Environment for Development

Discussion Paper Series May 2015 ■ EfD DP 15-09

The Costs of Coping with Poor Water Supply in Rural Kenya

Joseph Cook, Peter Kimuyu, and Dale Whittington

Sida

EfD
Environment for Development

RESOURCES FOR THE FUTURE

Environment for Development Centers

Central America
Research Program in Economics and Environment for Development in Central America
Tropical Agricultural Research and Higher Education Center (CATIE)
Email: centralamerica@efdinitiative.org

Chile
Research Nucleus on Environmental and Natural Resource Economics (NENRE)
Universidad de Concepción
Email: chile@efdinitiative.org

China
Environmental Economics Program in China (EEPC)
Peking University
Email: china@efdinitiative.org

Ethiopia
Environmental Economics Policy Forum for Ethiopia (EEPFE)
Ethiopian Development Research Institute (EDRI/AAU)
Email: ethiopia@efdinitiative.org

Kenya
Environment for Development Kenya
University of Nairobi with
Kenya Institute for Public Policy Research and Analysis (KIPPRA)
Email: kenya@efdinitiative.org

South Africa
Environmental Economics Policy Research Unit (EPRU)
University of Cape Town
Email: southafrica@efdinitiative.org

Tanzania
Environment for Development Tanzania
University of Dar es Salaam
Email: tanzania@efdinitiative.org

Sweden
Environmental Economics Unit
University of Gothenburg
Email: info@efdinitiative.org

USA (Washington, DC)
Resources for the Future (RFF)
Email: usa@efdintiative.org

The Environment for Development (EfD) initiative is an environmental economics program focused on international research collaboration, policy advice, and academic training. Financial support is provided by the Swedish International Development Cooperation Agency (Sida). Learn more at www.efdinitiative.org or contact info@efdinitiative.org.

Contents

The costs of coping with poor water supply in rural Kenya

Joseph Cook*[1], Peter Kimuyu[2], and Dale Whittington[3]

[1]*Associate Professor, Daniel J. Evans School of Public Affairs, University of Washington*
[2]*Professor, School of Economics, University of Nairobi*
[3]*Professor, Depts of Environmental Sciences & Engineering, and City & Regional Planning, University of North Carolina at Chapel Hill; and Manchester Business School*

May 2015

Abstract As the disease burden of poor access to water and sanitation declines around the world, the non-health benefits – mainly the time burden of water collection – will likely grow in importance in sector funding decisions and investment analyses. We measure the coping costs incurred by households in one area of rural Kenya. Sixty percent of the 387 households interviewed were collecting water outside the home, and household members were spending an average of two to three hours doing so per day. We value these time costs using an individual-level value of travel time estimate based on a stated preference experiment. We compare these results to estimates obtained assuming that the value of time saved is a fraction of unskilled wage rates. Coping cost estimates also include capital costs for storage and rainwater collection, money paid either to water vendors or at sources that charge volumetrically, costs of treating diarrhea cases, and expenditures on drinking water treatment (primarily boiling in our site). Median total coping costs per month are approximately US$20 per month, higher than average household water bills in many utilities in the United States, or 4.5% of reported monthly cash income. We estimate that coping costs are greater than 10% of income for one-quarter of households in our sample. They are also higher among poorer households. Even households with unprotected private wells or connections to an intermittent piped network spend money on water storage containers and on treating water they recognize as unsafe.

*Corresponding author: email: *jhcook@uw.edu*. We thank Annalise Blum, Josephine Gatua, Mark Mwiti, and John Wainana for valuable assistance in the field and in data analysis. Funding for the project was provided by Environment for Development - Kenya.

The costs of coping with poor water supply in rural Kenya

Abstract As the disease burden of poor access to water and sanitation declines around the world, the non-health benefits – mainly the time burden of water collection – will likely grow in importance in sector funding decisions and investment analyses. We measure the coping costs incurred by households in one area of rural Kenya. Sixty percent of the 387 households interviewed were collecting water outside the home, and household members were spending an average of two to three hours doing so per day. We value these time costs using an individual-level value of travel time estimate based on a stated preference experiment. We compare these results to estimates obtained assuming that the value of time saved is a fraction of unskilled wage rates. Coping cost estimates also include capital costs for storage and rainwater collection, money paid either to water vendors or at sources that charge volumetrically, costs of treating diarrhea cases, and expenditures on drinking water treatment (primarily boiling in our site). Median total coping costs per month are approximately US$20 per month, higher than average household water bills in many utilities in the United States, or 4.5% of reported monthly cash income. We estimate that coping costs are greater than 10% of income for one-quarter of households in our sample. They are also higher among poorer households. Even households with unprotected private wells or connections to an intermittent piped network spend money on water storage containers and on treating water they recognize as unsafe.

1 Introduction

Many people are not aware that childhood mortality rates are declining in developing countries, and that this decline is accelerating. The global decline from 1990 to 2010 is 2.1% per year for neonatal mortality, 2.3% for post-neonatal mortality, and 2.2% for childhood mortality (Rajaratnam et al., 2010). Across 21 regions of the world, rates of neonatal, post-neonatal, and childhood mortality are declining, and in 13 regions of the world, including all regions in sub-Saharan Africa and there is evidence of accelerating declines from 2000 to 2010 compared with 1990 to 2000.

Although the causes of these declines are not precisely clear, they have important implications for the economic appraisal of investments in improved water and sanitation services, especially piped network services. The economic benefits of improved water services consist of both health and nonhealth benefits. Estimates of the health benefits have two parts: people's valuation of the mortality and morbidity reductions due to improved water services. The nonhealth benefits include time savings from not having to walk to collect water from sources outside the home, reductions in averting or defensive expenditures, and aesthetic and lifestyle benefits. To the extent that baseline childhood mortality rates are lower, the health benefits of improved water and sanitation services will be lower, and the relative importance of non-health benefits (i.e., as a proportion of the total benefits) will increase. Rising household incomes also increase the magnitude and relative importance of nonhealth benefits.

In the past, most international donors and national governments have justified investments in improved water and sanitation services on humanitarian and moral grounds, based primarily on improved health outcomes. In the future, as the economic benefits from nonhealth outcomes become increasingly important in the appraisal of investments in new water supply and sanitation systems, water sector professionals will need to focus more attention on the magnitude and economic value of time and cost savings, and they will require improved tools and approaches to estimate these nonhealth-related benefits.

This paper provides estimates of the coping costs that households incur to obtain water supplies in one rural area in Kenya. We show that household coping costs are surprisingly large. Although we only measure one component of health benefits (reductions in diarrhea treatment costs), our results suggest that most of the benefits that would result from improved piped services in the home would be time savings. We also show that there is great heterogeneity in coping costs across households, both due to differences in the amount of time spent collecting water and the value of time saved. Rather than rely solely on assumptions about the value of travel time based on a fraction of wages, we use values predicted from a latent class model estimated with data obtained from a stated preference experiment. Because enumerators recorded GIS locations of houses and water sources, we are able to test the sensitivity of estimates of time spent collecting water based on reported and observed distances, the first such comparison we are aware of in the coping cost literature.

In the first section of the paper we provide a theoretical framework of household choices regarding water and health interventions that illustrates the different components of the total economic benefits of such investments. We use this theoretical framework to illustrate the coping cost component of the total economic benefits from a water intervention. The second section describes our study site in Kenya, the fieldwork and data collection activities, and the demographics of households in the study area. The third section describes households' existing water supply situations. In the fourth section we present our procedures for calculating households coping costs and the assumptions we make. The fifth section presents the results and in the sixth section we offer some concluding observations.

2 Theory and Existing Literature

To better understand households' behavioral decisions around coping with unimproved water supply, we begin with a household production function approach, drawing heavily from the one described in Pattanayak et al. (2005) and Pattanayak and Pfaff (2009). The goal is two-fold: to illustrate how coping costs are related to the total economic benefits of "improved" water supply, and to generate some simple theoretical predictions for how coping costs may be related to characteristics of households. We conclude this section with a discussion of other empirical coping cost estimates for water supply.

A household production function approach treats a household as combining exogenous and endogenous inputs into some output or good that households value, conditioned on that household's tastes and preferences. An improved water supply – defined here simply as one that provides water reliably, of potable quality, and of sufficient quantity to meet basic household needs like drinking, bathing, cooking, and washing around the house – is not that output. Rather, households value the good health and other amenities that water supply provides, namely, quenched thirst, clean homes and clothes, and perhaps productive and attractive kitchen gardens. These services provide households utility (eq. 1), which they maximize given their limited total time resources (eq. 2) and financial resources (eq 3):

$$U = (Z, T_l, W(G), S(C(T_c, M)), \theta) \tag{1}$$

$$T = T_w + T_c + T_l - T_s \tag{2}$$

$$I \geq N + wT_w - pM - Z] \tag{3}$$

Households gain utility from consuming a composite good Z, leisure time L_l, and water supply W. Water supply is a function of government policies G, such as infrastructure provision. Disutility comes from poor health and sickness S, which households can reduce by investing time coping T_c, for example by walking further to collect water from a protected source or by spending time treating water, and financial expenditures M on water treatment devices, higher prices at protected sources, or other capital costs. The parameter θ stands in a broad-brush way for heterogeneity in the tastes and preferences of households. Some may value leisure time more highly, for example, or place a higher priority on improving health (lowering S). Households allocate their total time T among leisure, time coping, and time spent earning wages T_w; they lose time T_s to illness (we assume a "unitary" household decisionmaking structure). Finally households are constrained by their budgets. Households must limit their total expenditures to the sum of income N not earned using time or labor (i.e. remittances, financial returns on investments) and their earnings from wage labor at a rate of w per unit time. Households spend money on the composite good Z (with a price normalized to one) and financial coping actions M which have a unit price of p.

Households maximize utility by deciding how much time to devote to wage labor T_w and water collection T_c , how much of the composite good Z to consume, and how many investments or expenditures to cope with poor water supply (M, at exogenous price p). Notice that $W(G)$ is exogenous to households; they cannot choose their current water supply regime. Only government policies affect the baseline water supply situation in our model. As W(G) changes, so too would the optimal decisions households choose to make about time, consumption and financial coping mechanisms. An important policy question, then, is: how valuable to households is a change in water supply regimes, say from the current status quo $W^0(G_0)$ to an improved situation $W^1(G_1)$ where the government or a regulated operator invests money to extend (or repair) a network, piped system? In other words, what is the total household willingness-to-pay (WTP) for this improvement? Using duality theory, we can define an "quasi-expenditure function" Ω that gives households the exact same utility U^0 that they had under the status quo water condition, and again optimize (take the first derivative) of this function (see Pattanayak et al. (2005) for a longer derivation). The household's total economic benefits, ${}^{\delta\Omega}/_{\delta W}$, of changing water supply situation W is:

$$\frac{\delta\Omega}{\delta W} = \left[p\frac{\delta M}{\delta W} + \frac{\delta T_c}{\delta W}\right] + w\frac{\delta S}{\delta W} - \lambda\frac{\delta U}{\delta S}\frac{\delta S}{\delta W} \tag{4}$$

The term ${}^{\delta\Omega}/_{\delta W}$ is the amount of income N that one could take away from a household at water situation W^1 and leave it with exactly the same utility as they had at water situation W^0. The three terms of this willingness-to-pay function are 1) the coping costs in terms of financial expenditures and lost wage income from time spent coping; 2) the costs of illness; and 3) any monetary value associated with pain and suffering from poor health S that is not associated with lost labor hours. One prominent component of this term would be the value households place on reducing mortality risk. Our analysis of coping costs will incorporate (theoretically) the first two terms but cannot capture the third, and is thus a lower bound on the total economic benefits of improving water supply. There are a number of other factors that may drive a further "wedge" between coping costs and total WTP, such as intrahousehold time allocation, suboptimal coping choices, possible income effects from improved water, and other factors (we again refer the reader to Pattanayak et al. (2005) for more discussion). Note also that ${}^{\delta}/_{\delta M}$ is the marginal utility of money, and if one makes the common assumption that the marginal utility of income is declining, one prediction from the model would be that coping costs increase with income, all else equal. Although the term "coping costs"' is sometimes used synonymously with the terms "averting" or "defensive" expenditures in the environmental economics literature, our "coping costs" are not simply averting expenditures. Rather, they combine both expenditures made defensively *ex ante* – boiling water to make it safe to drink

– with those made *ex post*, for example on treatment of diarrhea. The latter category are often called "cost-of-illness" studies in the health literature.

A number of studies have examined the strategies that households use to cope with poor water supply, though many have studied the treatment behavior of tap water consumers in industrialized countries where water collection time is very small (Larson and Gnedenko, 1999; McConnell and Rosado, 2000; Abrahams et al., 2000; Um et al., 2002). Furthermore, most existing studies in low-income countries have focused on urban customers, where intermittent supply tends to draw researchers' focus again to water treatment choices as well as water storage investments (Whittington et al., 1990; Zérah, 2000; Pattanayak et al., 2005; Jalan and Somanathan, 2008; Vásquez et al., 2009; Katuwal and Bohara, 2011; Vásquez, 2012; Nganyanyuka et al., 2014). Of these studies, we could find only two that included empirical estimates of the time spent traveling and waiting for water (Whittington et al., 1990; Pattanayak et al., 2005). Estimates of the total coping costs as a percent of income were not reported in all studies; they were 0.78% in Leon, Nicaragua (Vásquez, 2012), 1% of current income in Kathmandu, Nepal (Pattanayak et al., 2005), and 7.5% of income (primarily on bottled water) in Parral, Mexico (Vásquez et al., 2009). The evidence for coping strategies in rural areas - where most of the remaining water supply coverage gap lies – is quite thin. We know of only three studies examining coping strategies in rural areas (Pattanayak et al., 2010; Kremer et al., 2011; Jessoe, 2013), though there have been a number of studies using stated preference methods to measure willingness-to-pay for improved water supply directly. The focus of Jessoe (2013) is solely on water treatment choices in rural India, testing whether switching to protected, "improved" sources leads households to discontinue treating water, reducing expenditures but offsetting some of the quality benefits of protection. Kremer et al. (2011) uses both revealed and stated preference methods to estimate the value of source protection in rural Kenya; although coping strategies, especially water treatment and water collection times, are a central component of the study, coping cost estimates are mentioned only briefly. The best available rural estimates are from Pattanayak et al. (2010), who evaluated a community demand-driven water supply program in four districts of Maharashtra, India. The authors measured time costs of both water collection and poor sanitation, costs of illness, storage costs and treatment costs in panel data from nearly 10,000 households. Time costs comprised over half of total coping costs (Pattanayak et al., 2010, p.538).

3 Study Site and Demographics

We interviewed a total of 387 households near the small market town of Kianjai in September 2013, the dry season. Kianjai is approximately 20 miles from the city of Meru, in north-central Kenya. The study site was chosen purposefully because of the large number of existing water source options available (the work was part of a larger study on rural water source choices). The site was *not* chosen because we believed *a priori* that it had a water collection burden that was higher or lower than an average rural town in Kenya. Sample households were chosen randomly based on a transect approach. We provide more details on the sampling approach in Appendix A1.2. A team of seven trained enumerators asked households a number of detailed questions in Meru (the local language) about the water sources that households use during both the dry season and the rainy season[1], including water uses, prices paid, treatment of drinking water, and others. We interviewed the household member "who is mostly responsible for water-related decisions such as where to get water and how much to collect." In three quarters of the households, this person was also the one "who collected the most water in the past seven days."

A typical sample household is Catholic and has five members. The household is led by a married couple, both of whom are around forty years old and have each completed seven years of education. Eighty-one percent of households have a child under age 15; the average household has 1.8 children. Twenty-four percent of households have a child aged two or younger, and half have a child aged five or younger. They own their house and two acres of land. The household has a private pit latrine, but does not have electricity. Kerosene is used for lighting and firewood is used for cooking and heating. There are two rooms in the main

[1]There are two rainy seasons in central Kenya, one roughly from mid-March to June, and another shorter season from mid-October to early December.

house and three other buildings in the compound. Monthly household income from all sources is approximately 35,000 Ksh or 407 USD. The most common source of income by far is from farming. Thirty-nine percent of households, however, had at least one household member who earned income from full-time employment, part-time or seasonal employment, or business and self-employment. About 10% of households had more than one member earning income from these sources. Average food expenditure is 430 Ksh (5 USD) per household member per week or a total of 14,924 Ksh (174 USD) per month. Household assets include a cell phone (93% of sample households), bicycle (76%), and radio (82%). Thirteen percent own a motorbike. In terms of livestock, the typical household has four chickens, two goats, and two cows. Using data on durable assets, electricity connections, sanitation, building characteristics, and cooking fuel, we construct a wealth index using principal component analysis (PCA) following Filmer and Prichett (2001) and Filmer and Scott (2012), and assign households to wealth quintiles. Although water supply variables are often included in wealth indices, we exclude them to avoid potential confounding with other explanatory variables. (Appendix A1.3 provides more details on the construction of the wealth index.)

4 Water collection behavior

In this section, we describe the patterns of household water collection activities common in the study site. Households in our sample reported that they *could* use an average of 4.2 sources (median 4; range 2-7 sources); 91% of the sample falls in the range of 3-5 sources. They reported they *do* use an average of 1.9 sources (within the past 12 months); 68% of respondents said they had used two or more sources in the past 12 months.

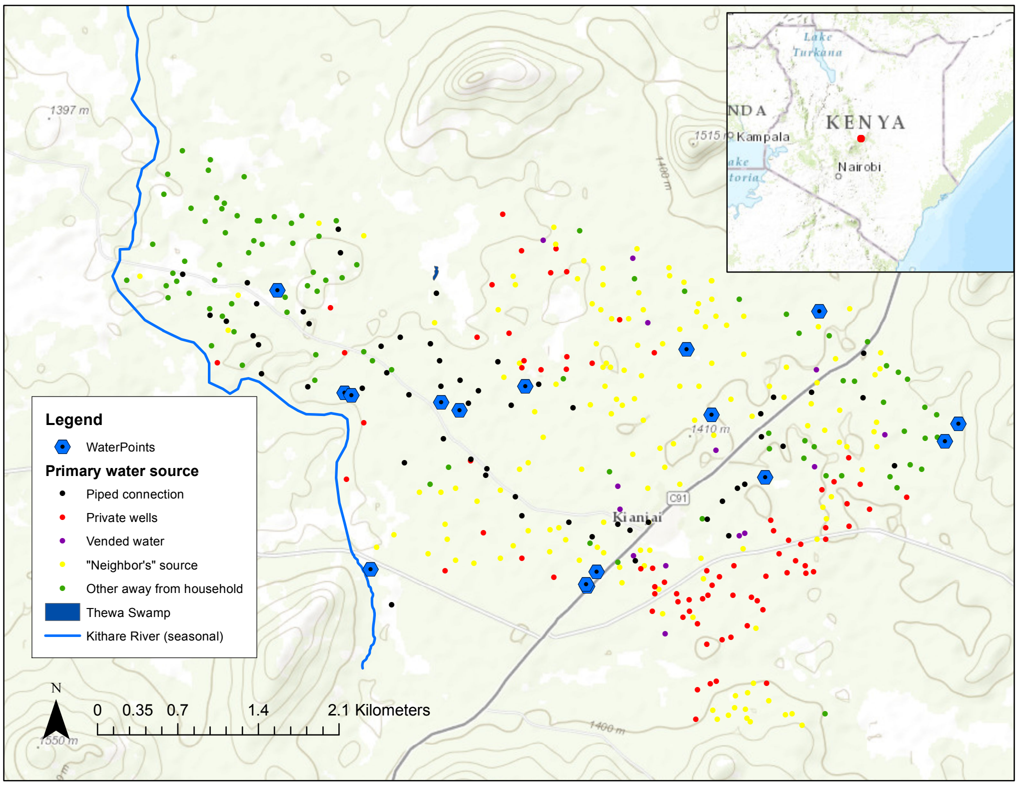
Because households may collect water from different sources to serve different purposes, we asked which water source the household primarily uses for different purposes, including drinking, washing around the house, cooking, bathing/personal hygiene, watering animals, and other productive activities. We also asked which source was their "primary" source for "most purposes". All but 11 respondents (2.8%) reported that their "primary" water source for these various purposes was the same, indicating that most households rely primarily on one source and use others as occasional or back-up sources. In other words, households do not dedicate water from different sources (perhaps with different qualities) to different water uses. We therefore organize our discussion of households by respondents' reported primary source for most purposes.

Figure 1 shows water sources and households in the study site, the latter grouped by primary source. The figure illustrates several features of the study site. First, a piped distribution network operated by a formerly-public, now-private water company (Imetha Water and Sanitation Company, IWASCO) serves the area. The system used to supply piped service to many households until the distribution network fell into disrepair in the 1990's and the raw water supply from the mountains east of the map became over-allocated. Many of the households in our sample *without* water supply at home were once served by this system and showed us their yard taps that were no longer working. IWASCO is now in the process of rehabilitating the network and expanding the raw water intake, but at the time of our fieldwork the households that had working IWASCO connections were clustered along the one distribution line which first ran along the main road and then branched off, running northwest from the main road.

Another group of households have piped connections to what is locally called "project" water. These are self-organized, self-financed distribution networks that typically divert untreated river water. Some households contribute labor and some cash for the construction and operation of these schemes. Once the pipes reach near to one's homestead, the household is expected to buy the connecting pipes in order to access the water. There is no volumetric charge for using the system, but households make periodic monetary contributions when the system breaks down and repair or maintenance is required. These small systems typically have little system storage, so the system is only as reliable as the surface water supply.

The study site's hydrology is reflected in Figure 1. Approximately 40% of our respondents reported that a well (either their own or a neighbor's) was their primary source, and these

Figure 1: Households and water sources in study site, grouped by “primary” source



are clustered where groundwater is relatively accessible, predominantly southeast of the main road (red dots show households whose primary source is their own private well). The northwest and northeast sections of the study area have less accessible groundwater, and households there are more likely to walk to water, either to a neighbor’s source (yellow) or to a source away from the household (green).

Table 1: Water collected from **all** sources, organized by the household's "primary source"

	Dry season			Rainy season	
	Median liters per capita per day	Median Monthly Collection (m^3)	Average % of all water collected from prim. source	Median Monthly Collection (m^3)	Average % of all water collected from rainwater
Private piped(60)	40	4.8	91%	4.8	67%
Private well(78)	42	7.3	96%	6.0	56%
Vended(15)	28	3.7	80%	3.4	76%
Public well(122)	25	4.2	75%	5.0	69%
Public borehole(72)	27	4.1	72%	5.1	70%
Public piped conn(32)	20	3.9	85%	4.2	75%
Surface, other public(6)	27	3.7	58%	4.7	69%

Notes: Numbers in parentheses in the first column refer to the number of households (of a total of 387) who reported that this type of source was their "primary" source for "most purposes". For example, 60 households said a piped connection was their primary source.

Thirty-six percent of households in our sample say their primary source is at home, either private piped connections, rainwater collection systems, or private wells (Table 1). Another 37% report that using a "neighbor's" source is their primary source. The remaining quarter of the sample rely primarily on water from vendors (n=15, 4%), or walk to public water points that we attempted to geolocate. Table 1 reports the total volumes (from all sources; not just the primary source) collected both in terms of monthly cubic meters and liters per capita per day. Because of concerns for both recall problems as well as day-to-day fluctuations in water collection behavior, we asked how much water was collected in the past 7 days as well as a "typical" week in the dry season and rainy season. The dry season calculations in Table 1 are based on collection data for the "past 7 days". As a point of reference, the World Health Organization's daily "minimum" is 40 liters per capita per day (LCD). Households with piped water are abstracting just this minimum, mainly because the system does not provide anything near 24-hr, 7-day service. The median household with a piped connection receives water for 12 hours per day, 3 days per week. Households with private wells are abstracting about the same amount per capita.[2] Households who travel for water ("neighbor's" and public sources) are collecting 25-

[2]Households with private wells abstract more in total than those with private connections (third column)

35 LCD. On average, households reported using approximately 72% of the water collected for washing around the home, bathing and cooking, and (plausibly) 2% (or 14 liters per household per week) for drinking. Eighty-five percent reported collecting water for animals. The median collected for animals is 40L per household per day, so this is likely for smaller livestock like chickens and goats and perhaps a supplemental supply for larger livestock.

Nearly all (96%) of respondents said they could and do use rainwater during the rainy season, and total volumes collected from all sources increase during the rainy season among households without water at home (Table 1). One-third of these households stored water in small buckets, but the remaining two-thirds had invested in larger storage containers (greater than 50L). The median household storage capacity was 200L. Only two households reported having sufficient storage capacity to use rainwater as their "primary" source. These two households had storage capacities of 9 m^3 and 18 m^3. Nevertheless, most households report that a majority of their water comes from rainwater during the rainy season (last column, Table 1).

We asked households to rate the sources they *could* use in terms of health risk of drinking from the source ("no risk", "some risk", or "serious risk"), taste of water from the source ("poor", "normal", "sweet", or "varies"), and whether using the source is likely to lead to a conflict with neighbors ("not likely at all", "somewhat likely", "very likely"). For any source that the household had actually used in the past twelve months - even if it was not their primary source - we asked whether respondents treated water before drinking. These results are shown in Table 2. Respondents judged the quality of water from piped systems to be poor. This is not surprising because the supply from these systems is intermittent. Sixty-one percent of respondents with working piped connections thought that drinking water posed some health risk, roughly similar to other households' perceptions of their *neighbor's* piped connection or a public tap on a piped system (Table 2). These respondents were more satisfied with the taste of the water than respondents using sources outside the home. Respondents also reported that water sources are a widespread cause of social conflict in the area (last column of Table 2). Three-quarters of respondents who said they could use a public well or tap thought using it

because of larger household sizes. The median household with a piped connection has four members versus six members for households with a private well.

would be “somewhat” or “very” likely to lead to conflict; so did 40 to 60 percent of respondents using a neighbor’s source. Among well-owners themselves, 85% reported allowing their neighbors to use the well and 28% of those said that sharing had led to conflict with neighbors (24% overall).

Table 2: Perceptions of water source characteristics among those who *could* use that source, and water treatment behavior

	Percent: *Some risk* from drinking water	Percent: *Serious risk* from drinking water	Percent: Water tastes poor	Percent: Using source likely or very likely to lead to conflict	Percent who treat water (if used source in past year)
Private piped	0.49	0.12	0.18	.	.
Rainwater	0.28	0.011	0	.	0.0081
Private well	0.44	0.25	0.33	0.24	0.54
Vended	0.51	0.27	0.31	.	.
Public well	0.56	0.27	0.38	0.65	0.39
Public borehole	0.31	0.084	0.20	0.53	0.28
Public piped conn	0.61	0.082	0.11	0.56	0.20
Surface, other public	0.31	0.40	0.43	0.35	0.17

Notes: Results for questions about health risk from drinking, poor taste, or the likelihood of conflict are based on responses from all households who said they *could* use the source, not only respondents who listed the source as their primary source. Respondents were asked questions about whether they treated water from a source if they had used the source in the past 12 months, even if it was not their primary source. Blank cells indicate that the question was not asked for that particular water source, i.e. respondents were not asked if collecting rainwater would lead to conflict with others.

Although households with private wells have largely solved their collection time and “quantity” problem, they have not solved their “quality” problems, and they are largely aware of this. Figure 2 shows three fairly typical private wells in the study site. They are hand-dug, i.e. not machine-drilled, and are fairly shallow; the median well depth as estimated by our respondents is 30 feet. The well at left is considered “protected” since it has a raised concrete cap and steel lid to prevent surface contamination, though the rope and bucket could quickly recontaminate the well. The wells in the middle and at right are both uncapped and prone to surface contamination; they are also both safety hazards for small children. Among the households with private wells, 69% thought that drinking water from the well would pose

Figure 2: Private, hand-dug shallow wells in the study site

"some risk" or a "serious risk" (Table 2). Among those using a neighbor's well or a public well, 82% and 87% similarly thought this drinking water source was a risk. In all of these groups, roughly a third thought the risk was "serious". Accordingly, 54% of those with private wells reported treating their water (discussed more below in Section 5), and 33% thought the water tasted poor. Thirty-six percent of those using a neighbor's shallow well and 44% of those using a public well reported treating water.

For households hauling water away from home, who is doing the collection work? These households have an average of 2.1 people who have collected water in the past 7 days. We asked which person collected the most water in the past 7 days. This person was a woman in 76% of households, with an average age of 34 years old (the full distribution is shown in Appendix Figure A3). Only 11% of these "primary" water collectors had paid employment outside the home.

5 Coping cost components

In this section we present the methodology and results for each individual component of coping costs: the time costs of collecting water outside the home, financial costs paid for vended water or at paid sources, amortized capital costs for storage and rainwater collection, diarrhea treatment costs, and money spent treating drinking water. Section 6 then reports total coping costs (the sum of these five categories), explores the pattern of coping costs by decile of self-reported income and asset/wealth quintile, and explores correlations between household characteristics and coping costs.

5.1 Collection Time

We use responses from several questions to calculate the time spent collecting water from outside the home. First, we asked respondents how long it would take to walk to a water source, one-way, with a full 20L "jerrican"[3]. Given that walking with an empty container should be faster, these times are multiplied by 1.75 to get roundtrip walking times.[4]. Households walking to get water reported an average roundtrip walking time of 43 minutes (median 35 minutes) for sources that they have used in the past year (Figure 3, Panel A). There is considerable variation, however: the standard deviation is 43 minutes and the histogram shown in Figure 3 omits 13 households who reported roundtrip walking times in excess of two hours. Panel A clearly shows modal responses corresponding to one-way times of 5, 10, 15, 30 (i.e. roundtrip = 30 * 1.75=52.5 minutes) and 60 minutes.

Because enumerators recorded GIS locations of houses and water sources, a second approach uses predicted walk times for those observations where we have GIS data. These are calculated with straight-line distances in GIS for 300 observations of household-water source pairs. Our results imply average walking speeds, encumbered with a 20L container of 2.75 kph or 1.7 mph. These are somewhat lower than other reported walking speeds in rural Africa, though most published studies do not distinguish between walking with and without the weight of jerricans (See Appendix A1.4 for more details and discussion). We do not have geolocations for "neighbor's" sources, however, which were reported by a significant fraction of households. A third approach is to supplement the GIS-predicted times with median reported distances (in minutes) to a "neighbor's" source by sublocation (to account for spatial differences in the den-

[3]The term jerrican is commonly used to denote rectangular, hard plastic water storage containers, the vast majority of which hold 20 liters of water.

[4]We are unaware of estimates comparing walking speeds with and without the burden of a 20L (20 kg) jerrican. See Appendix A1.4 for more discussion of this assumption, our distance calculations, and a comparison of walking speeds in the literature

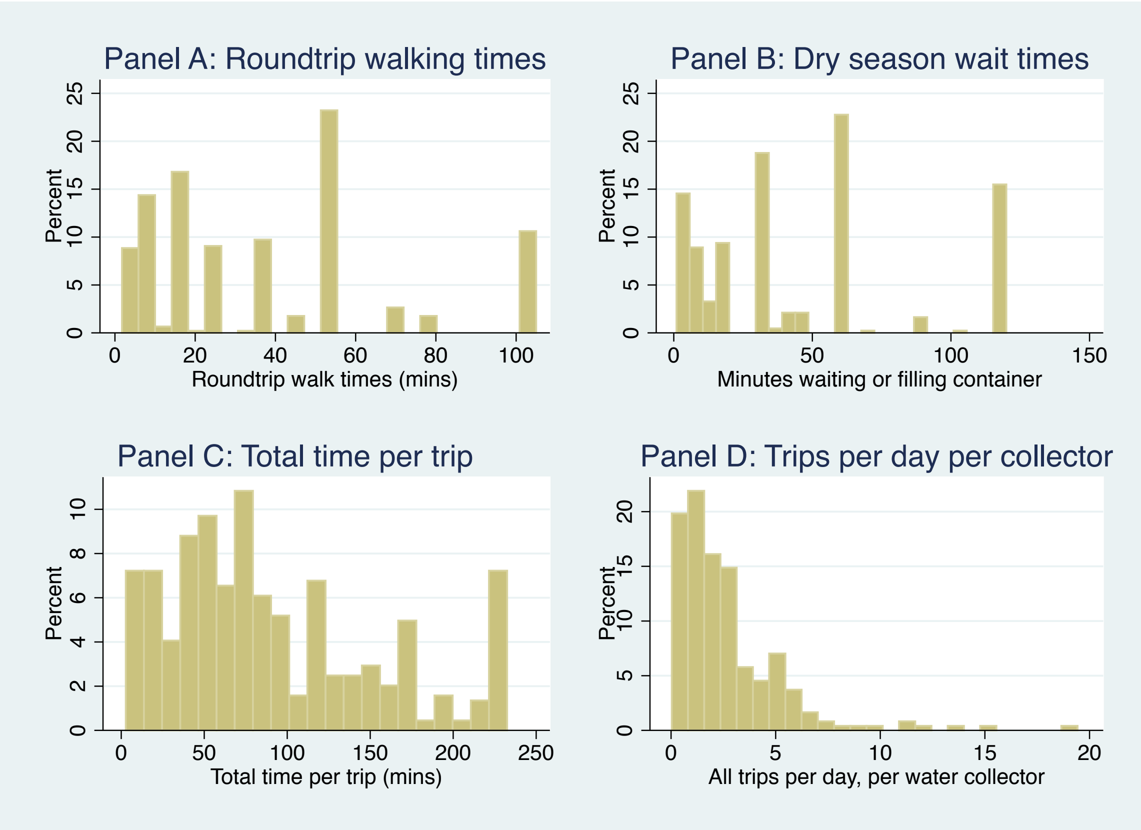
sity of houses). Rather than mixing reported and predicted distances, for clarity we focus primarily on using unadjusted but truncated times, described in more detail below. We provide results using the other approaches as a sensitivity analysis. We also cannot rule out that households may be combining activities on these trips, such as socializing or visiting the market.

Forty percent of households (n=165) reported collecting water by modes other than walking. The majority of these used bicycles, typically to carry one or two jerricans, though forty households used carts or wheelbarrows to haul larger quantities of water. One household used a motorbike, and three reported using a car to collect water. We did not ask about how many minutes it would take to travel to a source using these different modes. Roundtrip time for bicycles are assumed to be half the reported walking time, and the time for cars and motorbikes to be one-quarter the walking time.

We asked households how long they spent waiting in a queue or filling their container during an average week in the dry season and in the rainy season. These average 55 minutes (median 30 minutes) in the dry season (Figure 3, Panel B). Panel B again shows modal responses at 30 minutes, one hour and two hours, and omits 13 observations where respondents said a typical wait exceeded two hours. Although these reported times may be subject to recall bias and in some cases exaggerated, we cannot independently verify them without direct observation of sources. In a sensitivity check, we use median reported wait times (by source) to deal with the large outliers. We also do not observe whether households are able to leave the line – putting their jerrican in line as a placeholder – and do other things with their time. Reported wait times plummet to an average of 7 minutes (median 5 minutes) during the rainy season. There is no evidence that people who have more experience with a water source are more accurate in their estimate of wait times: the pair-wise correlation between wait times in the dry season and whether it is their primary source is small (0.035) and not significant at the 10% level. Figure 3 (Panel C) shows the distribution of total times per trip for all households collecting water from a source outside the home, combining the walking and waiting data.

Households also reported the number of trips they took to collect water in the past seven days, during an average dry season week and during an average rainy season week. We focus on

Figure 3: Reported walking and waiting times for sources away from home that were used by the household in the past 12 months

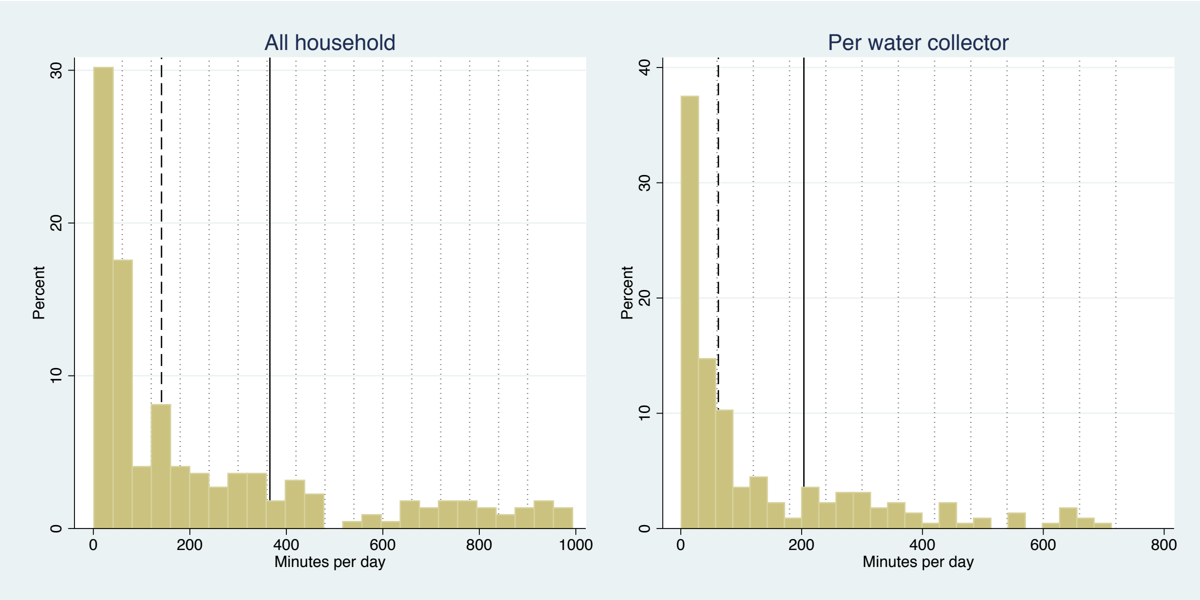


Notes: Panels A-C are based on 465 household-source combinations where the household reported using the source in the past 12 months. Panel A omits 13 observations exceeding two hours. Panel B omits 33 observations exceeding two hours per trip. Panel C excludes 32 observations exceeding four hours per trip. Panel D is based on data reported by households for the "past seven days" (n= 242 hhs getting water outside the home at least once in past 7 days) and includes trips on foot as well as by bicycle or other means.

trips made in the past 7 days, which are least likely to suffer from recall problems. Households who made trips on foot reported making an average of 4.8 trips per day (median 4.3) in the previous 7 days; households who made trips by other means made 2.2 trips per day (median 1.9). Twenty-nine households reported making more than 10 trips per day. Because some households may have more labor available to haul water, we calculate the total trips per day per water collector (Figure 3, Panel D). Households who did not have a primary source at home had an average of 2.2 household members who had collected water in the past week. Even after adjusting for water collectors, however, Panel D still shows a fraction of very high reported numbers of trips per day. Among the 25 households taking more than 5 trips per collector per day, one-quarter are collecting nearby, with roundtrip walking times to their primary source of

10 minutes or less. Four of the 25 households, however, report taking more than five trips per day to a source that takes over an hour to walk to and return.

Figure 4: Total collection time (minutes) per day, by household and per water collector



Notes: n=245 households that collected water outside the home and reported a trip within the past 7 days. Histogram on left does not display 23 values over 1000 minutes per day per household; histogram on right truncates 17 values that exceed daylight hours (12 hrs * 60 = 720 minutes). The two black vertical lines show the mean (solid) and median (dashed) collection times. Gray dotted lines show hourly-increments for ease of reference.

We next multiply the total number of trips by the travel and wait times per trip reported above to produce the implied total collection time (Figure 4); these are the key estimates that we multiply by the value of travel time below to get economic costs of time spent hauling water. The figure shows the distribution of daily collection time per household on the left and per water collector on the right. Households spend an average of 6.1 hours per day collecting water, and each water collector spends an average of 3.4 hours per day (these are shown as the solid black vertical lines in the two panels of Figure 4). Both figures show distinct right tails with some households reporting very large total collection times, and the figures do not display 17 households with collection times per collector that imply more than 12 hours per water collector (more than all daylight hours at our equatorial site). These households drive large differences in the mean and median times: median collection times are 2.4 hours per household per day and one hour per water collector (dashed black vertical lines in the two panels). Fifty-six percent of households have daily collection times of three hours or less, and 36% of households have total collection times per water collector of 30 minutes or less.

Rather than truncate values for travel times, wait times, or the number of trips that seem implausible, we truncate a household's total water collection time at what we feel is a reasonable upper bound. One upper-bound would be to assume that households cannot spend more than all 12 daylight hours for each water collector in their household; we would truncate any households with two water collectors at 2*12*60=1440 minutes. Instead, we assume, more conservatively, that water collectors spend no more than 6 hours per day collecting water, which truncates total collection times for 44 out of 245 households that collected water outside the home.

To place an economic value on this time, we use individual-level estimates of the value of travel time (VTT) based on results from a simple stated preference experiment where our respondents were asked to rank new hypothetical water sources that varied only in distance from home and price per jerrican (Cook et al., 2015). The results used here are from a latent-class multinomial logit model. Information criterion test statistics indicated that a model with four classes best fit the data, and we use the individual-level predicted class probabilities to assign respondents to one of these four classes. The first class (44% of respondents) valued time at 49 Ksh/hr. A second class (15%) was relatively unresponsive to the distance to source and has an implied VTT that is very low - 0.8 Ksh/hr - which we round to 1 Ksh for the coping cost calculations. The third class (32%) valued time at 8.6 Ksh/hr and the fourth class (9%) at 9.2 Ksh/hr. We combine the last two classes for our calculations, rounding to 9 Ksh per hour. The latent class model used no data on the household's socioeconomic characteristics, so use of these VTTs should not bias the models below. We did not estimate VTTs from this experiment for households with piped water or rainwater collection systems. We impute the median VTT from the experiments for these households, although this assumption is not critical since collection times for these households is typically very small.

Table 3: Economic costs of time spent collecting water collection time: Sensitivity analysis

	Dry season				Rainy season
	Reported	**Reported, Truncated**[a]	**GIS distances**[b]	**GIS + Medians**[c]	**Reported**
Total household collection times (minutes), per day					
Mean	366	259	367	279	112
Median	141	140	147	172	12
95th percentile	1401	950	1401	980	440
Monthly economic time costs (Ksh) - Individual value of travel times from stated preference					
Mean	4,026	2,978	4,074	3,185	1,130
Median	841	667	836	974	30
95th percentile	18,977	13,508	21,116	14,006	4,140
Median % income[d]	2.53%	2.21%	2.73%	2.67%	1.30%
Monthly economic time costs (Ksh) - Value of travel time = 50% of wages					
Mean	3,201	2,262	3,211	2,422	753
Median	1,236	1,225	1,289	1,508	101
95th percentile	12,262	8,312	12,262	8,571	3,150
Median % income[d]	3.46%	3.45%	3.93%	4.57%	2.75%

Notes: N=245. Households who reported no water collection times – those with piped water or private wells – are excluded from these calculations.[a] Assumes households cannot collect more than 6 hours per day per water collector. [b] Uses predicted walking times using GIS data and a linear fit between reported times and actual straight-line distances (see Appendix A1.4). Uses reported times for "neighbor's" sources where no geospatial data is available, and no changes are made to waiting times or number of trips. [c] Uses GIS-predicted distances where possible and median distances to "neighbor's" sources, calculated for each sublocation. Uses median wait times, calculated by source. [d] Monthly economic costs divided by reported total monthly income (missing in three observations).

Table 3 summarizes the results for the monthly economic cost of water collection under several scenarios; see table notes for descriptions of the four alternate methods of calculating travel times. Only households who reported non-zero collection times are included in the Table, and the estimates are based on responses for the 7 days prior to being interviewed, which occurred in the dry season. Median total daily collection times are approximately two and a half hours per household per day. The average times range from six hours (as reported) to 4.3 hours (truncated). Using GIS-predicted walk times but not adjusting trips, wait times, or distances to "neighbor's" sources does not change the means much, though using median distances and wait times does. Total reported water collection times fall roughly 70% during the rainy season to a mean of 112 minutes per day; the median is 12 minutes per day.[5]

[5]These estimates also truncate 15 households who reported more than 6 hours per collector per day and

Using our latent class value of travel time estimates, the median economic value of this collection time ranges from 667 Ksh to 974 Ksh per month depending on how collection times are treated, or 2.2 - 2.7% of reported monthly cash income. This income measure does not include the value of crops grown by households for their own consumption. It does, however, include self-reported income from farming in an "average month", and 96% of households reported some cash farm income. The economic costs of collection are much lower in the rainy season: the median falls to just 30 Ksh, and the average falls to 1130 Ksh.

For comparison, we also present estimates using a commonly-used benchmark of 50% of the unskilled wage rates (Pattanayak et al., 2005, 2010; Kremer et al., 2011). The field team estimated the unskilled wage rate to be 35 Ksh per hour in our site, implying a "benchmark" value of travel time of 17.5 Ksh/hr. This assumption results in lower average costs and higher medians. There are two factors affecting the comparison of results using the two different VTT estimates. The first is that they obviously imply different VTTs: using the latent class values, the (class-weighted) average VTT is 24 Ksh/hr, or 69% of unskilled wages versus 50% in the benchmark. Other things equal, this should increase total economic costs of water collection. The second factor is the amount of time spent collecting water among the respondents in each class. If they were distributed similarly across the three classes, again we would expect that using the latent class VTTs would increase total economic costs. They are not distributed similarly, however. The average (truncated) daily water collection time among the 142 households predicted to be in class 1, with a VTT of 50 Ksh per hour, is 171 minutes. The average time among the 133 households in class 3, with lower VTTs of 9 Ksh per hour, is 23 minutes longer. Among the 50 households predicted to have a VTT of 1 Ksh per hour, average daily collection times are 250 minutes, or over an hour longer. The median times for these three groups are 25, 55 and 67 minutes (Appendix Figure A1 shows the cumulative distribution of collection times by class.) In other words, the latent class model of stated preferences predicted households with longer water collection times would have lower VTTs. The results in Table 3 combine these two effects. The mean economic costs are indeed higher using these individual VTTs, but

exclude households with piped connections or private wells at home.

the medians are lower[6].

5.2 Financial water costs

Prices charged at public water points or at "neighbor's" sources were generally Ksh 2 or Ksh 2.5 per 20L jerrican[7]. Households report modest differences in prices at different sources. Although some of this dispersion could be recall or inaccuracy problems, it may reflect real differences in prices charged to different households. We therefore use reported prices rather than taking averages or medians. We found no cases of households paying for water away from the home periodically rather than volumetrically[8]. We multiply these volumetric prices by the number of jerricans the household reported collecting from that source in the past 7 days[9] and multiply by 4.29 (30/7) to convert to monthly costs. Given the total volumes collected (Table 1), the average monthly financial expenditure by households without primary sources at home is Ksh 387 (USD 4.5). This figure excludes household purchases from distributing water vendors.

We assume zero marginal financial cost for households with private wells. Only five of 87 households with wells used an electric pump; the remainder used a bucket and rope (n=79)

[6]Appendix Figure A2 may help explain this discrepancy. The figure plots each household's monthly travel cost using both approaches. The estimate using individual VTTs is plotted on the y-axis, and the one using the benchmark VTT is plotted on the x-axis. The red dashed line shows where they give the same result; points above and to the left of the line indicate households where using the individual-level VTT gives a larger monthly travel cost. The top panel plots all data while the bottom focuses on the households with monthly travel costs less than 2000 Ksh. The plotted points form clear lines; these correspond to each of the three latent classes (the points on the dashed line are mainly households where we have no individual VTT and assume it is equal to 17.5 Ksh). The "line" in the upper left corresponds to class 1 where the individual VTT of 49Ksh is much higher than 17.5 Ksh. The line closest to the x-axis is class 2 with an individual VTT of only 1 Ksh, and the line in the middle is class 3. The top panel shows ten or so points where the monthly travel costs implied by the individual VTTs are two to three times higher than those using the benchmark, driving up the average. The median will be less affected by the outliers, and driven more by how the VTT choice affects the bottom 50% of households, which is better displayed in the bottom panel of the figure. There are more points that lie below the dashed line than above it, driving down the median when using the individual VTT compared to the mean.

[7]Converted to USD at Ksh 86 per USD, Ksh 2 per jerrican is equivalent to USD 1.16 per cubic meter, or USD 0.31 per thousand gallons

[8]In 15 cases, households said they could use a source away from the home and reported paying per period rather than volumetrically, but none of them actually reported collecting water from that source in the past 7 days nor in an average dry or rainy season week, so it is likely they were mistaken about a source they do not use.

[9]Given the concern above with potentially unrealistic numbers of trips, volumes are truncated in a similar manner here to avoid inflating the average financial costs paid for water collected. We truncate households who said they collected more than 200 liters per household member per day. This affects 10 records, three of which are households buying from water vendors, 6 collecting from their private well (zero financial costs) and 1 household using its piped system (again, zero volumetric costs).

or a handpump (n=3). The cost of the electricity needed to use a pump is low, and without information on the model or efficiency of the electric pumps, our estimate of the electricity costs would be inaccurate. We thus choose to omit pumping costs for these five households.

Similarly, few households with piped connections face positive marginal water prices. Seventy-six households reported that they have a working piped connection, including 16 for whom it is *not* their primary source. Of these 76 households, 18% reported paying nothing for their water and 51% paid a one-time amount to obtain their connection, most commonly Ksh 300 (USD 3.5). Twenty percent (n=15) pay a non-volumetric monthly fee, typically Ksh 500. Only five respondents said they have a working meter, receive a bill and pay a volumetric charge. Only one respondent reported this charge, however (Ksh 5 per cubic meter). Rather than calculate volumetric bills, we used the previous month's reported bill to estimate the household's monthly financial cost[10].

Eighty percent of respondents said it was possible to buy vended water where they live. Eighty respondents (21%) said they bought vended water during the past seven days, and 189 (49%) said they purchase vended water in an average dry season week. As shown in Table 1 above, however, only 15 households said that they relied primarily on vended water, so most households use it as a supplemental source. Most of the 189 households who buy vended water buy less than 2 jerricans per person per day. Although a small number of respondents reported volumetric prices of Ksh 15, Ksh 20, or Ksh 30 per jerrican, 81% reported that the price per jerrican in the dry season was Ksh 10.[11] We again use prices reported by the household since a vendor may charge different households different prices depending on their distance from the vendor's water source. Average monthly expenditure on vended water in the entire sample is Ksh 254 (USD 2.95); the median is zero. Among those who purchased any vended water in the past 7 days, mean and median monthly expenditures are Ksh 1,227 (USD 14.3) and Ksh 879 (USD 10.2).[12] The study team was told that water vending was non-existent in the rainy

[10]The survey asked about arrears; in one case, the previous month's bill covered nine months of service. In this case, the bill is divided by nine, which may overestimate the volumetric costs if late payment penalties were added.

[11]Ksh 10 per jerrican is equivalent to USD 1.55 per thousand gallons, or USD 5.80 per cubic meter.

[12]These figures and the totals presented further in the paper censor four observations where the household reported buying more than 10 jerricans per day but did not list vending as their primary source. These amounts,

season, so the questionnaire did not ask about rainy season water vending expenditures.

Sixty-seven respondents reported buying bottled water in the past seven days; 75% of these bought three liters or fewer for the whole household over that time. A survey of local market prices for bottled water found a cost of 50 Ksh for one liter, the most common type of volume purchased by households.

5.3 Capital costs

The most common types of capital investments for coping with unreliable water supply are a) purchase of storage containers and b) construction of private wells. Investments in bicycles, carts and motorbikes, etc. might be seen as representing coping costs to ease the burden of water collection. These assets have multiple purposes, however, so lacking information on how important water collection was to the decision to purchase them, we omit them as coping costs. For all assets, we convert capital costs to monthly costs by amortizing using a 10% real discount rate and an assumed useful life of 20 years.

We asked households with piped connections and households who said they use rainwater about investments in storage equipment. Eighty-nine percent of households with working piped connections had invested in storage. The majority purchased tanks of 100 and 200L capacity, though a quarter purchased tanks that stored 1 cubic meter or more; average total storage capacity is 1,153 liters (median 340 liters). We asked about the cost (in today's dollars) to purchase their largest tank, which averages Ksh 5995 (USD 70). Amortized over 20 years at 10%, this is equivalent to a monthly cost of Ksh 59 (USD 0.69)[13] Among the 176 (70%) households who said they had built a rainwater collection system, the average system has a total storage capacity of 2.4 cubic meters and would cost Ksh 9,000 if purchased today. These capital costs are equivalent to an average monthly payment of Ksh 87. These averages are affected by a few

some of which imply monthly expenditures of USD 228, are unlikely. We censor the number of jerricans at ten for these four households.

[13]We first convert capital costs to annual values and then divide by 12, rather than compounding over 20*12 periods. The calculation just reported is as follows: the annuity factor for r=10% and t=20 years is (1 - 1/(1.1)^{20})/0.1 = 8.51. A capital cost of Ksh 5995 would be equivalent to 5995/8.51 = Ksh 704 per year, or 704/12 = Ksh 59 per month.

very large systems; median monthly capital costs are Ksh 7.

Table 4: Monthly coping costs (Ksh) during the **dry season**, excluding time costs, by primary source, mean (top) and median(bottom)

	Vending Costs	Bottled Water	Water Purchase	Capital Invest.	Diarrhea Trt.	Water Trt.
Private piped(60)	15	111	159	92	50	891
Private well(78)	38	185	16	363	48	817
Vended(15)	1167	271	47	67	50	214
Public well(122)	367	33	405	21	128	595
Public borehole(72)	368	185	454	18	32	613
Public piped conn(32)	48	33	397	3	0	461
Surface, other public(6)	657	36	246	6	0	462

Median

	Vending Costs	Bottled	Water Purchase	Capital Invest.	Diarrhea Trt.	Water Trt.
Private piped(60)	0	0	0	15	0	278
Private well(78)	0	0	0	343	0	246
Vended(15)	1114	0	0	4	0	123
Public well(122)	0	0	279	0	0	123
Public borehole(72)	0	0	300	0	0	2
Public piped conn(32)	0	0	257	0	0	74
Surface, other public(6)	300	0	240	0	0	467

Notes: N=325. 1 US$= 86 Ksh at the time of survey. Water treatment costs for vended and piped water are based on a predicted probability of treating water.

Households with private wells estimated that digging a similar well today would cost an average of Ksh 35,000 (USD 407), or Ksh 970 per foot of depth. Because eight respondents could not estimate the cost, and because of several outliers, we use predicted well digging costs from a simple linear model of reported depth on reported cost. Amortizing over 20 years at 10%, these imply average monthly capital costs for the 87 households with wells of Ksh 363 (USD 4.22).

5.4 Diarrhea treatment costs

We asked households how much money they had spent on diarrhea treatment costs in the past seven days, and multiplied these estimates by 4.29 to get monthly expenditures. At the level of an individual household, this is likely an overestimate of the diarrhea treatment costs given that diarrhea is acute and may not require a month of care and treatment. At the population level, however, the same logic implies that households who did not have a diarrhea case in the seven days prior to the survey may have had a case at some other time in the previous month. Nevertheless, diarrhea treatment costs are low. Twenty-nine respondents (7.5%) had spent money on diarrhea treatment and the average weekly expenditure among these households was 408 Ksh (USD 4.74). Among all households, including those who reported no diarrhea treatment costs, the average weekly cost is 31 Ksh (USD 0.36). The causes of diarrhea are complex, however, and these costs should not be attributed entirely to poor water and sanitation. For lack of a better estimate, we attribute half of these costs to water and sanitation.

5.5 Water treatment costs

The survey asked whether and how water was treated before drinking. We rely on a number of assumptions to calculate total treatment costs. First, we assume that only the volume of water reported used for drinking in the past 7 days was treated. For sources away from the household, we asked if they treat "always" or "sometimes", and assume that the latter implies half the volume of water is treated. This frequency question was not asked for water from rainwater or private wells; we assume that if a household reported treating water from these sources they "always" did. The treatment options "stand and settle" and "filter" are assumed to have no financial cost, since costly household point-of-use filtering devices are not common in the region. Furthermore, only 9 people reported filtering water.

For the 43% of households that reporting boiling water, we base the cost on the predominant fuel used for cooking, which is firewood for 80% of households. We assume that 1 kg of wood is needed to boil 1 liter of water (World Health Organization, 2015). Based on a survey of local prices, the cost of firewood is 800 Ksh for 70 kg of firewood, or 11 Ksh per liter (for purposes of comparison, recall that the local price of bottled water is 50 Ksh per liter). Fourteen percent of households use biomass (e.g. maize cobs, leaves) as their main source of energy for cooking. We maintain the assumption that 1 kg of biomass is needed to boil 1 liter, but assume the price is half that of firewood, since biomass is more often collected locally and because market prices of biomass were unavailable. Fifteen households (4%) reported using charcoal and 1% use electricity; we use the same treatment cost per liter as we use for firewood for lack of better information. For the 14% of households who reported chlorinating, we use a local price for Waterguard-brand chlorine of 25Ksh/1000L, or 0.025 Ksh per treated liter. A number of households with private wells reported adding chlorine directly to the well. We do not have information on what, how often, or how much they add to the well, although given chlorine's price this omission is unlikely to introduce a serious downward bias in our treatment cost estimates. Mean monthly average treatment costs among the entire sample, including those who do not treat water, are Ksh 658 (USD 7.7, median = Ksh 168).[14] Among the 69% who reported treating their water at all, average and median costs are Ksh 957 and Ksh 491. This figure is driven overwhelmingly by the 43% of respondents who said they boil water. Drinking water in household containers was not tested for microbiological quality, nor did we verify that the large fraction who said they boil water actually do. It seems likely that, on the whole, these estimates are biased upwards because of respondents wanting to give the "correct" answer regarding treatment costs (i.e. "social desirability" bias).

6 Total coping costs

Using our "truncated" approach to estimating travel times and valuing them with individual-level values of travel times, the median household in our sample incurs total economic costs of dealing with unreliable water supply in the dry season of 1737 Ksh (US$20.2). The total average cost is 3372 Ksh (US$39) per month. Expressed as a percentage of reported monthly cash income, the median household incurs costs equal to 4.5% of its reported cash income, and the average percentage is 15%. Eighty-four percent of respondents said coping costs were at least 1% of income, 47% said they were more than 5% of income, and 29% said they were more than 10% (Figure 5). Total monthly coping costs tend to be higher for respondents in lower quintiles

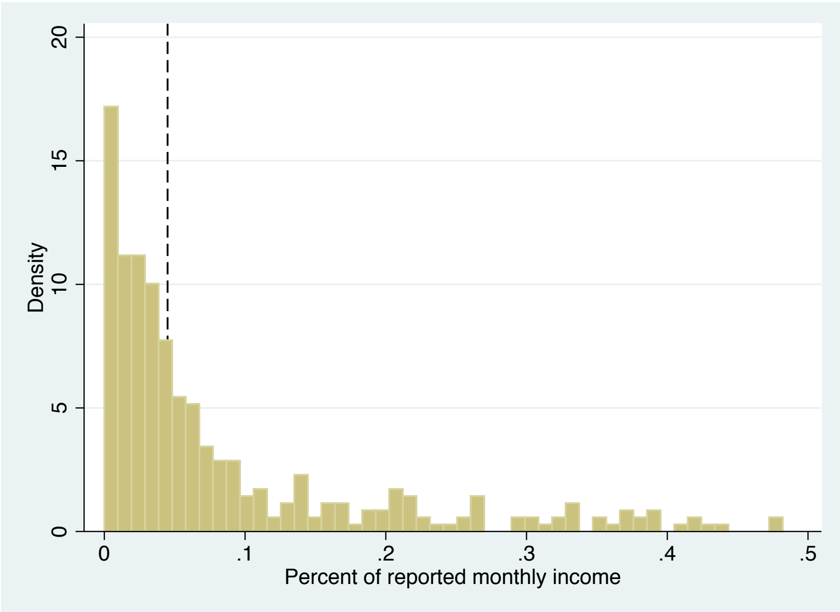
[14]Respondents were not asked about treatment of drinking water from vendors or from piped connections. We predict treatment costs based on a regression model of total monthly treatment costs for all other households (Appendix Table A2). As expected, treatment costs increase with perceived risk, education, and assets. They are negatively but weakly associated with diarrhea treatment costs. Predicting out-of-sample based on these covariates produces the monthly treatment costs in Table 4 for households whose primary source is vended or piped water.

of reported income, in contrast to the theoretical prediction outlined above, though the results are not monotonic (Figure 6, left). There is no clear pattern of results by wealth quintile (Figure 6, right). Total median coping costs fall to Ksh 814 (USD 9.47) during the rainy season; the mean coping costs in the rainy season are Ksh 1,879.

Figure 7 shows the breakdowns by category of coping cost; for respondents who rely primarily on sources away from the home, the majority of costs are travel time costs[15]. The figure again illustrates the impact on summary statistics of households who report large travel times - coping costs calculated by subgroup medians are much lower than means, though still a significant fraction of reported household income. One surprising result is that households who rely primarily on vended water have *lower* total coping costs (median = Ksh 1940) than households walking to collect water (median = Ksh 2304). If costs are really higher for households walking to get water, why would they not just hire vendors themselves? This is driven largely by differences in household size and thus total water collection. The fifteen “vended” households have an average household size of 4.1 members versus 5.7 members in households carrying water home, and total water collected in the past seven days is 710 liters among vended households and 1136 liters among households who walk to water. Households who rely on water vendors are indeed wealthier: their total monthly income is Ksh 50,170 compared to Ksh 38,900 for households carrying water. A further possibility is that, although households carrying water may incur large economic burdens, it is not easy for them to convert these economic costs into financial resources, i.e. convert saved time into money to pay vendors.

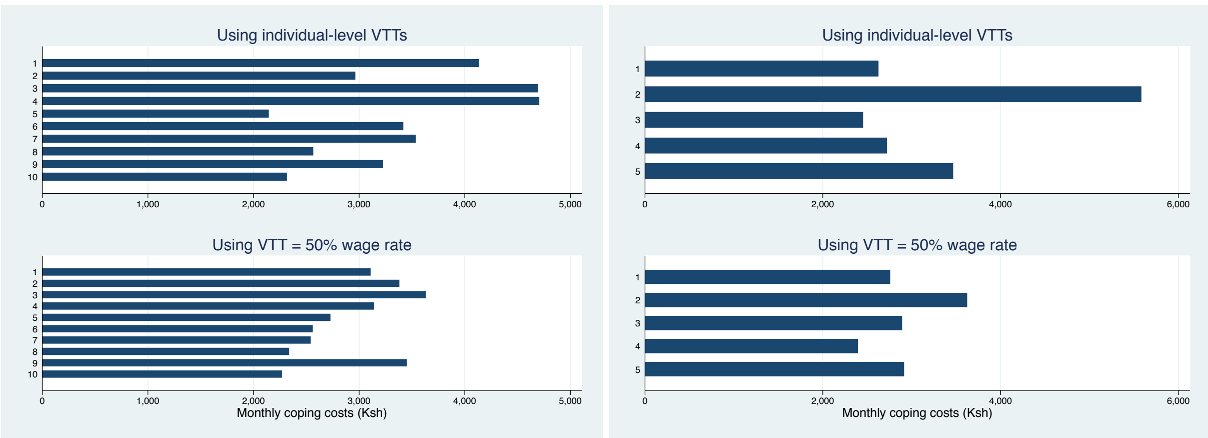
[15]These graphs are constructed by taking the median (or mean) value in each category within that primary source subgroup and then stacking them horizontally side by side.

Figure 5: Coping costs as a percent of reported household monthly income



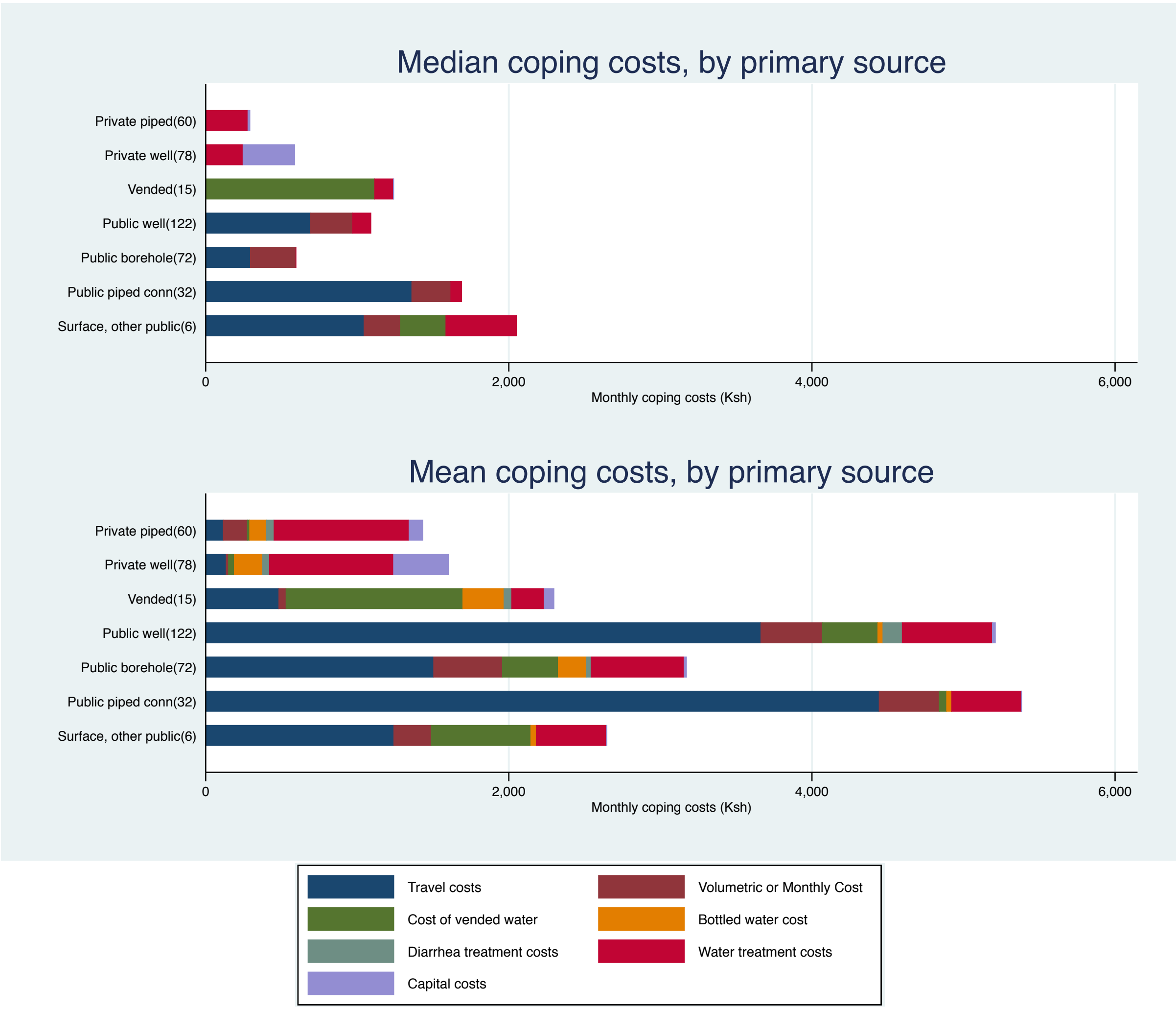
Notes: Uses truncated collection times and individual-level value of travel time estimates. Figure does not display 23 observations with coping costs over 50% of reported cash income.

Figure 6: Average total monthly coping costs, by **income decile** (left) and **wealth quintile** (right)



Notes: N=387. Uses truncated collection times, as described in the text. The first decile or quantile is the poorest group.

Figure 7: Mean and median total monthly coping costs (in Ksh), by primary source



Notes: N=387. Uses "truncated" travel time approach and individual-level value of travel time estimates. Two respondents with large rainwater collection systems not shown.

Finally, we regress total monthly coping costs on characteristics of the household (Table 5). Households with piped connections or private wells at home have lower monthly coping costs than those traveling to get water or buying vended water. There is no statistically significant relationship between decile of income or whether the respondent has a primary education, although the coefficient on income is negative (consistent with Figure 6). Larger households with larger needs for water collection (but also a potentially greater supply of water collectors) have higher monthly coping costs.

Table 5: OLS regression of total monthly coping costs[a]

Primary source is piped conn.	-2401.8***	(-3.20)
Primary source is private well	-2919.9***	(-4.38)
Primary source is vended water	-1181.1	(-0.86)
Income decile	-111.0	(-1.21)
Resp. has primary education	-334.6	(-0.59)
Age of respondent	14.9	(0.74)
Household size	688.2***	(5.65)
Constant	782.0	(0.61)
Observations	381	
R-square	0.16	

Notes: N=387. t-stats in parentheses; $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.001$. [a]Use "truncated" travel time approach and individual-level value of travel time estimates.

7 Conclusions

Using a carefully-constructed picture of households' daily water behaviors in this area of rural Kenya, we find that the total economic costs of dealing with unreliable or distant water supply are 3400 Ksh per month (US$38) (median = Ksh 1700 / US$20). The typical household incurs financial and economic costs equal to 4.3% of their reported income, and nearly a quarter incur costs greater than 10% of income. As a point of comparison, this total monthly coping cost is larger than the water bill a household would face in 25 of 30 major US cities[16]. It is also

[16]According to the 2014 survey of water rates in 30 major cities by Circle of Blue: http://www.circleofblue.org/waternews/wp-content/uploads/2014/05/WaterPricing2014TableInteractive.pdf.

much higher than the water bill of a typical household in Nairobi. According to the 2014 tariff schedule, a customer using 10m^3 (roughly double the consumption among households in our site) would pay \$2.76 for water service plus a wastewater charge of \$1.70. The tariff structure in Nairobi is much too low to cover operations and maintenance, let alone capital depreciation, but the comparison highlights an easily overlooked fact: the economic burdens of poor water supply often falls more heavily on unconnected rural customers than on households with piped connections.

This result is driven primarily by time costs, but treatment costs and monetary expenditures to public taps, neighbors, and vendors were also important. Forty-three percent of households reported boiling drinking water, implying non-trivial water treatment costs, and about a third of households in our study site had sources at home - either piped connections or private wells - where capital investments in storage capacity were also important. A small fraction of households incurred large costs relying on vended water. But for most households, the majority of costs are driven by the daily burden of bringing water home. The median household reported spending nearly two and a half hours per day collecting water. Although we acknowledge the potential pitfalls of relying on households' own reports of the time they spend collecting, a range of approaches for handling collection times and the value of time imply that, for households without water at home, the median time costs are Ksh 700 - 800 per month, approximately 2.5% of income.

As mentioned in Section 3, the water utility in the region (IWASCO) is in the process of rehabilitating the failed piped distribution system. Is this a good investment? Could households afford to connect to the new system? Lacking detailed cost data from IWASCO, we use a rough approximation of US\$0.80 per cubic meter in (Whittington et al., 2010, Table 2.1,pg.485); these do not include any networked wastewater collection system. When households are (re)connected to the distribution system, we would expect total volumes of water used to increase, though this will be a function of the price that IWASCO ultimately charges and other taste and structural characteristics of the household. As a rough estimate, assume that

The bill calculation is for a "low"-use household using 190 LCD, far higher than the water used in households in our sample (see Table 1).

households will use 110 liters per capita per day, or 20 cubic meters per month for a household of six members. We would estimate the total economic costs of providing piped water service to be US$16, or roughly 1400 Ksh per month. We estimate that 56 percent of households in our survey have total monthly coping costs in the dry season that exceed US$16. At a lower bound unit cost of US$0.35[17] for 20 cubic meters, we estimate that 78% of households have dry-season coping costs exceeding the monthly cost of US$7[18].

Another way of examining the question would be through the lens of the "5% rule": households cannot afford to spend more than five percent of their monthly income on water and sewer services. From this perspective, even if all of the economic coping costs could be converted to financial resources to pay for full-cost water tariffs, some in the sector would object to a piped system where many people were paying more than 5% of their income. What fraction of houses have coping costs over 1400 Ksh per month *and* could "afford" full-cost tariffs of 1400 Ksh per month? Only one-third of households fit this criteria, and 68% fit a similar criteria for the low-cost option (coping costs greater than USD 7 and USD 7 is less than 5% of income). This affordability criteria paints a more pessimistic picture for the financial sustainability of IWASCO's rehabilitation plan, but is in our view misguided. Our estimates suggest that many households are *already* spending more than 5% of monthly cash income on managing their water supply situation and coping with its effects, and would be likely to connect to a rehabilitated system even with full cost pricing. Recall also that coping costs are a lower bound on total economic benefits from improved water service, and in particular miss any benefits of reduced mortality risk. An even higher fraction than we predict may wish to connect, though coping costs are also lower during the approximately five months of the rainy season. Finally, the majority of coping costs are travel costs for most households, and the value of these losses will increase if the Kenyan economy continues to grow and the opportunity cost of water collectors' time increases. In terms of system sustainability, it may be more important to get the pricing policy for the rehabilitated system correct. Recall that we found none of the

[17] This excludes any opportunity cost of raw water, has minimal storage and simple chlorination treatment, and uses cheaper PVC pipe for the distribution network. See (Whittington et al., 2010, p.488).

[18] These estimates use individual-level VTTs. Using a value of time benchmark approach (50% of the unskilled wage rate), the corresponding percentages are 61% and 79%.

households with working piped connections were now paying volumetric water bills, and only twenty percent were paying a monthly non-volumetric fee of Ksh 500, far below our estimate of full cost pricing.

References

Abrahams, Nii Adote, Bryan J. Hubbell, and Jeffrey L. Jordan, "Joint Production and Averting Expenditure Measures of Willingness to Pay: Do Water Expenditures Really Measure Avoidance Costs?," *American Journal of Agricultural Economics*, 2000, *82* (May), 427–437.

Calvo, Christina Malmberg, "Case Study on the Role of Women in Rural Transport : Access of Women to Domestic Facilities," 1994.

Cook, Joseph, Peter Kimuyu, Annalise Blum, and Josephine Gatua, "A simple stated preference tool for estimating the value of travel time in rural Africa," 2015.

Filmer, Deon and Kinnon Scott, "Assessing Asset Indices," *Demography*, 2012, *49* (April), 359–392.

_ **and Lant Prichett**, "Estimating Wealth Effects Without Expenditure Data or Tears: An application to educational enrollments in states of India," *Demography*, 2001, *38* (1), 115–132.

Ho, Jeff C, Kory C Russel, and Jennifer Davis, "The challenge of global water access monitoring: evaluating straight-line distance versus self-reported travel time among rural households in Mozambique.," *Journal of Water and Health*, 2014, *12* (1), 173–83.

Jalan, Jyotsna and E. Somanathan, "The importance of being informed: Experimental evidence on demand for environmental quality," *Journal of Development Economics*, 2008, *87*, 14–28.

Jessoe, Katrina, "Improved source, improved quality? Demand for drinking water quality in rural India," *Journal of Environmental Economics and Management*, November 2013, *66* (3), 460–475.

Katuwal, Hari and Alok K. Bohara, "Coping with poor water supplies: Empirical evidence from Kathmandu, Nepal," *Journal of Water and Health*, 2011, *9*, 143–158.

Kremer, Michael, Jessica Leino, Edward Miguel, and Alix Peterson Zwane, "Spring cleaning: Rural water impacts, valuation, and property rights institutions," *Quarterly Journal of Economics*, 2011, *126* (1), 145–205.

Larson, Bruce and Ekaterina Gnedenko, "Avoiding health risks from drinking water in Moscow: An empirical analysis," *Environment and Development Economics*, 1999, *4* (1999), 565–581.

McConnell, Kenneth E. and Marcia a. Rosado, "Valuing discrete improvements in drinking water quality through revealed preferences," *Water Resources Research*, 2000, *36* (6), 1575–1582.

Nganyanyuka, Kapongola, Javier Martinez, Anna Wesselink, Juma H. Lungo, and Yola Georgiadou, "Accessing water services in Dar es Salaam: Are we counting what counts?," *Habitat International*, 2014, *44* (2014), 358–366.

Pattanayak, Subhrendu, Christine Poulos, Jui-Chen Yang, and Sumeet R Patil, "How valuable are environmental health interventions? Evidence from a quasi-experimental evaluation of community water projects," *Bulletin of the World Health Organization*, 2010, *88*, 535–542.

Pattanayak, Subhrendu K and Alexander Pfaff, "Behavior, Environment, and Health in Developing Countries: Evaluation and Valuation," *Annual Review of Resource Economics*, 2009, *1* (1), 183–217.

_ , Jui chen Yang, Dale Whittington, and K C Bal Kumar, "Coping with unreliable public water supplies : Averting expenditures by households in Kathmandu , Nepal," *Water Resources Research*, 2005, *41* (1), 1–11.

Rajaratnam, Julie Knoll, Jake R Marcus, Abraham D Flaxman, Haidong Wang, Alison Levin-Rector, Laura Dwyer, Megan Costa, Alan D Lopez, and Christopher J L Murray, "Neonatal, postneonatal, childhood, and under-5 mortality for 187 countries, 1970-2010: a systematic analysis of progress towards Millennium Development Goal 4.," *Lancet*, June 2010, *375* (9730), 1988–2008.

Tanser, Frank, Brice Gijsbertsen, and Kobus Herbst, "Modelling and understanding primary health care accessibility and utilization in rural South Africa: An exploration using a geographical information system," *Social Science and Medicine*, 2006, *63* (3), 691–705.

Um, Mj, Sj Kwak, and Ty Kim, "Estimating Willingness to Pay for Improved Drinking Water Quality Using Averting Behavior Method with Perception Measure," *Environmental and Resource Economics*, 2002, *21*, 285–300.

Vásquez, William F., "Reliability perceptions and water storage expenditures: Evidence from Nicaragua," *Water Resources Research*, 2012, *48*, 1–8.

_ , Pallab Mozumder, Jesús Hernández-Arce, and Robert P. Berrens, "Willingness to pay for safe drinking water: Evidence from Parral, Mexico," *Journal of Environmental Management*, 2009, *90*, 3391–3400.

White, Gilbert F, David J Bradley, and Anne U White, *Drawers of water: Domestic water use in East Africa* 1972.

Whittington, Dale, W Michael Hanemann, Claudia Sadoff, and Marc Jeuland, "The challenge of improving water and sanitation services in less developed countries," *Foundations and Trends in Microeconomics*, 2010, *4* (6-7).

— **, Xinming Mu, and Robert Roche**, "Calculating the value of time spent collecting water: Some estimates for Ukunda, Kenya," *World Development*, 1990, *18* (2), 269–280.

World Health Organization, "Managing water in the home: accelerated health gains from improved water supply," 2015.

Zérah, M. H., "Household strategies for coping with unreliable water supplies: The case of Delhi," *Habitat International*, 2000, *24*, 295–307.

A1 Appendix Materials

A1.1 Supplementary tables and figures

Table A1: Household demographics

	Mean	Std Dev
Household size	5.48	2.2
Water collectors	1.52	1.5
Female respondent	0.79	0.4
Years of education of female (head of hh or spouse)	7.22	3.7
Has working elec. conn.	0.11	0.3
Total monthly income (Ksh)	49896.4	70175.2
Weekly food exp. per person (Ksh)	435.9	322.1

Notes: N=387. One US dollar = 86 Kenyan shilling in the summer of 2013. Statistics for the number of water collectors includes households with primary water sources at home where no one reported collecting water in the past 7 days. Households collecting water away from the home reported an average of 2.2 members who collected water in the past week.

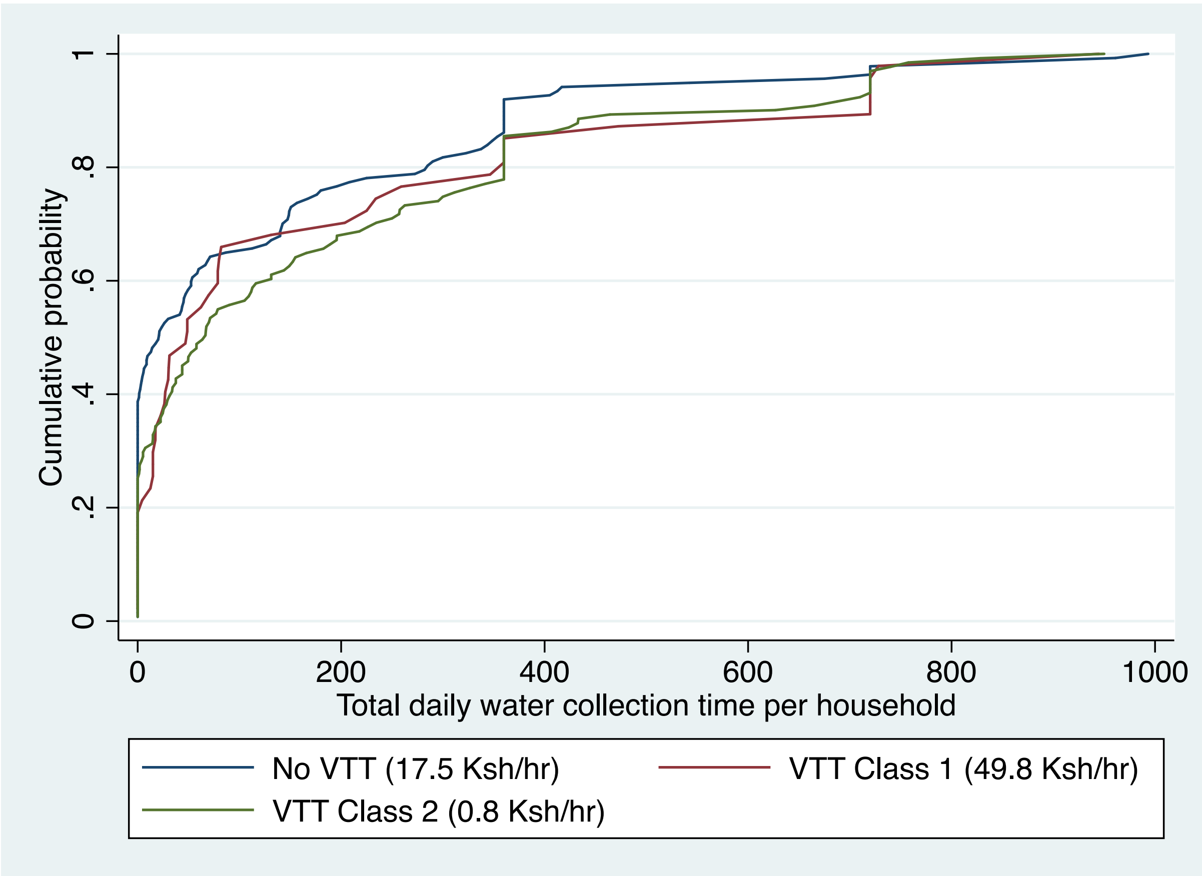
Table A2: Model explaining treatment costs

	Model A
=1 if 'some' or 'serious' risk of drinking water from source, =0 if no risk	2.834***
	(4.87)
Female respondent	-0.622
	(-0.89)
Respondent's years of education	0.246***
	(3.28)
HH member with diarrhea in past 7 days?	-1.424
	(-1.50)
Number of listed assets owned	0.372*
	(1.66)
Constant	1.204
	(1.16)
Observations	312
R-square	0.127

t statistics in parentheses

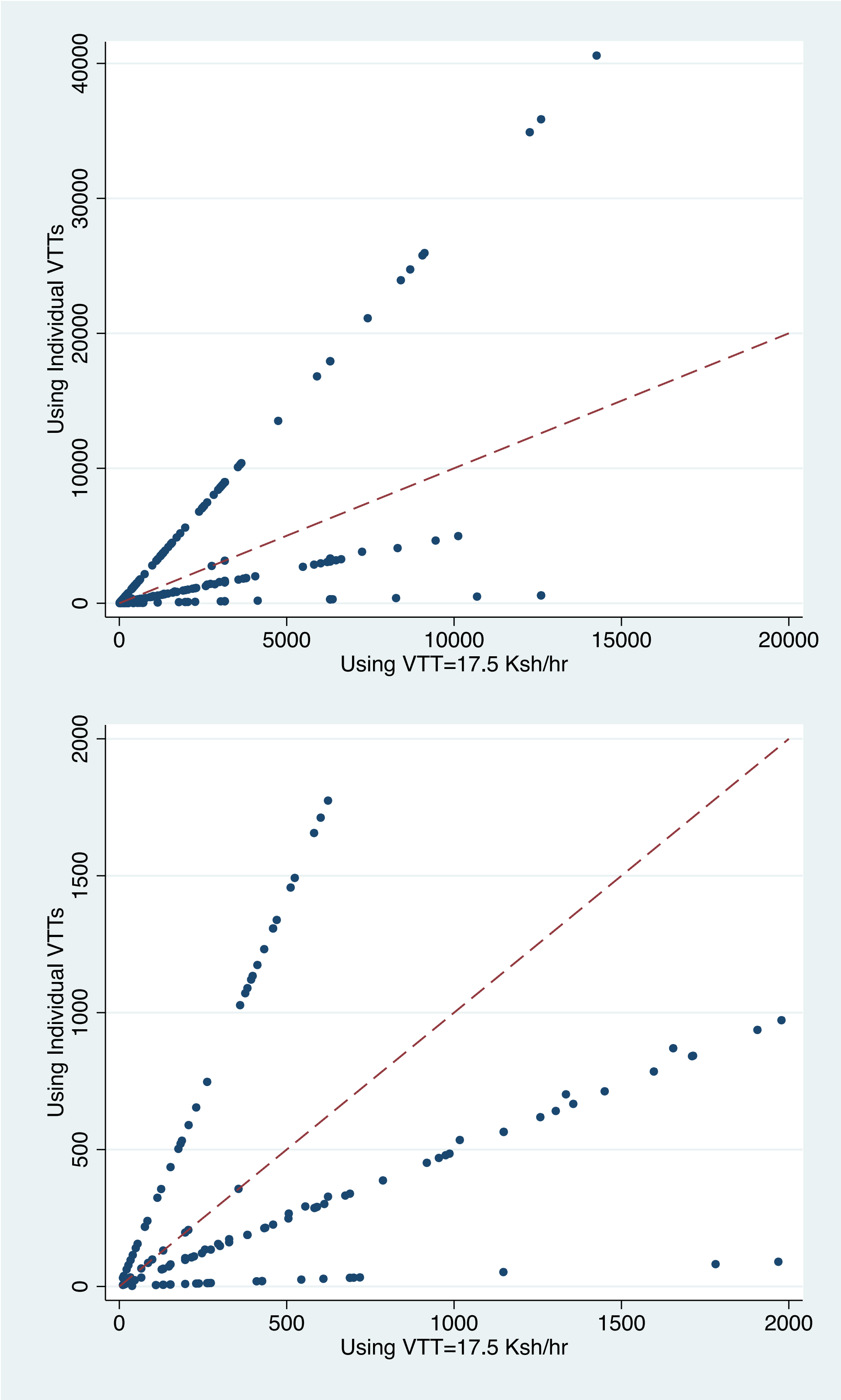
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure A1: Cumulative density function of total daily collection times (mins) per household, by VTT latent class



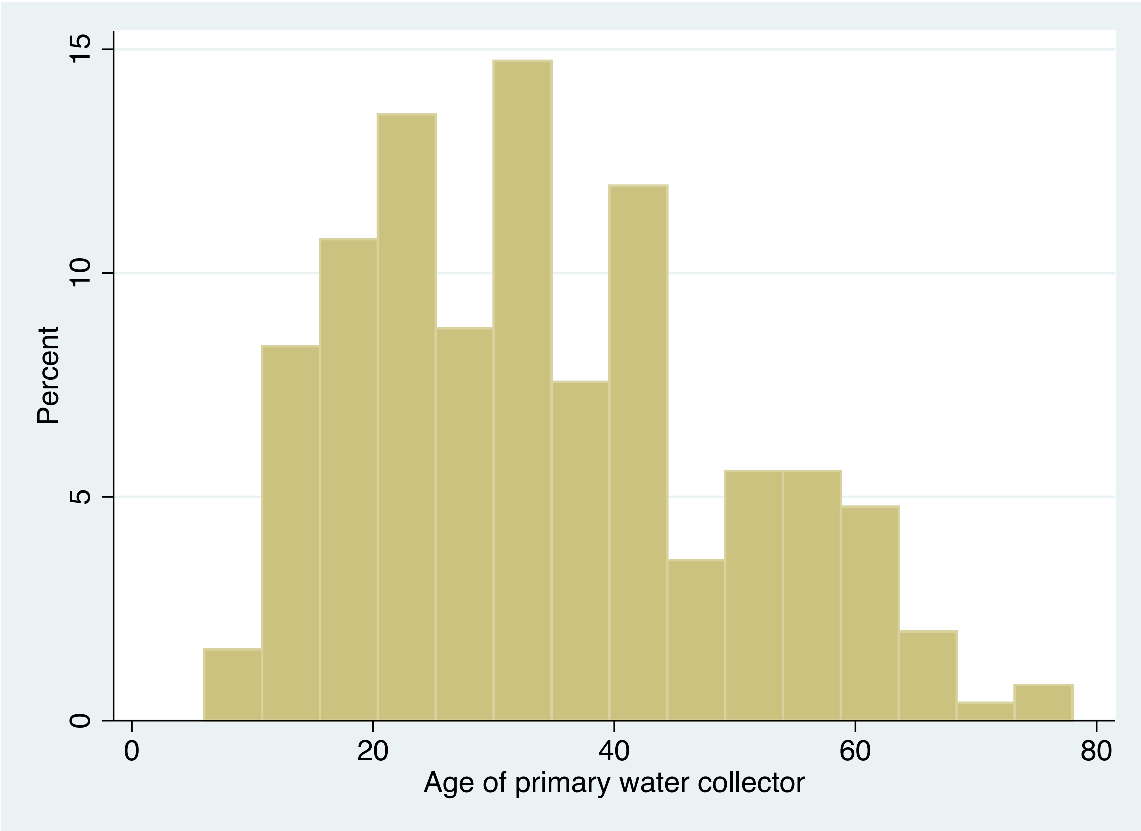
Notes: N=377. Uses truncated collection times, as described in the text. Figure omits 10 observations over 1000 minutes to display more of the useful range of the data. Of these ten, five were in class 1, three were in class 2, and two were in class 3.

Figure A2: Comparing total monthly travel costs by value-of-travel-time approach



Notes: N=387. Uses truncated collection times, as described in the text. Lower panel displays same data as upper panel for smaller range.

Figure A3: Age of primary water collector



Notes: N=255 households who reported at least one water collector.

A1.2 Sampling

Our fieldwork took place in Meru County, Kenya. The most recent census in 2009 estimates there are 320,616 households in an area amounting to 6,936 square kilometers, giving an estimated population density of 196 persons per square kilometer. The county has a 12 percent urban population compared to a national average of about 32 percent (Kenya Open Data Survey, 2014). The elevation is approximately 5,000 feet and average annual temperatures range from 62-69 degrees F. Considered one of the most fertile parts of Kenya, this agricultural area produces staple crops, such as wheat, potatoes, and maize, as well as cash crops, including tea, coffee, and bananas. Rice is sold for 85 Ksh (1 USD) per kilogram while the price of maize is 30 Ksh (0.35 USD) per kilogram (or 2.2 pounds). Average annual rainfall is fifty-four inches and there are a variety of surface and ground water sources.

We sampled households in four “sublocations” in the Tigania West “location” within Meru County: Kianjai, Mutionjuri, Machako and Nairiri. Although Meru County is in the top quarter of Kenya’s income distribution, Tigania West has many poor households which may represent the entire income distribution in Kenya somewhat better. According to the 2009 census, the populations of these sublocations were 1102, 1056, 337 and 398 households in Kianjai, Mutionjuri, Machako and Nairiri respectively.

The field team selected households by using access roads and paths as transect lines. Households were then randomly selected on either side of these paths for interview based on pre-determined skip patterns. Since our target sample was 400 households and the most recent census indicated 3,005 households in these four sublocations, we targeted approximately 13% of the population in each of these sublocations, or every fifth household. In 23 sampled households, the respondents in the household were unavailable so that call backs had to be scheduled. In 15 of these 23, an interview was later completed. The remainder were replaced after three unsuccessful attempts. Six households declined to be interviewed. This means that of 402 households contacted, 387 were interviewed giving a response rate of 96%. The final sample sizes by sublocation are given in Table 2.

Table A3: Interviews conducted and total number of households in each sublocation

Sublocation	Households interviewed	Total households in 2009 Census
Kianjai	141	1091
Mutionjuri	129	992
Machaku	44	341
Nairiri	74	581
TOTAL	388	3005

A1.3 Wealth Index

We construct a wealth index using principal component analysis (PCA) following Filmer and Prichett (2001) and Filmer and Scott (2012). Data on durable assets, electricity connection, sanitation, number of rooms and number of buildings, and main cooking fuel were included (see Table A4). Although water-related variables are often included in wealth indices constructed in this manner, we exclude them to avoid potential confounding with explanatory models in the main paper. All variables were converted to either dummy (0/1) or continuous variables. We had only two instances of missing observations. One respondent left blank any information about her ownership of livestock; we assume zero for this observation. A second respondent did not report the number of buildings in the compound; we assume it is one.

The first column of Table A4 reports the first principal component from the PCA analysis, which corresponds to the underlying latent variable of wealth. This first principal component had an estimated eigenvalue of 3.54, explaining 15.4% of variation in these 22 variables. In the case of binary variables, this score can be interpreted as the marginal change in the household's wealth score by moving from not owning the asset to owning (for example, owning a cell phone increases the household's wealth score by 0.221). Similarly, the three percent of housholds with no on-site sanitation option have a wealth score that is 0.163 lower than those with on-site sanitation.

The distribution of predicted wealth scores is relatively smooth and normally-distributed (Figure A4). Households are ordered on this predicted score and divided into five equal quintiles of 77 or 78 households each; the breakpoints in scores for each quintile are shown in the figure as vertical red lines. The remaining columns of Table A4 display the summary statistics for each of the component variables by the predicted quintile of wealth. For example, 71% of those in the lowest wealth quintile own a mobile phone, while all households in the fourth or fifth (highest) quintile own mobile phones. The results display face validity for the wealth index, with some small exceptions. Those in the lowest quintile own an average of 0.41 sheep, while those in the wealth quintile just above them own 0.19 sheep. There are similar patterns of non-monotonicity for owning a cart, owning a radio, owning a vehicle, and no sanitation at home.

Figure A4: Distribution of factors scores for wealth index

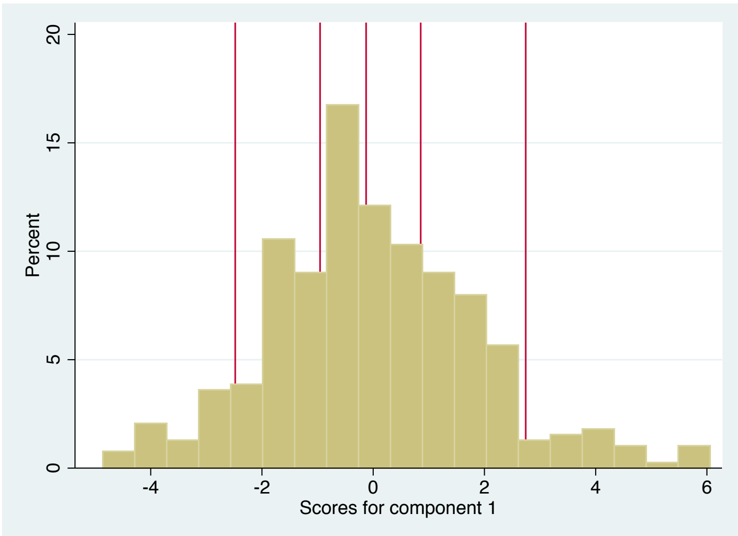


Table A4: Factor scores and descriptive statistics of components of the wealth index, by predicted wealth quintile

	Factor Score	Lowest	Second	Middle	Fourth	Highest	All
Own cell phone	0.221	0.71	0.97	0.99	1.00	1.00	0.93
Own bicycle	0.257	0.36	0.69	0.83	0.91	0.99	0.76
Own cart	0.158	0.05	0.04	0.09	0.27	0.35	0.16
Own radio	0.217	0.50	0.79	0.92	0.88	1.00	0.82
Own TV	0.289	0.05	0.19	0.30	0.62	0.79	0.39
Own motorbike	0.190	0.01	0.05	0.06	0.14	0.36	0.13
Own vehicle	0.206	0.00	0.01	0.04	0.03	0.27	0.07
Num. cattle	0.307	0.73	1.60	1.82	2.54	3.87	2.11
Num. goats	0.108	0.96	1.29	2.40	1.85	2.65	1.83
Num. sheep	0.128	0.41	0.19	0.42	0.96	1.27	0.65
Num. chickens	0.231	1.85	3.91	5.61	7.97	10.00	5.86
Own home	0.077	0.95	0.95	0.97	0.99	1.00	0.97
Has working elec. conn.	0.210	0.01	0.03	0.05	0.10	0.36	0.11
Num bedrooms	0.261	1.42	1.65	1.96	2.51	2.78	2.06
Num buildings	0.301	3.12	3.94	4.51	4.76	6.26	4.51
Acres land owned	0.281	0.95	1.07	1.60	1.92	4.47	2.00
No sanitation at home	-0.163	0.15	0.01	0.00	0.00	0.00	0.03
Owns non-shared toilet	0.190	0.60	0.79	0.92	0.94	0.97	0.85
Ventilated pit latrine	0.259	0.04	0.08	0.13	0.35	0.57	0.23
Cook w/ elec	0.054	0.00	0.00	0.01	0.01	0.03	0.01
Cook w/ biomass	-0.160	0.35	0.15	0.13	0.06	0.01	0.14
Cook w/ wood	0.160	0.58	0.78	0.82	0.88	0.96	0.80
Cook w/ charcoal	-0.065	0.08	0.06	0.04	0.04	0.00	0.04

Notes: N=387. Each quintile has 77 or 78 households.

A1.4 GIS

We collected GPS locations for all households using GPS-enabled Android mobile phones[19] and an Android app called "GPS Essentials". We attempted to collect locations of all water source locations. Of the 30 water sources that households said they could or do use, however, nine are geographically indeterminate (i.e "private well", "neighbor's well" "neighbor's piped connection") and six are surface water sources. Of these surface sources, many are seasonal or small rivers that do not show up on a satellite image, and we are not aware of hydrological GIS data layers for this section of Kenya. Three can be identified on satellite images: the Loria River, Thewa "swamp" and Mbututia "swamp". We hand-digitized the length of the Loria River through the study site, and the extent (as a polygon) of the two swamps. Of the 19 water sources with specific locations, we were able to record GIS data on thirteen. The field supervisor rather than the enumerator was responsible for geotagging a "new" water point, and in four cases the supervisor was not able to locate the water point with the name given by the

[19]The phone was the Samsung Galaxy Pocket, which retailed for around USD90 in Nairobi in July 2013. The phone's GPS unit is in theory accurate to 5m, though accuracy is affected by weather conditions.

household. Only 10 respondents reported using these four sources.

We then calculated the Euclidean (straight-line) distance between each household and all 13 GIS-coded water points as well as the nearest point on the three surface locations. Our distance measures will be an underestimate of the true distance since water carriers may not simply cross fields and homesteads but must use roads and paths. Ho et al. (2014) finds that straight-line distance calculations underestimate modeled travel along the road network in rural Mozambique (hand-digitized from satellite photography) by an average of 23%, and that self-reported distances to water sources were poorly correlated with these modeled times (R^2=0.12). Ho et al. (2014) use an average walking speed of 3.75 km per hour (kmh), averaging estimates from Calvo (1994) of 5 kph and Tanser et al. (2006) of 2 kph for travel on paths to unpaved "tracks", 3 kph for travel on "tracks", and 4 kph for travel on roads. White et al. (1972) uses an estimate of 4 kph (2.5mph). Only White et al. (1972) discusses the differences in walking speeds with and without the 20kg of a 20L jerrican. They briefly discuss the issue (White et al., 1972, p.95) in the context of calculating the energy expenditure of a water carrier with and without a load. They cite two studies from the early 1950's, noting that in one study subjects were reported to have slowed down 'a trifle' when walking with water on their heads. White et al. (1972) estimate that water carriers use approximately 12% more energy when walking 20L than walking unburdened, though the slope and condition of the path are important.

A more sophisticated routing analysis is possible, but requires hand-digitizing the entire road network from satellite maps in our site, which we considered a low priority. The area is dense with small roads and paths and a straight-line distance is reasonably accurate (Ho et al., 2014). Our study site is flat, so we ignore slope.

We merge this matrix of distances (16 per household) back with the household survey data, and keep only those distances where the person actually reported being able to use the water source. Of the 2,036 instances that a household reported they could use any water source (including "at-home" and "neighbor's"sources), we have household-source distances for 337. This is primarily driven by the fact that households in our sample most commonly report sources that are geographically indeterminate.

For those 337 instances, Figure A5 plots reported time spent walking to that source and the calculated GIS distance (in meters) for two subsets of households. The first subset (blue dots) either reported using the source during the current dry season, or reported that the source was their primary source, or both of these things. The second subset (green dots) did not report that the source was its primary one nor that they used it in dry season. Figure A5 adds simple linear fits for the two subsamples, and one can see that, although the fit is different for these two subsamples of distance calculations, the difference is not dramatic: households who are more familiar with the source are no more accurate about distances that those who are not.

Figure A5: Reported time walking vs. GIS-calculated distance with linear fits

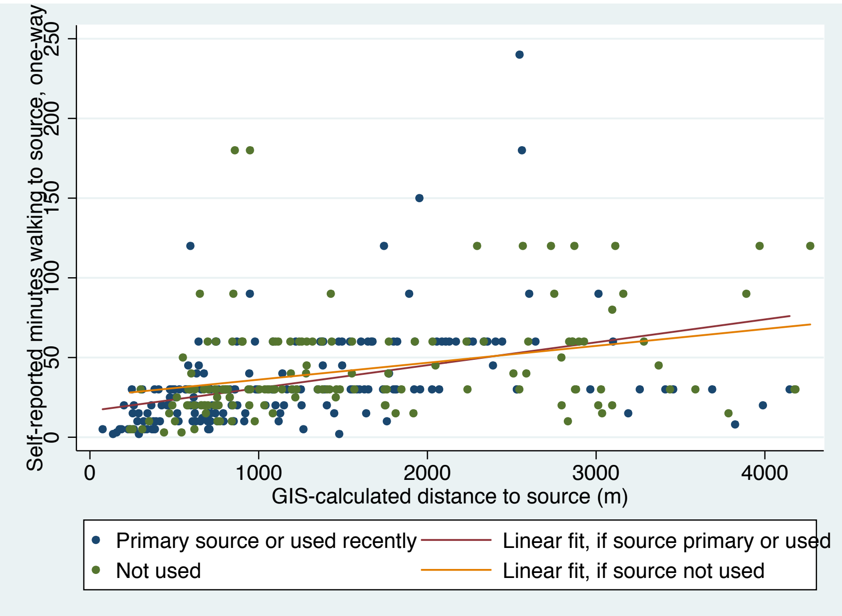


Table A5 shows two simple regression models (Models A and B) that fit the linear and quadratic forms to all data points in the scatter plot. We add a quadratic term to explore whether respondents have a more difficult time reporting travel times accurately as the source becomes farther away. As expected, the relationship between GIS-calculated distance and reported travel times is highly statistically significant in Models A and B, but the R^2 of 0.152 and 0.199 are somewhat poor. Ho et al. (2014) find a slightly-worse relationship (R^2=0.12) for a similar model of self-reported one-way travel times vs straight-line distance. The coefficient on the squared term in Model B is tiny but statistically signficant. The coefficient on distance (main term) in the linear model (A) means that a 1 meter increase in distance increases reported walking times by 0.0128 minutes (78.1 meters per minute), or 4.69 km per hour (2.91 mph). The coefficient on the quadratic model (B) implies a much slower walking speed of 1.6 kph.

Table A5: Regression models of GIS distance and reported walking times

	Model A	Model B	Model C	Model D
GIS-calculated distance to source (m)	0.0128***	0.0376***	0.0218***	0.0237***
	(7.72)	(6.41)	(9.90)	(2.80)
GIS Distance - Squared		-0.00000666***		-0.000000677
		(-4.40)		(-0.24)
Constant	20.02***	4.408	11.29***	10.27**
	(7.30)	(0.99)	(3.85)	(1.98)
Observations	334	334	298	298
R-square	0.152	0.199	0.249	0.249

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The distance calculations from households to **surface** water sources are imperfect. For the river, it implies that households go to the nearest segment of river, though some segments may be easier to access than others. In addition, the two swamps fill seasonally such that the pool of standing water may be at the edge of the swamp (as viewed from the satellite) during the wettest part of the season, but retreat into the"middle" of the swamp as the weather dries and groundwater levels fall. It is not clear from the satellite images where the "pool" is, so our GIS calculation is from the edge of the visible swamp. For Mbututia, the distance from the edge to the middle is approximately two hundred meters. Using the edge of the swamp will underestimate walk times, and using the centroid of the swamp will overestimate them. For these reasons, Models C and D show the same set of models but drop calculated distances between households and surface sources. This reduces the number of household-distance combinations to 300, but model fit improves (R^2=0.25). The squared term is no longer significant, so Model C is our preferred specification. These slope coefficients imply speeds of 2.75 kph/1.7 mph. This implied speed is slower than all of the speed estimates discussed above with the exception of the estimate of 2 kph for travel on paths to unpaved "tracks" (Tanser et al., 2006). They are, on the other hand, the only speed estimates that use responses where people are explicitly asked about time walking with a full container. They are also more reasonable than the speeds implied by a similar prediction exercise (0.5 kph) in Ho et al. (2014), who conclude that reported times are generally not reliable. We note, though, that straight-line distances will underestimate real travel distances on a network of roads and paths, so a similar model of network distance on self-reported times would imply somewhat faster walking speeds.

Figure A6 plots the predicted walking times (from Model C) against the reported walking times (on the y-axis). The red dashed line shows where the model predicted the same walking time as reported. Values above the dashed line indicate the predicted value is smaller than the one reported; values below the line indicate the model predicts *longer* walking times. As expected because of the regression context, the values are scattered around the dashed line. Two things are apparent in the graph. First, one can easily see the modal answers of 10 minutes, 20 minutes, 30 minutes, 60 minutes, etc. in the horizontal cluster of points at those levels. Second, there is a mass of points between 15 and 45 minutes where the model actually predicts longer walking times.

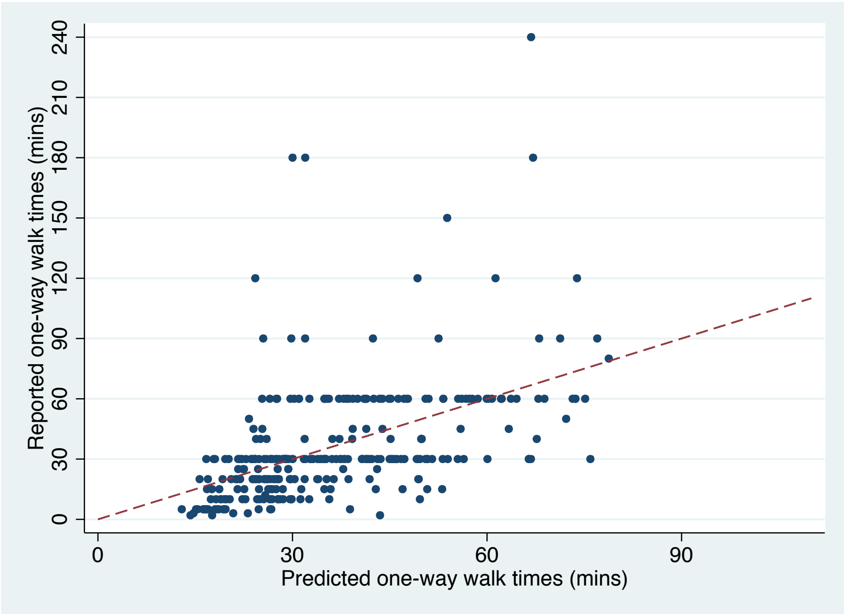
In the main paper, the "GIS distances" specifications use predicted walk times for those observations where we have GIS data, including private wells.[20] We do not predict times (out-of-sample in Model C) for surface water sources, but leave these as reported.

Lacking information in the published literature on the difference in walking speeds with and without 20L of water, we assume that unencumbered speeds are 75% faster (i.e we multiply our predicted full-container times by 1.75 to get roundtrip walk times). This is roughly consistent with White et al's 12% increase in energy expenditure when carrying water. It is also consistent with the ratio of our predicted, encumbered walking speed of 2.75 kph to the preferred speed in Ho et al. (2014) of 3.75 that would include unencumbered speeds (2.75/3.75 = 0.73).

Where we do not have any distance information, we try two approaches. First, we simply use self-reported walk-times, again multiplying by 1.75 to get roundtrip times. Second,

[20]We asked households how far away their private wells are from the house (in meters), though we did not ask them how long it takes to walk there. Model C was used to predict walking times for 74 households that reported distances to private wells. Because the model's statistically significant constant implies an 11-minute walk for a source even 1 meter away, we subtract the constant for these 74 households, leading to walk times all (plausibly) under 2 minutes.

Figure A6: Reported one-way walking times vs. predicted times (based on distance)



Notes: Dashed line shows where actual would equal predicted. Based on Model C in Table A5.

we use the median reported walk times for neighbor's sources. This will obviously increase reported walk times for half of the sample, but is intended to trim outliers. Because the density of houses varies by sublocation (and hence the distance to a neighbor), we calculate the medians by sublocation. These are 10 minutes for the three dense sublocations and 15 minutes for Nairiri, the sublocation furthest from the main road.