgar Press, 1999.

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

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

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

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

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.* (2000), "Reducing deaths from diarrhea through oral rehydration therapy," Bulletin of the World Health Organization 78(10), 1246-56.

Viscusi, W. Kip, and Joseph E. Aldy. (2003). "The Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World", *Journal of Risk and Uncertainty*, 27(1): 5-76.

Wade T. J. *et al*.(2003), "Do Environmental Protection Agency water quality guidelines for recreational waters prevent gastrointestinal illness? A systematic review," Environmental Health Perspectives 111(8):1102-09.

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.

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.

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.

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	3.77	95	0.13
	(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 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 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	4.04	736	3.93	719	0.11
	(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	48.03	733	47.99	716	0.04
	(36.51)		(38.48)		(2.51)

	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	-0.01
	(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 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	1047	1.72	995	-0.02
	(0.95)		(0.97)		(0.04)
Child male (=1)	0.52	1047	0.50	995	0.02
	(0.50)		(0.50)		(0.02)
Child had diarrhea in past week indicator	0.23	996	0.20	961	0.03
	(0.42)		(0.40)		(0.02)
Child height (cm)	76.10	870	76.13	835	-0.03
	(11.67)		(12.16)		(0.57)
Child weight (kg)	9.98	864	10.02	810	-0.05
	(3.04)		(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 the spring level when using household level data), significantly different than zero at * 90% ** 95% *** 99% confidence.
Diarrhea is defined as three or more "looser than normal" stools per day.
Assigned spring is the project sample 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 youngest adult 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 (2004-2007)

	Dependent variable: ln(Spring water *E. coli* MPN)			
	(1)	(2)	(3)	(4)
Treatment (protected) indicator	-1.08	-1.08	-1.03	-1.09
	(0.28)***	(0.27)***	(0.23)***	(0.24)***
Baseline ln(Spring water *E. coli* MPN)		0.43	0.96	0.98
		(0.04)***	(0.04)***	(0.05)***
Baseline ln(Spring water *E. coli* MPN)			-0.17	-0.15
* Treatment indicator			(0.12)	(0.13)
Baseline latrine density				-0.19
				(0.61)
Baseline latrine density * Treatment indicator				0.86
				(1.75)
Baseline diarrhea prevention score				-0.04
				(0.07)
Baseline diarrhea prevention score				-0.31
*Treatment indicator				(0.25)
Baseline boiled water yesterday density				0.42
				(0.66)
Baseline boiled water yesterday density				0.82
*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)	-0.25	-0.33	-0.37	-0.30
	(0.30)	(0.20)*	(0.17)**	(0.20)
Treatment group 2 (phased in late 2005)	-0.20	-0.24	-0.27	-0.22
	(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	3.64 (1.94)	3.64 (1.94)	3.64 (1.94)	3.64 (1.94)

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 some of 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 (survey round and wave) 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 its interaction with the treatment indicator are included as additional control variables (not shown in the table).
The -108 log point effect in column 1 is equivalent to a 66% reduction in *E. Coli* fecal coliform units per 100ml.

Table 3: Spring protection household water quality impacts (2004-2007)

	Dependent variable: ln(Home water *E. coli* MPN)		
	(1)	(2)	(3)
Treatment (protected) indicator	-0.26	-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		0.04	0.07
		(0.25)	(0.25)
Baseline latrine density	-0.83	-0.84	-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			-0.06
			(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			0.51
			(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	-0.03
	(0.14)	(0.18)	(0.27)
Treatment group 2 (phased in late 2005)	-0.13	-0.17	-0.22
	(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 (survey round and wave) 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.
The -26 log point effect in column 1 is equivalent to a 23% reduction in *E. Coli* fecal coliform units per 100ml.

Table 4: Health outcomes for children under age three at baseline or born since 2004 (2004-2007 data)

	------------Dependent variable: Diarrhea in past week ------------						Dependent variable Weight (kg)		Dependent variable Body mass index, BMI (kg/m^2)	
	(1)	(2) Probit	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment (protected) indicator	-0.044***	-0.052***	-0.044*	-0.046**	-0.088***	-0.026	-0.011	-0.022	0.12	0.17
	(0.012)	(0.019)	(0.023)	(0.023)	(0.029)	(0.039)	(0.073)	(0.094)	(0.13)	(0.16)
Treatment (protected) indicator * Male					0.081**			-0.022		-0.089
					(0.040)			(0.123)		(0.186)
Treatment (protected) indicator * Baseline latrine density						0.101				
						(0.123)				
Treatment (protected) indicator * Baseline diarrhea prevention score						-0.0079				
						(0.0073)				
Treatment (protected) indicator * Baseline mother's years of education						0.0023				
						(0.0044)				
Child fixed effects	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Treatment group fixed effects	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month of year controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Gender-age controls	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.00	-	0.53	0.53	0.53	0.53	0.96	0.96	0.69	0.69
Child-year observations	6759	6758	6758	6663	6678	6604	5746	5746	5652	5652
Mean (s.d.) of the dependent variable in the comparison group	0.19	0.19	0.19	0.19	0.19	0.19	11.38	11.38	17.0	17.0
	(0.39)	(0. 39)	(0. 39)	(0. 39)	(0. 39)	(0. 39)	(3.52)	(3.52)	(2.2)	(2.2)

Notes: Column 2 estimated using probit (marginal effects presented), columns 1 and 3-10 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. The gender-age controls include linear and quadratic current age (by month), and these terms interacted with a gender indicator. Columns 2-10 also contain survey round controls. In column 6, additional control variables are number of children under 12 living in the home, home has iron roof indicator, iron roof density within spring community, and the boiled water yesterday indicator (all measured at baseline), all interacted with the treatment indicator.

Table 5: Treatment effects on household water source choice and health behaviors (2004-2007)

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
Panel A: Water transportation and storage	(1)	(2)	(3)	(4)
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 B: 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 has soap in the home indicator	-0.01 (0.02)	-0.01 (0.02)	0.01 (0.03)	0.89 (0.31)
Fingers with bacterial contamination (fecal *Streptococci* colonies)	0.09 (0.13)	0.08 (0.17)	0.26 (0.24)	0.71 (1.26)
Panel C: Spring amenities (recorded by enumerators) [(d)]				
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)
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)
Panel D: Water collection and source choice				
Fraction of trips to assigned spring	0.09 (0.03)***	0.03 (0.02)*	0.20 (0.05) ***	0.76 (0.40)
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)
Trips made to get water (all uses, members, sources) past week	-2.46 (2.15)	-0.86 (2.39)	-4.52 (3.51)	31.77 (24.42)
Self-reported distance to nearest water (min.)	-1.41 (0.44) ***	-1.71 (0.48) ***	-1.03 (0.80)	7.92 (7.23)
Calculated distance (GPS) to assigned spring (km)	0.03 (0.03)	0.06 (0.05)	0.01 (0.01)	0.36 (2.52)

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 the dependent variable is reported in the 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 project sample spring that we believed households used at baseline based on spring user lists. The fingertip contamination results are for the respondent's main hand (so values range from 0-5).

(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.

(d): Panel C contains spring level information, so the breakdown into sole source and multi-source households is not possible.

Table 6: Discrete choice models (conditional and mixed logit) of water source choice (2007 surveys)

	---------------- Revealed Preference ----------------					--- Stated Ranking ---	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treatment (protected) indicator	0.56***	0.09	0.41***	0.54***		0.96***	
	(0.05)	(0.07)	(0.08)	(0.08)		(0.24)	
Mixed logit – Mean (normal):					0.56***		0.97***
					(0.05)		(0.24)
Mixed logit – Std. dev. (normal):					0.11***		0.22
					(0.04)		(0.19)
ln (source water E. coli MPN)		-0.13***					
		(0.01)					
Water quality at source perceived to be above average		0.95***					
		(0.06)					
Distance to water source (minutes walking)	-0.058***	-0.060***	-0.042***	-0.058***	-0.058***	-0.033***	-0.033***
	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.009)	(0009)
Distance * Children aged 0-5 with diarrhea last week			-0.016***				
			(0.003)				
Treatment indicator * Children aged 0-5 with diarrhea last week			0.23***				
			(0.06)				
Treatment indicator * Baseline latrine ownership				1.8***			
				(0.2)			
Treatment indicator * Baseline diarrhea prevention score				0.027			
				(0.018)			
Treatment indicator * Baseline mother's years of education				0.027**			
				(0.010)			
Source type: Borehole/piped	-0.14*		-0.11	-0.18**	-0.14*	0.07	0.06
	(0.07)		(0.07)	(0.07)	(0.07)	(0.25)	(0.25)
Source type: Well	-0.20***		-0.26***	-0.23***	-0.21***	-0.43*	-0.44*
	(0.06)		(0.06)	(0.06)	(0.06)	(0.23)	(0.24)
Source type: Stream/river	-0.52***		-0.48***	-0.44***	-0.53***	-2.19***	-2.19***
	(0.09)		(0.09)	(0.09)	(0.08)	(0.50)	(0.51)
Source type: Lake/pond	-0.13		-0.15	-0.17	-0.13	-2.82	-2.84
	(0.19)		(0.20)	(0.18)	(0.19)	(1.80)	(1.76)
Log likelihood at convergence	-5775	-2654	-5400	-5346	-5770	-363	-363
Number of observations	53445	29086	51006	50042	53445	2114	2114
Number of households	453	330	429	423	453	483	483

Notes: Conditional logit model in columns 1-4 and 6 and mixed logit model in columns 5 and 7 (grouped by trip or choice situation, and weighting each household equally). Significantly different than zero at * 90 ** 95 *** 99% confidence. In columns 1-5 each observation represents a unique household-water source pair in a given water collection trip. In columns 6-7, 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 “non-program spring”. The coefficient estimate on the indicator for being the assigned program sample spring, the spring that we believe households used at baseline based on spring user lists, is not shown. In column 3, additional controls are included for children aged 0-5 and 5-12 at baseline, and the distance to water source term, directly and interacted with the treatment indicator (not shown). In column 4, additional control variables are number of children under 12 living in the home, home has iron roof indicator, iron roof density within spring community, and the boiled water yesterday indicator (all measured at baseline), directly and interacted with the treatment indicator.

Table 7: Valuation of one year of spring protection (2007 survey)

	One year of spring protection			
Panel A: Revealed preference valuation (from mixed logit – Table 6, column 5)	Mean	Std. dev.		
Work days (8 hour days)	12.7 days	6.7 days		
Assume value of time is 50% Kenyan worker average wage	$9.03	$4.75		
Assume value of time is 25% Kenyan worker average wage	$4.52	$2.38		
Equate mean stated preference ranking and contingent valuation	$5.79	$3.05		
Panel B: Stated preference ranking valuation (from mixed logit – Table 6, column 7)				
Work days (8 hour days)	38.3 days	20.2 days		
Assume value of time is 50% Kenyan worker average wage	$27.52	$15.76		
Assume value of time is 25% Kenyan worker average wage	$13.76	$7.88		
Equate mean stated preference ranking and contingent valuation	$17.64	$10.10		
Panel C: Contingent Valuation			Full Round	Final Wave, emphasizing trade-offs
Proportion willing to pay this for spring protection:				
US$3.57 (250 Kenya Shillings)			0.94 [308]	0.80 [98]
US$7.14 (500 Kenya Shillings)			0.90 [316]	0.79 [204]
US$14.29 (1000 Kenya Shillings)			-	0.60 [204]
	One year of spring protection			
	Mean	Std. dev.		
Sample: Final Wave, emphasizing trade-offs	$17.64	$13.09		
Subsample with 250 KSH starting value	$12.62	$11.06		
Subsample with 500 KSH starting value	$23.91	$14.28		

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

Table 8: Property Rights Institutions: Counterfactual Policy Simulations

	Proportion of springs protected	Average price per water trip (USD)	NPV profits, per land owner (USD)	NPV household welfare, per spring (USD)	Proportion households with lower utility than status quo	NPV of public spending on springs (USD)	Social welfare, per spring (USD)
Status quo	0.00	0.000	0.00	0.00	0.00	0.00	0.00
Social planner	0.26	0.000	0.00	528.6	0.00	-371.4	157.2
Public investment (including tax deadweight loss)	0.11	0.000	0.00	293.0	0.00	-165.9	127.1
Springs social planner does not protect	0.00	0.000	0.00	0.00	0.00	0.00	0.00
Springs social planner protects	0.45	0.000	0.00	1140.9	0.00	-638.2	494.8
"Pure" privatization	0.03	0.0172	2656.1	-3885.9	1.00	0.00	-1229.8
Springs social planner does not protect	0.00	0.0164	1749.7	-2635.7	1.00	0.00	-886.0
Springs social planner protects	0.13	0.0195	5235.8	-7444.0	1.00	0.00	-2208.2
"Conditional" privatization	0.72	0.0148	1943.9	-3356.2	0.73	0.00	-1412.3
Springs social planner does not protect	0.62	0.0126	934.7	-2073.7	0.63	0.00	-1139.0
Springs social planner protects	1.00	0.0209	4816.1	-7006.1	1.00	0.00	-2190.1
"Conditional" privatization, access to unprotected water	0.20	0.0007	117.7	11.1	0.00	0.00	128.8
Springs social planner does not protect	0.00	0.000	0.00	0.00	0.00	0.00	0.00
Springs social planner protects	0.79	0.0027	452.5	42.7	0.00	0.00	495.2

Notes: The status quo assumes that water prices are zero at all sources, and that all springs remain unprotected. The "pure" privatization scenario allows land owners complete freedom in charging water prices and in deciding whether or not to protect their springs. The "conditional" privatization scenario only allows land owners to charge positive prices at protected springs. "Conditional" privatization access to unprotected water prohibits land owners from charging for this unprotected water. Household utility is normalized to take on a value of zero in the Status Quo case. Net present values are discounted at 5% annually for both households and land owners, over 10 years. Household utility values are converted into USD using the value of time that equates mean stated preference ranking and contingent valuation (as in Table 7 above).

Appendix Table 1: The elasticity of household water quality with respect to spring water quality (2004-2005)

	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: Rural Water Project (RWP) study region and sample springs

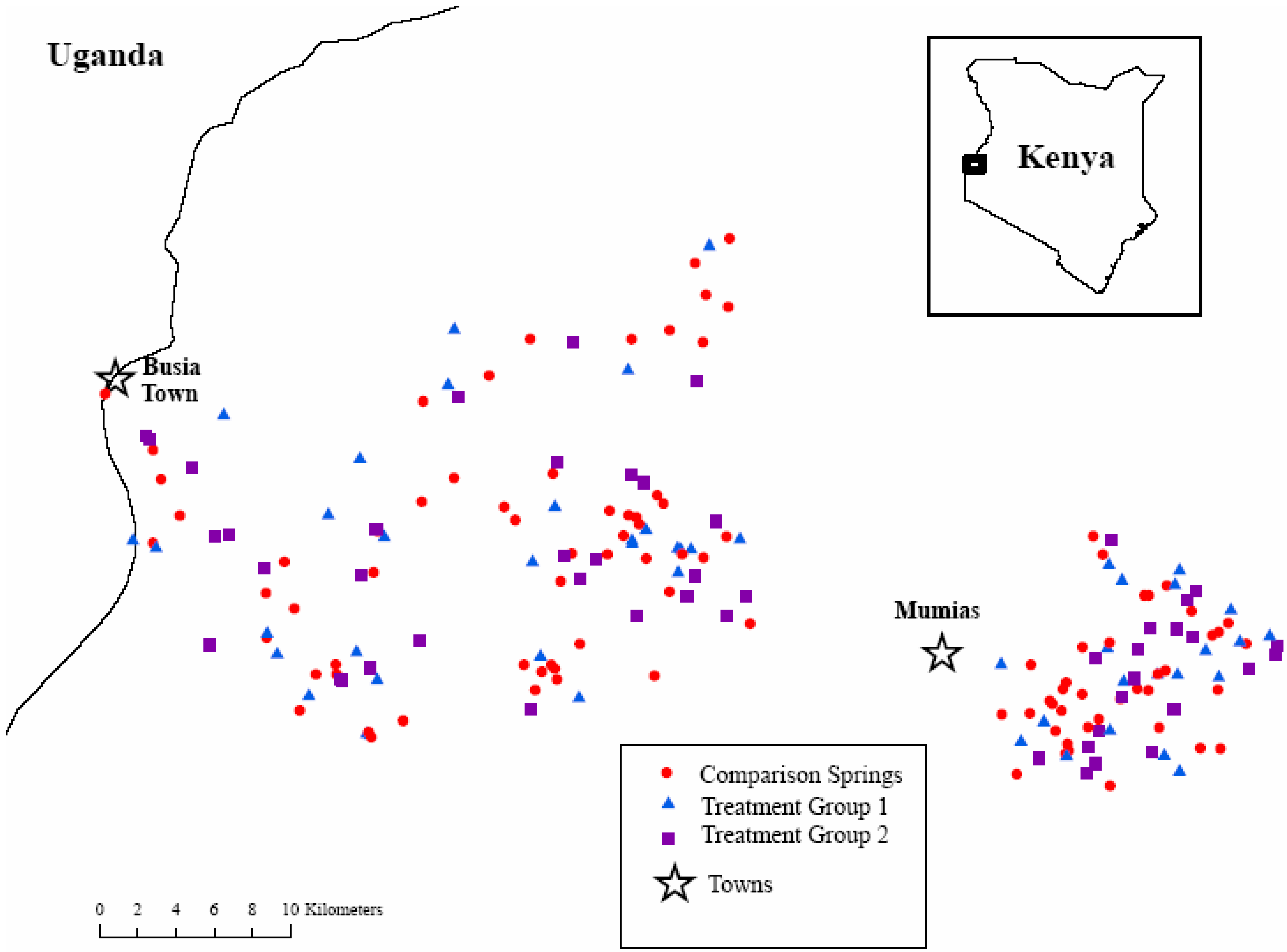


Figure 2: Rural Water Project (RWP) Timeline 2004-2007

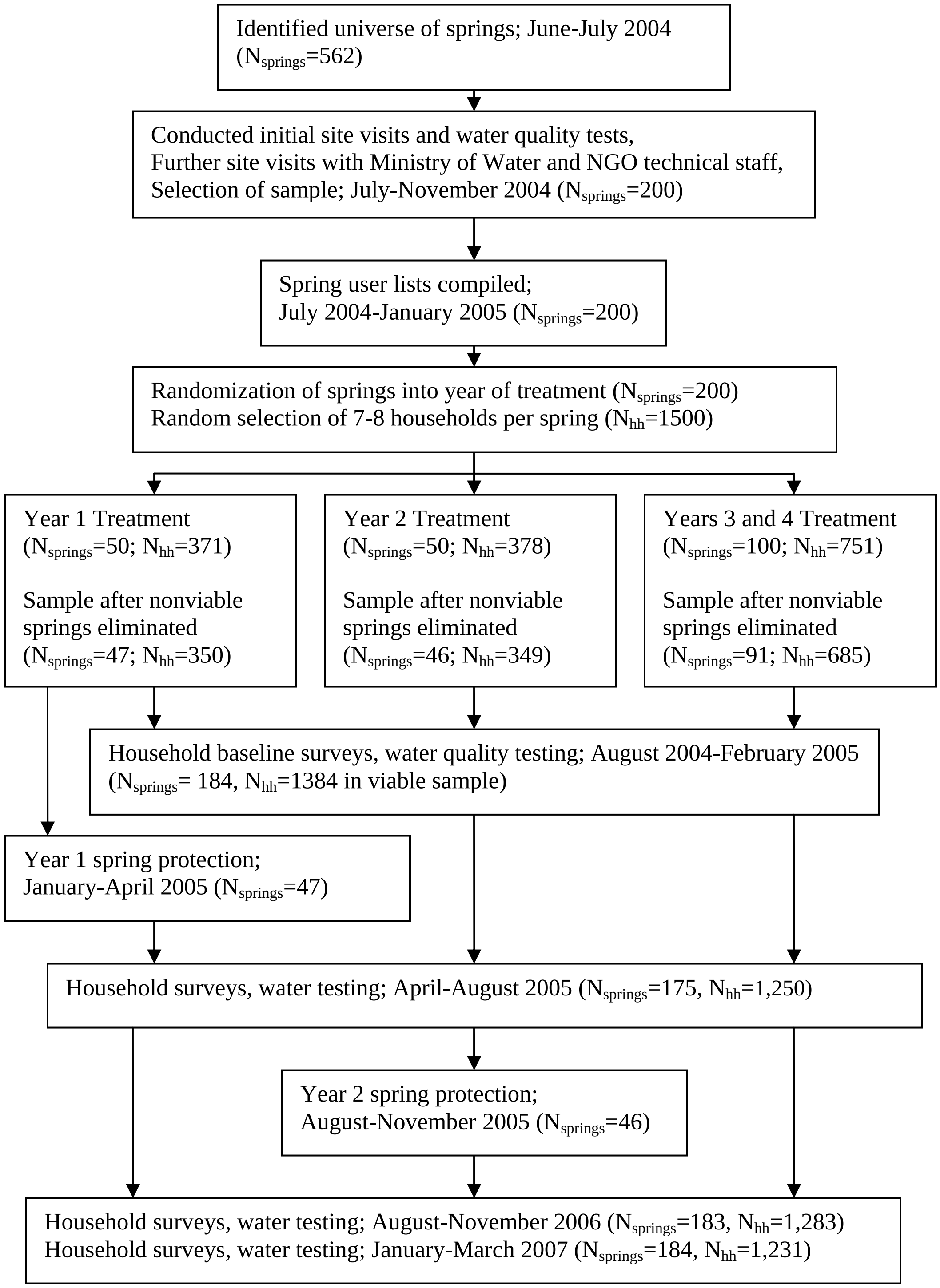
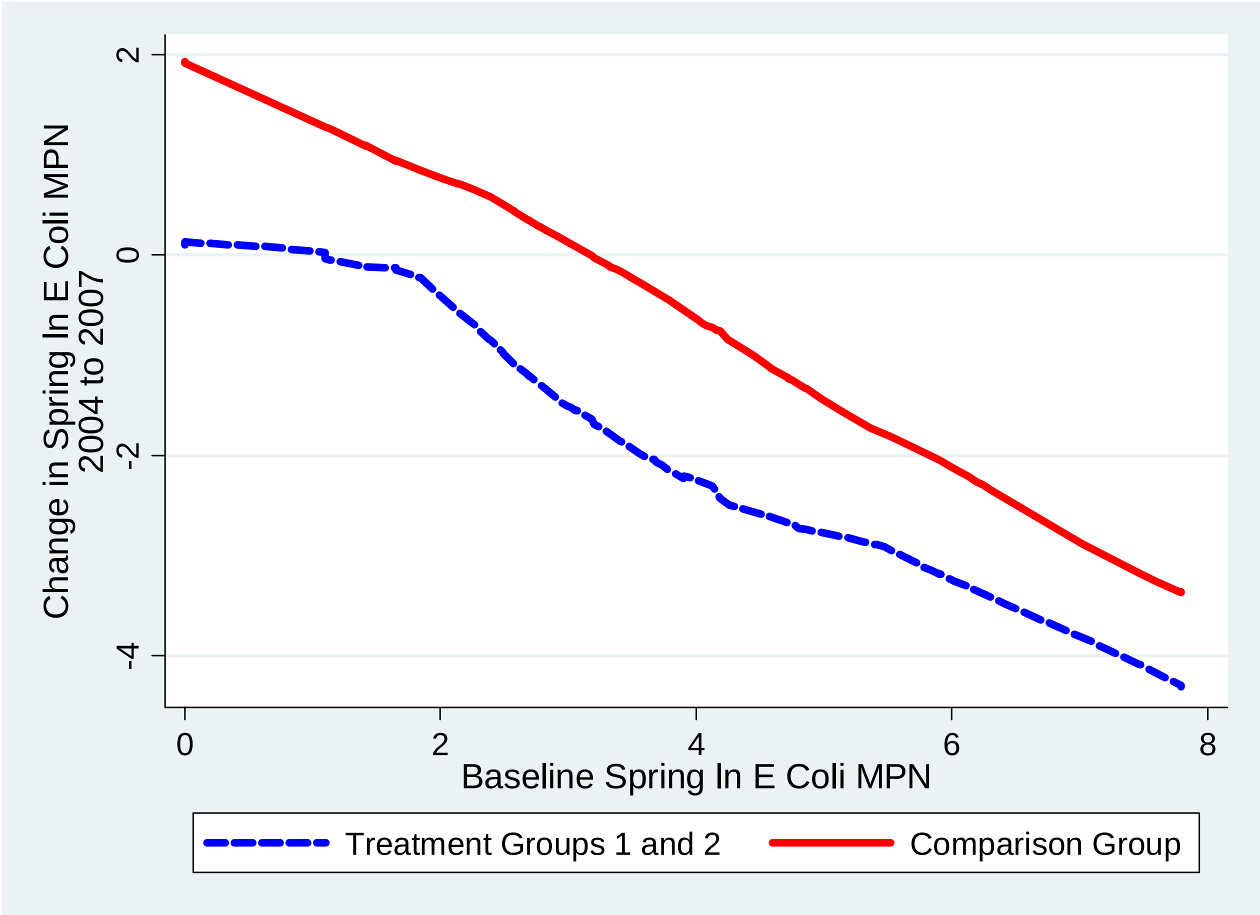
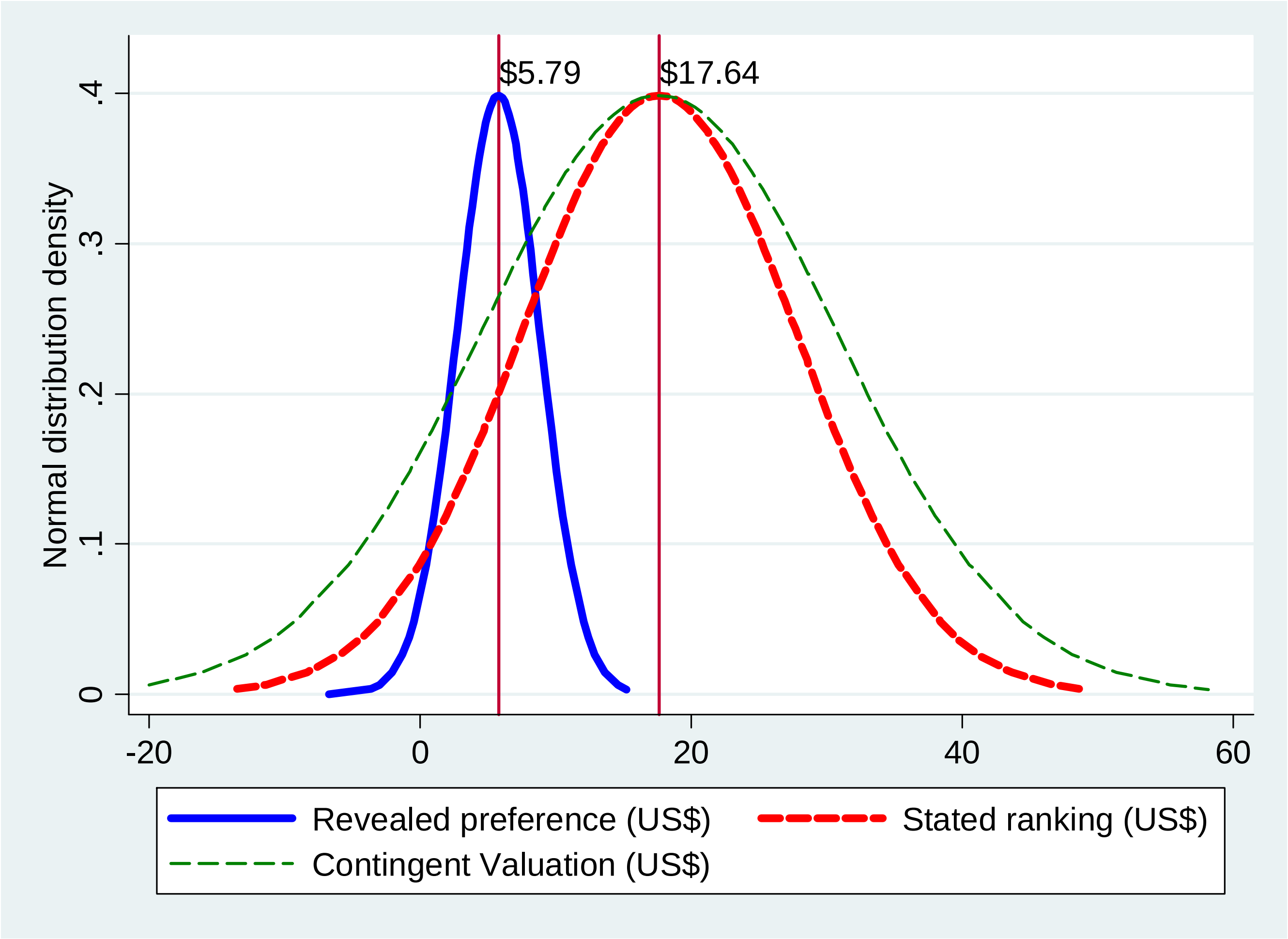


Figure 3: Change in spring water contamination from 2004 to 2007 versus baseline (2004) water contamination



Notes: To 10-90 range in Baseline ln (*E Coli* MPN) is [1.1, 6.3]. MPN stands for "most probable number" coliform forming units (CFU) per 100ml.

Figure 4: Household revealed preference and stated preference valuations of one year of spring protection (2007)



Notes: The revealed preference estimates are from the mixed logit results in Table 6, regression 5, and the stated preference ranking results are from the mixed.logit results in Table 6, regression 7. The contingent valuation data are presented in Table 7, Panel C.