The short-run impacts of reducing water collection times on time use, well-being and education in rural Kenya

Joseph Cook [a], Jane Kabubo-Mariara [b] and Peter Kimuyu [c]

[a]School of Economic Sciences, Washington State University, Pullman WA USA
[b]Partnership for Economic Policy, Nairobi, Kenya
[c]Commission on Revenue Allocation, Kenya

Version July 2021

Abstract

Millions of households around the world devote significant time to bringing water to their home. This paper examines the impact that water collection has on the time allocation patterns and emotional wellbeing of water carriers and children in rural Kenya. We exogenously reduced water collection times to zero for a randomly-chosen subset of 195 households by having water vendors deliver water to their door each day over 4 weeks. Data on time use and affect (happiness, safety, energy, sociability, etc) come from short surveys the main water collector completed on a mobile phone at several randomly-chosen times each day over the 4 weeks of the treatment period as well as a 4-week baseline data collection period. Parents also self-reported school attendance, chores, and minutes spent studying for all school-aged children in the household, and we matched children to school-recorded attendance records. We find that of the approximately 95 minutes per day in water collection time that the vending treatment eliminates, water collectors reallocated approximately half to other household chores, 20% to working on the household's own farm, and 25% to leisure. We find no evidence of an increase in paid work. Water collectors report feeling happier, more energetic, more safe, and less likely to be in physical pain. Treatment increased school-recorded attendance by 3.6 percentage points, from a base of 92%. Data from the survey on school-aged children confirmed that receiving vended water reduced the probability that children collect water, but their time is reallocated to other chores, particularly cleaning and cooking. Nevertheless, children in treated households spent roughly 15% more minutes studying. Our results have implications for estimating the benefits of improving access to water supply in rural areas.

Introduction

Improving access to safe, convenient drinking water remains a key development challenge in many parts of the world, particularly rural areas of the Global South. The Joint Monitoring Programme of WHO and UNICEF estimate that one in four households globally do not have water at home ((UNICEF and WHO 2021, pg. 36). Sixty percent of households in least developed countries, and 69% of households in sub-Saharan Africa, lack at-home access (UNICEF and WHO 2021). Among policy-makers and researchers in the water sector, the focus has been predominantly on the health benefits of improving access to safe water, particularly in preventing diarrheal disease (see Wolf et al. (2018) for a recent review). The establishment of community or household water taps can, however, also significantly reduce the time spent collecting water, typically a task for women and girls (Sorenson, Morssink, and Campos 2011; Graham, Hirai, and Kim 2016). Water project beneficiaries themselves often perceive the increased convenience and time savings of water infrastructure improvements as equally or more important than health benefits (Hope 2006; Winter, Darmstadt, and Davis 2021; Cairncross and Valdmanis 2007; Bisung and Elliott 2019).

A number of studies have measured how much time is saved when water access becomes more convenient or reliable using either cross-sectional or quasi-experimental approaches (Cairncross and Cliff 1987; Ilahi and Grimard 2000; Ngongang 2008; Gibson and Mace 2006; Pattanayak et al. 2010; Arku 2010; Boone, Glick, and Sahn 2011; Chen, Chindarkar, and Zhao 2019; Meeks 2017; Winter, Darmstadt, and Davis 2021)[1]. All rely on water collection times as recalled by respondents, typically asked in isolation (e.g. "how many minutes did you spend yesterday collecting water") and sometimes as part of a recollection of an entire day.[2] Relatively few, however, have tried to measure whether and how the freed time is re-allocated to paid work, education, agricultural labor, or leisure (Devoto et al. 2012; Koolwal and Van de Walle 2013; Meeks 2017; Gross et al. 2018; Chen, Chindarkar, and Zhao 2019). Since children are often involved in resource collection (Cockburn and Dostie 2007; Porter et al. 2012), a related question

[1] There are also a number of conceptually similar studies on the time savings from switching from open defecation to toilet use (e.g. Dickinson et al. (2015)) and fuelwood collection time (see Jeuland et al. (2021) for a recent review). A number of studies have documented time savings when fuel is more available (for example by establishing community forestry plantations (Kohlin and Amacher 2005), though how the time is used is less clear. The connection between reduced wood collection times and schooling is not well established. The evidence is more convincing, however, that electrification increases women's off-farm employment and income, and improves school attendance for both boys and girls (Dinkelman 2011).

[2] See Chandrasekaran et al (2021) for a recent review of these studies, including a discussion of time use elicitation approaches.

is whether this prevents their development through lower school attendance or poorer performance[3]. Answers to these questions are an important but sparsely-populated piece of the economic evidence base for investment in the water supply sector, and will become more so as the child mortality rates continue to decline globally (Rajaratnam et al. 2010; Jeuland et al. 2013) and drinking water quality improves with point-of-use technologies.

The majority of existing studies have been primarily cross-sectional and non-experimental, jeopardizing identification of causal effects if households and/or water points are located non-randomly. The placement of water infrastructure may be influenced by a number of factors that might be correlated with time use patterns (e.g., socioeconomic status, regional or tribal differences, political allegiance), and households with a higher opportunity costs of time may consider water supply availability when making locational decisions. Although it is relatively straightforward to randomize water quality by randomizing access or prices for household water treatment devices like chlorine or filters (reviewed in Ahuja et al. (2010)) or protected wells (Kremer et al. 2011), we are aware of only one published study that has directly randomized quantity or distance to water sources. Devoto et al. (2012) offered a randomly-chosen subset of households in the city of Tangiers, Morocco a simplified procedure for connecting to the piped distribution system. Households who chose to connect and cease using the system of free public taps consumed significantly more water and saved a substantial amount of time. This time was used for increased leisure and social activities, but not productive activities, and no effects on school attendance were observed. Respondents in treatment households did, however, report less conflict with their neighbors and higher life satisfaction. Other studies have used quasi-experimental approaches to examine how water collection time is reallocated. Meeks (2017) and Gross et al. (2018) both exploit staggered timing of water infrastructure investments to examine time reallocation in rural Kyrgyzstan and Benin, respectively. Meeks (2017) finds that approximately half of saved water collection time is reallocated to farm labor; Gross et al. (2018) find no evidence that saved time was reallocated to income-generating activities.

[3] Most studies have been cross-sectional and relied on self-reported school attendance or enrollment (Akabayashi and Psacharopoulos 1999; Haile and Haile 2012; Ndiritu and Nyangena 2011; Nankhuni and Findeis 2004; Koolwal and Van de Walle 2013). Nauges (2017) built a panel dataset for rural Ghana to control for village fixed effects, and Gross et al. (2018) and Ashraf et al. (2021) use quasi-experimental approaches. Devoto et al. (2012) is the only experimental study, using self-reported school absenteeism. As we discuss in the conclusions, the evidence from these studies is mixed.

Furthermore, observing the behavioral response to a reduction in distance to a water point is complex. In a common situation where households have more than one public or surface water source to choose from, they make a water collection decision on two interrelated margins. The first margin is a discrete choice about which source or sources to collect from, balancing distance, water quality, financial price and other quality or availability factors (Mu, Whittington and Briscoe (1990); see also Nauges and Whittington (2010) for a review). They also make a continuous choice of how many trips to take and therefore how much water to collect (Wagner, Cook, and Kimuyu 2019; Gross and Elshiewy 2019). Installation of a more proximate source will lower the time price of collecting water, but this may be offset by a higher financial price for the improved source or impacted by the household's willingness to pay for higher water quality. Households may therefore react to installation of more proximate water sources by spending the same amount of time but collecting more water; collecting the same amount of water but reallocating time savings; or something in between. Indeed, this partially explains why Gross et al. (2018) found no effects of water collection time on labor or schooling outcomes. They find that use of improved water sources increased only modestly in treatment villages, with many households using the new source only intermittently. Daily time savings were also a modest 41 minutes (compared to a baseline of three hours per day) in part because households collected more containers per day and began using smaller containers because they were easier to carry.

This paper makes four contributions. First, we exogenously reduce water collection times to zero for a randomly-selected subset of 195 households in one subdistrict in rural Kenya. We used water vendors to deliver a generous quantity of water to treatment households for four weeks and succeeded in mainly eliminating water collection for those households. This allows us to generate causal estimates for the short-term effects of providing at-home water without the complication of endogenous changes in which source to collect from or how many collection trips to make, as discussed above. Although the existing quasi-experimental studies on time reallocation like Meeks (2017), Gross et al. (2018), and Ashraf et al. (2021) provide high-quality causal evidence, we add a second purely experimental approach, and the only experiment on time reallocation in the rural areas of the Global South where at-home water service is worst.

Second, we use a high-frequency approach called the Experience Sampling Method borrowed from psychology (Larson and Csikszentmihalyi 1983; Stone and Shiffman 1994) to measure time use without recall bias. We had the "main" water carrier in each household (93%

women) complete a short survey on their activities at four randomly-chosen times per day on a study-provided smartphone over the eight-week study period. This allows us to estimate panel difference-in-difference models with individual fixed effects to control for possible endogeneity between individual unobservable characteristics and outcomes.

Third, we also ask a number of questions in the high-frequency survey about well-being, including energy, safety, sociability, pain, and agency, dimensions which have received more attention in anthropology (Wutich and Ragsdale 2008; Bisung and Elliott 2019; Chindarkar, Jie, and Yogendra 2019; Wutich, Beresford, and Carvajal 2016). We ask about moment-by-moment (hedonic) happiness, a dimension of well-being that is less studied than overall life satisfaction but has found recent applications in economics (Bryson and Mackerron 2017; Allcott et al. 2020).

Finally, we contribute to the literature on how child schooling outcomes are affected by resource collection in our experimental setting. Like prior studies, we ask parents to report school attendance for children in their households, again using a high-frequency (daily) survey. We also, however, ask them to report the chores that children were asked to do and how much time they spent studying that day. We also linked children in study households to school-recorded attendance records to avoid relying solely on parents' reports, which are likely to suffer from a social desirability bias.

Our water vending treatment successfully reduced water collection times from a baseline of approximately 1.5 hours per day to nearly zero. Freed of this water collection burden, we find that water carriers reallocate approximately 20% of the time savings to working on the household's farm, an activity which is generally income-generating in our study site (rather than subsistence agriculture). The remaining time savings are reallocated to other household chores (roughly 50%) and leisure (roughly 25%). We find no evidence of an increase in paid work. Women did not report enjoying water collection; treatment participants reported feeling happier, more energetic and more safe, and less likely to be in physical pain. They also have increased agency, as measured by whether the person wishes they had been doing something else. We find no statistically significant impact on perceptions of sociability. School-reported attendance increased moderately (3.6 percentage points) for children from treatment households. The impact on attendance as self-reported by parents was larger. Data from the survey on school-aged children show that reducing water collection times to zero in the household reduces the probability that children collect water as expected, though their time is mainly reallocated to other chores, particularly cleaning and cooking. Nevertheless, we find children in treated households are more likely to have been reported as attending school that day, and spent 15% more minutes studying.

Study site and respondents

We conducted baseline interviews with a total of 248 households in four "sublocations" in the Tigania West political constituency, 19 kilometers from the larger town of Meru in north-central Kenya. The area is an important agricultural area, growing grams, peas, cassava, mangoes, and commercial livestock. We chose the study site purposefully in 2013 for a separate study of households' water source and collection decisions because of the diversity of existing water source options available (Cook, Kimuyu, and Whittington 2016; Wagner, Cook, and Kimuyu 2019). Sample households in 2015 were chosen randomly based on a transect approach (see Appendix D), and we re-contacted surveyed households to participate in this study. Because the experiment would reduce water collection times, we excluded households who had water at home, such as a private well or piped water. We also excluded households with no school-aged children at home.

Households reported that, on average, total water collection times for all household members combined was nearly three hours the prior day. We believe this collection burden is broadly representative of many rural Kenyan households without piped water at home, though we cannot confirm this because a similar figure is not collected in nationally representative surveys[4]. Although it would have been preferable to sample fewer households per location but at a broader scale, the logistics of delivering water to treated homes forced us to concentrate efforts in a small area.

A team of ten trained enumerators conducted in-person, baseline interviews in Kimeru, the local language, in early August 2016, the beginning of the dry season. Enumerators spoke with the household member "who is mostly responsible for water-related decisions such as where to get water and how much to collect". They also asked which household member spends the most time collecting water. The latter, who we call the "main water collector", is the person who reported on time use and affect in the ESM survey. Of the 248 households, 12 dropped out of the study and did not provide any meaningful number of ESM surveys. We dropped an additional 16 households

[4] The 2014 Demographic and Health Survey in Kenya found that the total roundtrip time to collect water from the house's main drinking water source is "less than 30 minutes" for 33% of rural households and "30 minutes or more" for 40% of rural households (Kenyan Bureau of Statistics 2014). To compare with our estimates, one would also need to know the number of trips taken per day by all collectors.

because enumerators expressed multiple concerns that the person carrying the phone was not the main water collector. Further data cleaning, described below, dropped all of the ESM records for some households, leaving a total sample of 195 households.

In 90% of the 195 households remaining, the person who was responsible for making water-related decisions was also the main water collector. The typical (median) main water collector was a 37-year-old woman with eight years of education who could read “with difficulty”. Ninety-three percent of respondents were women. Sixty-five percent of respondents were age 40 or under; the full age distribution is shown in Appendix Figure A 1. Only three main collectors were under age 18 (the youngest was 12 years old), though a typical household had one or more children supplementing the water collection of the main collector, typically the mother. These estimates are roughly in line with rural Kenya overall. The 2014 Kenyan DHS found that the person “who usually collects water” is a woman over age 15 in 77% of rural households, a man over 15 in 19% of households, and a child younger than 15 in only 5% of households.

Nearly all main water collectors (98%) said that they work on the household’s own farm, and 39% said they have worked for wages in the past two weeks. On average, they had worked for 4.7 days in the past 14 days and earned an average daily wage of 250 Kenyan shillings (Ksh), about USD 2.45 (~102 Ksh/USD in August 2016). Among those who worked for wages, 29% said the wage work was casual labor, 40% performed wage labor on someone else’s farm, and 23% were self-employed entrepreneurs. Only four respondents were employed in the formal sector.

The water carriers in treatment and control households were not statistically different from each other on all observables (Table A 1). At the level of the household, by random chance our treatment households were less likely to have a dirt floor, owned more acres of land and were more likely to report treating drinking water (Table A 2).

Methods

Experience Sampling Method (ESM) survey

Each participating household was given a low-cost (USD 20) smartphone, a solar charger, and a SIM card which was loaded daily with enough airtime credit to transmit any completed forms. ESM surveys were conducted on a custom Open Data Kit (ODK) app[5] that first asked about the "primary activity" the respondent was doing when the phone buzzed, with 18 time use categories. The time use categories and descriptions (Appendix Table A 3) were carefully explained to each participant at baseline as part of the recall-based time use elicitation[6]. Since half of respondents said they could read "with difficulty" and 28% said they could not read at all, we made the app flexible for users to interact in the way most comfortable to them. We assigned a photo for each time use category, and users could push a button in the app to have the program play back a recording of each of the detailed descriptions for each activity, read in Kimeru. Since many educated Kenyans prefer to read in English, the category headings could be shown in either English or Kimeru.

The ESM survey asked a number of questions about that primary activity, including a) follow-up questions for households who said they were farming, going to market, or working that asked the type of work, if they worked for wages, and if they worked for themselves or others, b) how much they enjoyed the activity, c) whether it was important to them and others, d) whether they wished they had been doing something else, e) a secondary (concurrent) activity (if any), f) who they were with when the phone buzzed, g) whether they were in any physical pain or discomfort, h) whether they felt safe, and i) affect on three dimensions (happy vs. sad, tired vs. energetic, lonely vs. sociable). Each of the affect measures were asked of respondents as a 7-point Likert scale, which we transform to a 0 to 100 scale for ease of interpretation, following similar cardinality assumptions made in Bryson and MacKerron (2017)[7]. The question on safety (very unsafe, somewhat unsafe, somewhat safe and very safe) was similarly transformed to a 0 to 100 scale. We collapsed responses to the question on physical pain ("none", "slight pain" and "severe pain") into a dummy variable that is equal to one if the respondent reported either slight or severe pain since we believe these response options are less likely to be cardinal.

---

[5] The underlying file needed to replicate the ODK survey are available at https://tinyurl.com/y8o33vcm , as is a video showing how the app would have looked and sounded to respondents (on the actual phones).

[6] In this exercise, enumerators asked respondents to "reconstruct" the prior day, minimizing possible recall bias where contemporaneous diary approaches are infeasible due to low literacy rates (see Masuda et al.(2014) and Hoque and Hope (2018) for other pictorial diary approaches). We focus here on time use as derived from ESM.

[7] For the happiness measure, for example, the options were "very sad" (coded in the data as 1), "quite sad" (2), "a bit sad" (3), "neither happy for sad" (4) and "a bit happy" (5), "quite happy" (6) and "very happy" (7). Transformed onto a 0 to 100 scale, 0 is "very sad", "quite sad" is assigned 16.7 (100/6 intervals), "a bit sad" is 33.3, etc. up to "very happy" (100).

We adapted an existing Android app to generate a randomly-timed prompt to complete the ESM survey at four randomly-chosen times during waking hours, six days a week. Because timestamps from this reminder app were not saved on the phone (as planned), we are unable to verify that respondents completed the ESM survey just after being prompted. The ESM app, however, did capture timestamps of when surveys were started and submitted.

To guard against the phone being taken by another household member, we asked the main collector in the baseline survey to choose a pictorial password (a picture of an East African animal) that would be her "secret animal" that should not be shared with others. The ESM survey first asked for the "secret animal" to help us identify that the right respondent was answering (though we cannot rule out that the real respondent simply forgot). We monitored ESM data for incorrect passcodes on an ongoing basis, and had a member of the team visit suspect households and warn them that their participation in the program could end (losing the smartphone and solar charger) if the main water collector was not the person completing the surveys.

As described in Appendix C, we dropped a total of 4,450 ESM records that a) could not be matched back to a household based on the phone's unique ID; b) were records likely generated during training or programming of the phones; c) had implausible time stamps from the network; or d) had an incorrect "password" and were thus likely to have been filled out by another household member. In an additional 9,934 ESM records, the timing and length of ESM records suggests that the survey was either not completed immediately after being prompted or was completed without any prompting at all (e.g. completing more than four surveys per day). Rather than discard these records, we flag them as potentially problematic. The results presented below do <u>not</u> use these records and are estimated on a final sample of 9,559 records from 195 respondents, though sample sizes vary by measure because of item-level missing data

Appendix Figure A 2 shows that the average number of submitted surveys was stable over the main six weeks of the study period (Panels A and B). The average submissions per day is less than four: as expected, respondents did not on average respond to all four prompts from the phone each day. Panel C shows that surveys were submitted throughout the day, and Panel D shows balance by day of the week (the program did not prompt ESM surveys on Sundays). Finally, Panel E shows that the time between when the ESM program was started by the respondent and when it was submitted as complete was under 10 minutes for 95% of surveys. The average and median times were 4.2 minutes and 2.7 minutes.

Survey and school attendance for school-aged children

We also asked the main water decision-maker (the respondent to the baseline survey) to fill out a different form on the ODK app once each evening during the week. The reminder app generated an alert each weekday evening at 8:00pm. This form asked the respondent to report, for each school-aged child in the household, whether the child attended school that day, how many hours the child spent studying (seven categories, transformed to a continuous measure using the midpoint of the intervals), and whether the child performed any household chores that day. After removing names which were either blank or could not be matched to a name in our household roster, we collected 3,891 daily records from 389 school-aged children (183 control, 206 treatment) in 205 households (100 control, 105 treatment).

We worked with all 10 primary and secondary schools in the area to obtain their written attendance logs during the study period (Aug.29 – Oct.14, 2017, or 27 school days), including the period before treatment. We were able to manually match student names in the attendance logs to 242 children (n=124 treatment and n=118 control) in 154 of our study households (n=79 control, n=75 treatment). By chance, school attendance before treatment was higher for children in control households (91.5% attendance) compared to children in treatment households (88%), a statistically-significant difference.

Randomized water delivery

We held two meetings of participants to randomize them into treatment groups. Participants drew numbers from an urn to determine treatment status. Control households were given cooking oil as an incentive for participation (beyond the phone and charger), and vended households were given a 100L storage tank to ensure that they would have enough capacity on site to hold one day's worth of vended water. At project end, control households were given another bottle of oil and the same 100L storage tank for equity[8]. Because cooking oil has cash value, our intervention may have induced control households to purchase more vended water themselves, biasing our treatment effects downwards.

[8] The study received human subjects' approval from the University of Washington Institutional Review Board.

To deliver water to the 92 treated households, we rented a 1000L tanker truck and negotiated with a team of approximately 20 water vendors who followed the tanker and walked or biked the water the last distance from the road to the household. We encountered logistical problems during the first two weeks[9], and consider this time period as partially treated: some households who should have received water on a specific day did not. We have information on which households were missed, and drop those households from the analysis for periods in which they should have been receiving water but were not. Because the main limiting step was the tanker truck, a second was brought from Nairobi at the end of the second treatment week. Treatment was then consistent for the remaining 3 weeks of the study period. Appendix B explores the sensitivity of our results to two alternative definitions of treatment.

Model

We exploit the panel nature of the dataset to explore the impact of the randomly-assigned water deliveries on affect and time use for the household's main water carrier, and on school attendance, minutes spent studying, and chores for all school children in the household. For three affect measures (happy, sociable, energetic), the dependent variable is the response to each of the affect questions, scaled from 0 to 100 as described above. Although it is not strictly speaking a percentage, we will refer to these coefficients as percentage changes, so that a positive coefficient of 6 points on the 0-100 scale is a 6% increase in happiness. For dichotomous variables (time use categories, school attendance, probability of chores, and perception of safety) we estimate a linear probability model.

By estimating a model with person fixed-effects, we exploit variation within each person's reporting of her time use, happiness, safety, etc. as the ESM found them doing, and as a function of treatment status. Similarly, we report variation in outcomes for children as a function of treatment status.

The estimation model is:

(1)

$$h_{it} = \propto_i + \delta \cdot Treat_t + \epsilon_{it}$$

[9] These included a) the water source at the local water utility being open half as often as the utility told us initially, b) the pump on the truck being unable to pull water from the tank, c) flat tires, and d) vendors who did not consistently arrive for work.

where *h* is the outcome measure of individual *i* at time *t*. Standard errors are clustered by respondent[10]. Models for affect also include controls for time of day (e.g. morning, afternoon, evening). We also explored random-effects models to increase efficiency and reduce standard errors. Hausman tests indicated that a random-effects estimator would be consistent for most but not all outcome measures, so to avoid confusion we report only results from the fixed-effects estimator[11].

Results

Baseline survey – do women enjoy collecting water?

Before turning to the ESM data, we begin by briefly discussing survey questions added specifically to understand attitudes towards water collection. Based on a few simple questions, we find no evidence in our sample to support the belief that women enjoy the activity of water collection. First, only 22% percent reported that they combine trips collecting water with other activities such as stopping to visit friends, going to the market, or doing other work. The vast majority of respondents collect water by themselves; only 11% regularly reported collecting water with another person. Second, we directly asked respondents how much they enjoy collecting water. Three-quarters said they “dislike it a lot”, and 19% dislike “a little”. Only 6% said they enjoy it a “little” or “a lot”.

Interestingly, we find more indications of the social nature of water collection in the ESM data than in our simple question above, where only 11% reported regularly collecting with someone else. Among the ESM records where the primary activity was collecting water, half (49%) reported being with someone else. In two-thirds of these cases, respondents said they were with a friend.

[10] Implemented with *xtreg,fe vce(cluster respondent_id)* in Stata 15.
[11]Hausman tests of fixed-effects vs. random-effects models for each of the six affect measures indicated that random-effects models would be consistent for the happiness (fail to reject null of no systematic difference in coefficients at p=0.18), sociability (p=0.42), safety measures (p=0.90), and some pain (p=0.40). We weakly reject the null for energy (p=0.08) and ”wish do else” (p=0.09), which would indicate that only a fixed effect estimator is consistent. Similar tests for time use categories found that we would reject the null for ”household work” (p=0.09) and ”bathing and washing” (p=0.02).

What changes did households report?

During the midline and endline surveys, households were asked to describe any changes, including to wages or working hours, since the project was implemented. Twenty four percent of treatment households said that they worked for fewer hours, 52% of the households reported working more, and 15% reported having started working. Twenty-seven percent said they had increased incomes. Twenty two percent of the households said that they had more time for farming, 57% reported more time for household chores, 13% reported that children had more time to study, and 7% reported more time to do their business. Additionally, 7% of the households said that they had more energy to work on other activities. Seventeen percent perceived improved hygiene and sanitation due to more water available. Four percent of treatment households reported having less social time since there was less time spent collecting water and interacting with others in the process.

Econometric results

Table 1 reports the treatment effect of receiving vended water at home: water carriers in treatment households felt 2.7% happier, 5.6% more energetic, and 4.4% more safe than control households. They are 7.9 percentage points less likely to report feeling some or severe pain, compared to a mean in the control group of 25%. We find no evidence that having water delivered at home reduces water collector's perception of sociability: the coefficient is positive but is not statistically different than zero. We also find impacts on self-reported agency: the respondent is 7.5 percentage points less likely to report that she wished she had been doing something else.

Table 1. Impact of receiving vended water treatment on affect

	Happy	Sociable	Energetic	Safe	Some pain	Wish Do Else
Treatment	2.65**	1.61	5.63***	4.44***	-0.078***	-0.075***
	(1.07)	(1.60)	(1.67)	(1.29)	(0.021)	(0.019)
Alone	-0.54	-12.44***	-3.41***	-0.94*	0.021	0.023**
	(0.53)	(1.49)	(0.79)	(0.52)	(0.012)	(0.010)
N=	8,454	8,425	8,440	8,458	8,466	8,434
Groups =	195	195	195	195	195	195
Overall R^2	0.0035	0.039	0.047	0.011	0.012	0.0064
Mean in control group	79.1	63.1	34.5	70.7	0.24	0.17

Notes: Fixed-effects linear panel data model (Stata: *xtreg, fe*). Robust standard errors clustered at the respondent level. Controls for time of day (morning, afternoon or evening) when the ESM record was completed are included in the model but not reported (results available on request). * = 90%, **95%, ***99%.

Table 2 shows the impact on time use (the activity the respondent reported doing when the ESM survey went off). As expected, the treatment was effective in reducing water collection times. The control group reported water collection in 12.4% of ESM records, and treatment reduces this probability nearly to zero (by 10.4%). Multiplying these fractions by the average waking hours in the sample (15.5) gives an estimate of the number of minutes saved. This time savings of approximately 95 minutes was largely reallocated to household work (45 minutes), leisure (25 minutes) and work on the household's own farm (20 minutes). We find no impact on paid labor, though the fraction of respondents with paid jobs is low at baseline (3.3% in the control group).

Table 2. Impact of receiving vended water on time use of main water collector

	Water collection	Household Work	Leisure	Social	Own farm	Meetings	Caring for others	Going to Market
Treatment	-0.104*** (0.013)	0.051*** (0.017)	0.0203** (0.0080)	0.0025 (0.0089)	0.0252* (0.013)	-0.0068 (0.0083)	$5.5e^{-3}$ ($8.0e^{-3}$)	$3.5e^{-3}$ ($6.8e^{-3}$)
Mean in control group	0.128 (119 mins)	0.296 (275 mins)	0.083 (78 mins)	0.081 (76 mins)	0.065 (60 mins)	0.070 (65 mins)	0.052 (48 mins)	0.051 (48 mins)
Treatment effect in minutes: mean (95% CI)	-96 mins (-72,-120)	+47 mins (17,78)	+19 mins (4,33)	n/a	+23 mins (0,47)	n/a	n/a	n/a

Table (cont'd)

	Paid work	Livestock	Washing	Wood collection	Reading, studying, school	Private	Informal unpaid work
Treatment	$1.47e^{-3}$ ($7.6e^{-3}$)	0.011 (0.0074)	$3.3e^{-3}$ ($6.0e^{-3}$)	$2.7e^{-3}$ ($5.4e^{-3}$)	$-5.3e^{-4}$ ($4.3e^{-3}$)	$2.2e^{-3}$ ($4.8e^{-3}$)	$-1.9e^{-3}$ ($3.2e^{-3}$)
Mean in control group	0.033 (31 mins)	0.038 (35 mins)	0.034 (31 mins)	0.024 (22 mins)	0.017 (16 mins)	0.012 (11 mins)	0.013 (12 mins)
Treatment effect in minutes: mean (95% CI)	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Notes: All results based on 8,561 ESM records from 195 groups. Fixed-effects linear probability panel data model (Stata: *xtreg, fe*). Robust standard errors (shown in parentheses) clustered at the respondent. "Leisure" includes sleeping, playing sports/hobbies, listening to TV/radio, socializing or doing "not much of anything". "Paid work" includes paid work in the formal or informal sector, including paid work on someone else's farm, and time spent in one's own private business. * = 90%, **95%, ***99%.

Tables 3 and 4 report the impact of treatment on outcomes for all school-aged children in the household (not just those who collect water). Reducing water collection time to zero at the household level increases teacher-recorded daily school attendance by 3.6 percentage points (Table 3, Model A). Model B shows that treatment effects are concentrated in children aged 5-10 (the omitted category). We find no difference in treatment effects by gender. The effects on attendance as self-reported by parents in the daily phone survey are larger: treatment increased attendance by 14.1 percentage points (Model C); the average daily attendance rate in the control group is 82%. We find no differential impacts by gender or age (Model D).

Table 3. Impact of treatment on school attendance and time spent studying by school-aged children (household-fixed effects OLS)

	(A) School-reported attendance	**(B)** School-reported attendance	**(C)** Self-reported school attendance	**(D)** Self-reported school attendance
Treatment	0.036 ***	0.065***	0.141***	0.125***
	(0.013)	(0.025)	(0.027)	(0.035)
Trt*Girl		-0.009		0.042
		(0.032)		(0.038)
Trt*Age10-15		-0.061**		-0.012
		(0.027)		(0.038)
Trt*Age16-19		0.070		0.0247
		(0.059)		(0.083)
N (groups)	6,483 (241)	6,483 (241)	3,831 (382)	3,831 (382)
Mean in control group	0.92	0.92	0.82	55 minutes

Notes: Columns A and B use school-reported attendance; columns C and D use data as self-reported by households. Robust standard errors clustered at the respondent/household level for all models. * = 90%, **95%, ***99%.

Compared to an average of 55 minutes spent studying per day in the control group, treatment increases studying by 8 minutes, a 15% increase (Table 4). We find no differential treatment effects by gender or age.

Table 4. Impact of treatment on parent-reported time spent studying by school-aged children (household-fixed effects OLS)

	(A) Minutes spent studying	**(B)** Minutes spent studying
Treatment	7.85**	9.73**
	(3.07)	(4.84)
Trt*Girl		-5.07
		(4.65)
Trt*Age10-15		-2.35
		(4.57)
Trt*Age16-19		11.47
		(10.45)
N (groups)	3,707 (381)	3,707 (381)
Mean in control group	55 minutes	55 minutes

Notes: Minutes studying as self-reported by parents on a mobile phone survey conducted at the end of each school day. Robust standard errors clustered at the respondent/household level for all models. * = 90%, **95%, ***99%.

As expected, treatment reduces the probability that a child is reported having collected water in the previous day (). In 31% of control group records, children were reported as collecting water that day; treatment reduced this by 25.6 percentage points, an 83 percent reduction. Our results suggest that children were not freed from all chores: treatment has no statistically significant impact on the probability of doing "no chores". Rather, children in treated households are more likely to reallocate some of the water collection time savings to cleaning and cooking chores (Table 5). We find no differential effects for any of the measures in Table 5 except for water collection. The reduction in the probability of collecting water is smaller for girls and larger for those aged 10-15 compared to those aged 6-10 (results available on request).

Table 5. Impact of treatment on chores performed by school-aged children

	Water collection	No Chores	Wood collection	Cleaning	Cooking	Farm/ Livestock
Treatment	-0.256*** (0.032)	0.0425 (0.030)	0.029 (0.038)	0.087*** (0.021)	0.044* (0.023)	0.011 (0.017)
Mean in control group	0.31	0.28	0.20	0.13	0.10	0.069

Notes: All regressions based on 3,891 survey responses for 384 school-aged children. Robust standard errors are clustered at the household level. * = 90%, **95%, ***99%.

One possible validity threat arises if treated households share some of their vended water with other households, a practice that is quite common globally (Rosinger et al. 2020). If treatment households were in fact receiving more water than they could productively use, and if they shared with neighbors who were not part of the study, this sharing should not affect our results. Two different circumstances could bias our results. First, if treatment households shared excess water with households in the control group, thereby reducing the counter-factual water collection times and affect measures. Second, if treatment households felt guilty about receiving vended water, gave some of it away to neighbors, and as a result did not reduce their collection times to zero. To measure this threat we asked about water sharing during both the midline and endline surveys.

Treated households did share water. Fifty-eight percent of treated households reported sharing water at midline, falling to 48% at endline after asking these households not to share water. All but one household shared water without asking for payment, and the majority shared 20 liters per day (median 20L, mean 25L). We find little evidence that guilt-driven sharing led households to do more water collection during the time they were receiving vended water: using our preferred definition of treatment, sharers did not spend more time collecting water than non-sharers (point estimate of difference=1.8 minutes, t=0.44).

Five households told us in the midline survey that they shared with a household who was also participating in the study (two at endline). We also asked households in the control group directly if they were receiving water from treatment households: sixteen percent said they were at midline, and 8% at endline. To be conservative, we re-estimated our models dropping 23 households in the control group who reported receiving water and an additional seven who did not but were mentioned by treatment households as receiving their vended water. Because our preferred approach uses a fixed-effects model, dropping control households shifts the estimated intercept of the model and slightly increases the standard error of our point estimates but leaves the coefficients unchanged. When estimating random-effects models, the results for water carrier affect were unchanged and changed in only one respect for time use: the point estimate of the predicted increase in leisure time falls from +25 minutes to +18 minutes. The estimates for schoolchildren also show identical effects for school attendance and minutes studying and results for chores are unchanged (all results available on request).

Appendix B shows that our results are generally robust to different definitions of treatment status, given that the uneven logistics of our water vending early in the treatment period (discussed above). A "conservative" definition drops all records from treated households in the first two weeks. Because most of the problems happened in the first two days of water vending as logistical issues were worked out, a second "simple" treatment definition drops all observations from all treated households only for those first two days. Our well-being results show the same sign and pattern of significance in these two alternate definitions, with most treatment effects increasing slightly in strength (Table A 4). The results for time reallocation are also similar (Table A 5). The impact on school-recorded attendance increases from 3.6 percentage points to 4.8 percentage points under the conservative definition and 4.4 under the simple definition Table A 6). The interactions with gender and age remain. Our results on the number of minutes spent studying are somewhat smaller and statistically significant in the simple treatment definition, but are insignificant in the conservative definition (Table A 7). The effect of treatment on children's chores is very similar Table A 8).

Finally, Appendix C explores the sensitivity of our results to our decisions in dropping ESM records that we suspected could have been either duplicates or incorrectly completed. The pattern of results for well-being is largely the same when we estimate models that included these "flagged" responses, with positive treatment effects increasing for happiness, safety and likelihood of pain and decreasing for energy. Although our preferred results show no effect on sociability, results using the larger dataset show a positive 2.23% increase in feelings of sociability as a result of treatment (Table A 9). The results for time allocation including flagged records are similar for water collection, but show a somewhat smaller reallocation to household work and own-farm work. These models also imply less time reallocated to leisure and more reallocated to socializing, caring for livestock and washing/bathing, though the treatment effects are small in magnitude and only marginally significant (Table A 10).

Limitations

We acknowledge several limitations of the study. First and most importantly, we studied only the short-term impacts of reducing water collection times. One might expect short-term impacts to be similar to long-run impacts for our measures of emotional well-being and perhaps for children's school attendance and studying. But households who have "solved" their water collection problem more permanently are likely to readjust time use patterns and labor supply in a manner that may not approximate the pattern our short-term solution provided them. Second, our study was conducted in the dry season when water is scarcest on the landscape and at a relatively slack time in the agricultural calendar. Households in the area rely much more heavily on rainwater collection during the rainy season (Cook, Kimuyu, and Whittington 2016). On one hand, this implies the benefits of improving water supply access may be lower than during the dry season we observed. On the other hand, demand for female labor on the family farm may be higher in those times, and water carriers may reallocate more saved collection time to farming and less to leisure and other household chores than we observed. Third, the external validity of our results may be limited by its small geographic scope. This represented a tradeoff between internal and external validity, given that the logistics involved in effectively randomizing water quantity changes precluded a larger geographic scope. Fourth, our high-frequency ESM data collection approach was logistically difficult to implement and may have been burdensome to respondents. Because of malfunctions with the time-stamp feature of our ESM program malfunctioning, we rely on several assumptions to clean the data which could be questioned, though we demonstrate our results are not overly sensitive to those assumptions. Still, we feel the approach is promising for studying time use and for applications where moment-by-moment well-being are important. We know of only one other ESM application in a low-income setting, and the data was collected with frequent short phone calls rather than self-administered (Miñarro et al. 2021). Finally, we acknowledge the possibility of Hawthorne effects. All study participants were well-aware that the study focused on the impacts of water collection. It is possible that households selectively reported activities accordingly, perhaps over-stating the time spent collecting water. Private activities may have been under-reported (recall that the survey included a category called "private"), though this under-reporting should have been unrelated to treatment status. Our time reallocation results are also unlikely to be affected by Hawthorne effects or social desirability bias. We observed that the impacts on school attendance as reported by parents were much higher than those derived from school records. Households in the treatment group may also have shaded their responses to emotional well-being upwards; households in the control group may have given lower scores partly in protest for not being chosen for the treatment arm.

Conclusions

We conclude with three messages for policy analysts and researchers working on the benefits of improving water supply access and reliability. First, there is now substantial evidence on the negative physical, emotional and psychological impacts of women's water collection burdens, though translating these impacts into monetized welfare changes will be challenging. Devoto et al. (2012) found that reduced collection times led to higher overall evaluative life satisfaction scores and less conflict with neighbors. They also found a positive but statistically insignificant effect on a summative index of emotional (momentary) well-being along dimensions of sadness, worry and satisfaction. Our study finds positive and significant effects on momentary happiness. We also find women have more energy, are in less pain, and feel safer and more in control of their time. These echo findings from non-experimental studies. Chindarkar (2019) found that households in Kathmandu (Nepal) with higher coping costs for unreliable, poor quality water supply had lower emotional well-being. Bisung and Elliott (2017) review fifteen studies finding negative psychosocial impacts of water collection.

Although there is anecdotal evidence that women value socializing while collecting water, we find little evidence that reducing water collection times reduced how frequently women reported feeling sociable. In our preferred model, the effect was positive but not statistically significant, suggesting women were able to substitute other less physically-taxing ways of socializing (the effect was positive and statistically significant in one of our sensitivity analyses). It is possible, of course, that this result is unique to our study setting in rural Kenya. We did not specifically explore household bargaining power or women's roles in the household. It may be the case that in other settings where women's autonomy is very low, water collection may be both sociable and an escape for women. Women experiencing domestic violence at home may be more likely to appreciate the opportunity to collect water: Fajardo-Gonzalez (2017) finds that incidence of domestic violence increased the labor supply of women in Colombia. The effect of socializing versus being alone on overall happiness is also likely to vary culturally: in other sites the utility of being with others may outweigh the physical discomfort of water collection. However, without careful empirical study - directly asking women themselves – development professionals should be especially careful not to assume that women enjoy water collection, or that infrequent use of an improved water source is attributable to this type of preference. It is also possible for development policies to support other types of activities that provide women autonomy and time socializing without the physically-demanding drudgery of water or fuelwood collection.

Second, the evidence remains mixed about whether children's resource collection work hinders their ability to attend and succeed in school, but the baseline water collection burden of children surely matters. Most evidence has been from observational studies. In rural Tanzania, Akabayashi and Pacharopoulos (1999) found that increased distance to the water source was associated with increased work hours for both boys and girls, and girls worked approximately 45 minutes more than boys. They found no association, however, with school enrollment or study hours. In Ethiopia, Haile and Haile (2012) found that girls in households farther from water sources were more likely to engage in domestic tasks and not attend school, while boys were more likely to attend school and not work. In Kenya, Ndiritu and Nyangena (2011) found children spending more than two hours collecting water or firewood were 21 per cent less likely to be attending school. In Malawi, Nankhuni and Findeis (2004) examined both urban and rural areas and found that piped water access in the home significantly reduced children's time spent collecting water among rural households, and was also positively associated with child school enrollment. Koolwal and Van de Walle (2013) found a statistically significant and positive effect of water access on school enrollment in Yemen, Morocco, Nepal and Pakistan, but no effect in the four Sub-Saharan African countries studied. Moving beyond cross-sectional analyses, Nauges (2017) built a panel of GPS-identified communities in rural Ghana and examined the effects of water access on girls' school attendance. They found girls in male-headed households had lower school attendance, and estimated school attendance to increase by, on average, 2.4 per cent if time spent collecting water were reduced by half in all communities. In urban Morocco, Devoto et al. (2012) found no impact on school absenteeism or the hours children spend studying. In a quasi-experiment in urban Zambia, however, Ashraf et al. (2021) finds suggestive evidence that water system outages reduce the amount of time girls spend studying, though the effect was not statistically significant. We find that reducing collection burdens increased school-observed attendance from 92% to 96% and increased the amount of time children spent studying.

Finally, water sector professionals cannot assume that time freed from water collection will be devoted solely to income-generating activities, for example by valuing reduced collection times at 100% of rural unskilled wages in a benefit-cost analysis. It is understandable that many sector advocates would like to believe that freed time can be converted to cash to either improve household incomes or contribute to the cost-recovery of improved infrastructure through user fees, but the reality is more complex. Although cross-sectional studies have found that women report having more time to work (Winter, Darmstadt, and Davis 2021; Crow, Swallow, and Asamba 2012), these rely on retrospective assessments by women rather than detailed time use diaries or labor supply data. Several quasi-experimental and experimental studies have found no impact of water collection on the probability of women undertaking paid work or household income (Devoto et al. 2012; Gross et al. 2018; Koolwal and Van de Walle 2013). However, in a quasi-experimental approach in rural Kyrgyzstan, Meeks (2017) found that the roughly 170 minutes of time savings (based on detailed time diaries) were reallocated roughly equally to leisure (80 minutes) and farm labor (90 minutes). Because the time savings resulted in measurable increases in cereal production, Meeks (2017) is able to assign a shadow value to the time of roughly 50% of reported farm wages. Although we find no impact on paid labor, we also find that some of the time is reallocated to farm labor, though a smaller fraction (one-fifth of saved time). The majority is spent on leisure and other household chores.

Like the gains to emotional well-being, any time reallocation resulting from lower water collection times is undoubtedly welfare-improving (see Koolwall and van de Walle (2013) for a theoretical model and proof) but difficult to value economically in a benefit-cost context. This points to the general need for more studies examining the value of time in non-market activities in low-income countries, particularly rural areas with few wage-earning opportunities (Whittington and Cook 2019). It also suggests the importance of examining the entire time budget of the household, including firewood collection, unpaid caregiving, household chores, etc. Over three decades ago, Briscoe (1987) argued that improvements in water and sanitation infrastructure were a necessary but insufficient condition to dramatically improve childhood survival from diarrheal disease because the etiology of diarrhea is so multi-faceted. It was only once a sufficient number of disease transmission pathways had been closed by earlier interventions that the "last" intervention would cause large, observable health effects. It may be that until women have access to a suite of labor-saving technologies like modern energy services, at-home water service, and washing machines, much of their time will unfortunately be spent on physically taxing drudgery. Reductions in one type of drudgery (e.g. water collection) will mainly allow more time for other drudgery (e.g. firewood collection or manual farm labor).

References

Ahuja, Amrita, Michael Kremer, and Alix Peterson Zwane. 2010. "Providing Safe Water: Evidence from Randomized Evaluations." *Annual Review of Resource Economics* 2 (1): 237–56. https://doi.org/doi:10.1146/annurev.resource.012809.103919.

Akabayashi, Hideo, and George Psacharopoulos. 1999. "The Trade-off between Child Labour and Human Capital Formation: A Tanzanian Case Study." *The Journal of Development Studies* 35 (5): 120-140. https://doi.org/10.1080/00220389908422594.

Allcott, Hunt, Luca Braghieri, Sarah Eichmeyer, and Matthew Gentzkow. 2020. "The Welfare Effects of Social Media." *American Economic Review* 110 (3): 629–76. https://doi.org/10.2139/ssrn.3308640.

Arku, Frank S. 2010. "Time Savings from Easy Access to Clean Water: Implications for Rural Men's and Women's Well-Being." *Progress in Development Studies* 10 (3): 233–46. https://doi.org/10.1177/146499340901000303.

Ashraf, Nava, Edward Glaeser, Abraham Holland, and Bryce Millett Steinber. 2021. "Water, Health and Wealth." *Economica* 88: 755–81.

Bisung, Elijah, and Susan J. Elliott. 2017. "Psychosocial Impacts of the Lack of Access to Water and Sanitation in Low- and Middle-Income Countries: A Scoping Review." *Journal of Water and Health* 15 (1): 17–30. https://doi.org/10.2166/wh.2016.158.

———. 2019. "Community Water Supply Improvement and Wellbeing: A Pre-Post Photovoice Intervention Study in Kenya." *Habitat International* 85 (January): 14–20. https://doi.org/10.1016/j.habitatint.2019.01.005.

Boone, Christopher, Peter Glick, and David E. Sahn. 2011. "Household Water Supply Choice and Time Allocated to Water Collection: Evidence from Madagascar." *Journal of Development Studies* 47 (12): 1826–50. https://doi.org/10.1080/00220388.2011.579394.

Briscoe, John. 1987. "A Role for Water Supply and Sanitation in the Child Survival Revolution."

Bulletin of the Pan American Health Organization 21 (2).

Bryson, Alex, and George Mackerron. 2017. “Are You Happy While You Work?” *Economic Journal* 127 (599): 106–25. https://doi.org/10.1111/ecoj.12269.

Cairncross, S, and J L Cliff. 1987. “Water Use and Health in Mueda, Mozambique.” *Transactions of the Royal Society of Tropical Medicine and Hygiene* 81 (1): 51–54. http://www.sciencedirect.com/science/article/B75GP-4C0DTJC-1FD/2/67b6936379e7768f95202bfe1cc603d9.

Cairncross, and Vivian Valdmanis. 2007. “Water Supply, Sanitation and Hygiene Promotion.” In *Disease Control Priorities in Developing Countries*, edited by Dean Jamison, Joel Breman, Anthony Measham, George Alleyne, Mariam Claeson, David B Evans, Prabhat Jha, Anne Mills, and Philip Musgrove, 2nd ed. Oxford University and the World Bank.

Chandrasekaran, Maya, Joseph Cook, and Marc Jeuland. 2021. “The Evidence Base for Time Savings Benefits in Water and Sanitation Interventions.” *Oxford Research Encyclopedia of Global Public Health*.

Chen, Yvonne, Namrata Chindarkar, and Jane Zhao. 2019. “Water and Time Use : Evidence from Kathmandu , Nepal.” *Water Policy* 21 (S1): 76–100. https://doi.org/10.2166/wp.2019.082.

Chindarkar, Namrata, Yvonne Jie, and Chen Yogendra. 2019. “Subjective Well - Being Effects of Coping Cost : Evidence from Household Water Supply in Kathmandu Valley, Nepal.” *Journal of Happiness Studies* 20: 2581–2608. https://doi.org/10.1007/s10902-018-0060-6.

Cockburn, John, and Benoit Dostie. 2007. “Child Work and Schooling: The Role of Household Asset Profiles and Poverty in Rural Ethiopia.” *Journal of African Economies* 16 (4): 519–63. https://doi.org/10.1093/jae/ejl045.

Cook, Joseph, Peter Kimuyu, and Dale Whittington. 2016. “The Costs of Coping with Poor Water Supply in Rural Kenya.” *Water Resources Research* 52: 841–59.

Crow, Ben, Brent Swallow, and Isabella Asamba. 2012. “Community Organized Household Water Increases Not Only Rural Incomes, but Also Men’s Work.” *World Development* 40 (3): 528–41.

Devoto, Florencia, Esther Duflo, Pascaline Dupas, William Parienté, Vincent Pons, William Pariente, and Vincent Pons. 2012. “Happiness on Tap: Piped Water Adoption in Urban Morocco.” *American Economic Journal: Economic Policy* 4 (4): 68–99.

https://doi.org/10.1257/pol.4.4.68.

Dickinson, Katherine L, Sumeet R Patil, Subhrendu K Pattanayak, Christine Poulos, and Jui-chen Yang. 2015. "Nature's Call : Impacts of Sanitation Choices in Orissa , India." *Economic Development and Cultural Change* 64 (1): 1–29.

Dinkelman, Taryn. 2011. "The Effects of Rural Electrification on Employment : New Evidence from South Africa †." *The American Economic Review* 101 (December): 3078–3108.

Fajardo-gonzalez, Johanna. 2017. "Domestic Violence , Decision-Making Power and Female Employment in Colombia 1 Introduction."

Gibson, M A, and R Mace. 2006. "An Energy-Saving Development Initiative Increases Birth Rate and Childhood Malnutrition in Rural Ethiopia." *PLoS Medicine* 3: e87.

Graham, Jay P., Mitsuaki Hirai, and Seung-Sup Kim. 2016. "An Analysis of Water Collection Labor among Women and Children in 24 Sub-Saharan African Countries." *Plos One* 11 (6): e0155981. https://doi.org/10.1371/journal.pone.0155981.

Gross, Elena, and Ossama Elshiewy. 2019. "Choice and Quantity Demand for Improved and Unimproved Public Water Sources in Rural Areas: Evidence from Benin." *Journal of Rural Studies* 69: 186–94. https://doi.org/10.1016/j.jrurstud.2019.02.021.

Gross, Elena, Isabel Guenther, Youdi Schipper, and Van Der Walle. 2018. "Women Are Walking and Waiting for Water: The Time Value of Public Water Supply." *Economic Development and Cultural Change*.

Haile, Getinet, and Beliyou Haile. 2012. "Child Labour and Child Schooling in Rural Ethiopia: Nature and Trade-Off." *Education Economics* 20 (4): 365–85. https://doi.org/10.1080/09645292.2011.623376.

Hope, R. 2006. "Evaluating Water Policy Scenarios against the Priorities of the Rural Poor." *World Development* 34 (1): 167–79. https://doi.org/10.1016/j.worlddev.2005.07.014.

Hoque, Sonia Ferdous, and Robert Hope. 2018. "The Water Diary Method – Proof-of-Concept and Policy Implications for Monitoring Water Use Behaviour in Rural Kenya." *Water Policy* 20 (4): 725–43. https://doi.org/10.2166/wp.2018.179.

Ilahi, Nadeem, and Franque Grimard. 2000. "Public Infrastructure and Private Costs: Water Supply and Time Allocation of Women in Rural Pakistan." *Economic Development and Cultural Change* 49: 45–75.

Jeuland, Marc, T. Robert Fetter, Yating Li, Subhrendu K. Pattanayak, Faraz Usmani, Randall A.

Bluffstone, Carlos Chávez, et al. 2021. “Is Energy the Golden Thread? A Systematic Review of the Impacts of Modern and Traditional Energy Use in Low- and Middle-Income Countries.” *Renewable and Sustainable Energy Reviews* 135 (October 2020). https://doi.org/10.1016/j.rser.2020.110406.

Jeuland, Marc, David E Fuente, Semra Ozdemir, Maura C Allaire, and Dale Whittington. 2013. “The Long-Term Dynamics of Mortality Benefits from Improved Water and Sanitation in Less Developed Countries.” *PloS One* 8 (10): e74804. https://doi.org/10.1371/journal.pone.0074804.

Kohlin, Gunnar, and Gregory S Amacher. 2005. “Welfare Implications of Community Forest Plantations in Developing Countries: The Orissa Social Forestry Project.” *American Journal of Agricultural Economics* 87 (November): 855–69.

Koolwal, G, and D Van de Walle. 2013. “Access to Water, Women’s Work, and Child Outcomes.” *Economic Development and Cultural Change* 61 (2): 369–405. http://www.jstor.org/stable/10.1086/668280.

Kremer, Michael, Jessica Leino, Edward Miguel, and Alix Peterson Zwane. 2011. “Spring Cleaning: Rural Water Impacts, Valuation, and Property Rights Institutions.” *Quarterly Journal of Economics* 126 (1): 145–205.

Larson, R, and M Csikszentmihalyi. 1983. “The Experience Sampling Method.” In *New Directions for Methodology of Social and Behavioral Science*, edited by H T Reis. Vol. 15. San Francisco: Jossey-Bass.

Masuda, Yuta J, Lea Fortmann, Mary Kay Gugerty, Marla Smith-Nilson, and Joseph Cook. 2014. “Pictorial Approaches for Measuring Time Use in Rural Ethiopia.” *Social Indicators Research* 115 (1): 467–82.

Meeks, Robyn C. 2017. “Water Works.” *Journal of Human Resources* 52 (4): 1119–53. https://doi.org/10.3368/jhr.52.4.0915-7408R1.

Miñarro, Sara, Victoria Reyes-García, Shankar Aswani, Samiya Selim, Christopher P. Barrington-Leigh, and Eric D. Galbraith. 2021. “Happy without Money: Minimally Monetized Societies Can Exhibit High Subjective Wellbeing.” *PLoS ONE* 16 (1 January): 5–7. https://doi.org/10.1371/journal.pone.0244569.

Mu, Xinming, Dale Whittington, and John Briscoe. 1990. “Modelling Village Water Demand: A Discrete Choice Approach.” *Water Resources Research* 26 (4): 521–29.

Nankhuni, Flora J., and Jill L. Findeis. 2004. "Natural Resource-Collection Work and Children's Schooling in Malawi." *Agricultural Economics* 31 (2-3 SPEC. ISS.): 123–34. https://doi.org/10.1016/j.agecon.2004.09.022.

Nauges, Celine. 2017. "Water Hauling and Girls ' School Attendance Some New Evidence From Ghana." *Environmental and Resource Economics* 66 (1): 65–88.

Nauges, Celine, and Dale Whittington. 2010. "Estimation of Water Demand in Developing Countries: An Overview." *World Bank Res Obs* 25 (2Nauges, C., and D. Whittington. 2010. "Estimation of Water Demand in Developing Countries: An Overview." World Bank Res Obs 25(2):263–294. Available at: http://wbro.oxfordjournals.org/cgi/content/abstract/lkp016v1.): 263–94. https://doi.org/10.1093/wbro/lkp016.

Ndiritu, Simon Wagura, and Wilfred Nyangena. 2011. "Environmental Goods Collection and Children's Schooling: Evidence from Kenya." *Regional Environmental Change* 11 (3): 531–42. https://doi.org/10.1007/s10113-010-0168-7.

Ngongang, Elie. 2008. "Economics Of Water, Health, And Households' Labor Market Participation In Central African Economic And Monetary Community (CAEMC) Countries: The Case Of Cameroon." *Journal of Sustainable Development in Africa* 10 (3): 15–50. http://www.jsd-africa.com/Jsda/V10N3_Fall2008/PDF/ARC_Economics_of_WaterHealthHouseholdLaobrMarket.pdf.

Pattanayak, Subhrendu, Christine Poulos, Jui-Chen Yang, and Sumeet R Patil. 2010. "How Valuable Are Environmental Health Interventions? Evidence from a Quasi-Experimental Evaluation of Community Water Projects." *Bulletin of the World Health Organization*. Vol. 88.

Porter, Gina, Kate Hampshire, Albert Abane, Alister Munthali, Elsbeth Robson, Mac Mashiri, Augustine Tanle, Goodhope Maponya, and Sipho Dube. 2012. "Child Porterage and Africa's Transport Gap: Evidence from Ghana, Malawi and South Africa." *World Development* 40 (10): 2136–54. https://doi.org/10.1016/j.worlddev.2012.05.004.

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. 2010. "Neonatal, Postneonatal, Childhood, and under-5 Mortality for 187 Countries, 1970-2010:

A Systematic Analysis of Progress towards Millennium Development Goal 4." *Lancet* 375 (9730): 1988–2008. https://doi.org/10.1016/S0140-6736(10)60703-9.

Rosinger, Asher Y., Alexandra Brewis, Amber Wutich, Wendy Jepson, Chad Staddon, Justin Stoler, and Sera L. Young. 2020. "Water Borrowing Is Consistently Practiced Globally and Is Associated with Water-Related System Failures across Diverse Environments." *Global Environmental Change* 64 (July): 102148. https://doi.org/10.1016/j.gloenvcha.2020.102148.

Sorenson, Susan B, Christiaan Morssink, and Paola Abril Campos. 2011. "Safe Access to Safe Water in Low Income Countries: Water Fetching in Current Times." *Social Science & Medicine* 72 (9): 1522–26. https://doi.org/-9536/$ e see front matter Ó 2011 Elsevier doi:10.1016/j.socscimed.2011.03.010.

Stone, Arthur, and Saul Shiffman. 1994. "Ecological Momentary Assessment (EMA) in Behavioral Medicine." *Annals of Behavioral Medicine* 16 (3): 199–202.

UNICEF, and WHO. 2021. "Progress on Household Drinking Water, Sanitation and Hygiene."

Wagner, Jake, Joseph Cook, and Peter Kimuyu. 2019. "Household Demand for Water in Rural Kenya." *Environmental and Resource Economics* 74 (4): 1563–84.

Whittington, Dale, and Joseph Cook. 2019. "Valuing Changes in Time Use in Low- and Middle-Income Countries." *Journal of Benefit-Cost Analysis* 10: 51–72. https://doi.org/10.1017/bca.2018.21.

Winter, James C., Gary L. Darmstadt, and Jennifer Davis. 2021. "The Role of Piped Water Supplies in Advancing Health, Economic Development, and Gender Equality in Rural Communities." *Social Science and Medicine* 270: 113599. https://doi.org/10.1016/j.socscimed.2020.113599.

Wolf, Jennyfer, Paul R. Hunter, Matthew C. Freeman, Oliver Cumming, Thomas Clasen, Jamie Bartram, Julian P.T. Higgins, et al. 2018. "Impact of Drinking Water, Sanitation and Handwashing with Soap on Childhood Diarrhoeal Disease: Updated Meta-Analysis and Meta-Regression." *Tropical Medicine and International Health* 23 (5): 508–25. https://doi.org/10.1111/tmi.13051.

Wutich, Amber, Melissa Beresford, and Cinthia Carvajal. 2016. "Can Informal Water Vendors Deliver on the Promise of A Human Right to Water? Results From Cochabamba, Bolivia." *World Development* 79: 14–24. https://doi.org/10.1016/j.worlddev.2015.10.043.

Wutich, Amber, and Kathleen Ragsdale. 2008. "Water Insecurity and Emotional Distress:

Coping with Supply, Access, and Seasonal Variability of Water in a Bolivian Squatter Settlement.” *Social Science and Medicine* 67 (12): 2116–25. https://doi.org/10.1016/j.socscimed.2008.09.042.

Acknowledgments

We gratefully acknowledge funding from the Environment for Development Initiative, with support from SIDA. We thank Victor Mose, Mark Mwiti and especially Shireen Tabrizi for excellent and tireless work in the field. Rose Kathure, Gladys Kithinji, Felix Kairanya, Eric Murithi, Frida Gacheri, Martin Mutua, Dorothy Mwari, Dominic Murithi, Winniejoy Muthoni, and Millicent Kamungi collected baseline data. Judith Mutheu assisted with data cleaning and analysis. Yaw Anokwa and Pierre Biscaye provided useful advice on implementing the project in ODK and James Morris generously worked with us to adapt his Randomly Remind Me program. We thank Christine Lee for a helpful conversation about ESM, and members of the EfD Scientific Advisory Committee for advice on research design. The project would also not have been possible with the support of Andrew Kamundi, Edward Ngagi and David Kirimania at Imetha Water Company, Boniface Kirimi, the Chief of Kianjai sublocation David Nabea and sub-chiefs Joshua Arya, Bernard Gikundi, Benjamin Choker, and John Thiane. We thank Kenneth Berger for useful feedback.

APPPENDICES for "The short-run impacts of exogenously…"
Appendix A: Additional figures and tables
Appendix B: Sensitivity analysis using different definitions of treatment
Appendix C: Additional details on ESM data cleaning and sensitivity analysis
Appendix D: Additional details on sampling

Appendix A: Additional tables and figures

Figure A 1. Age distribution of phone carriers

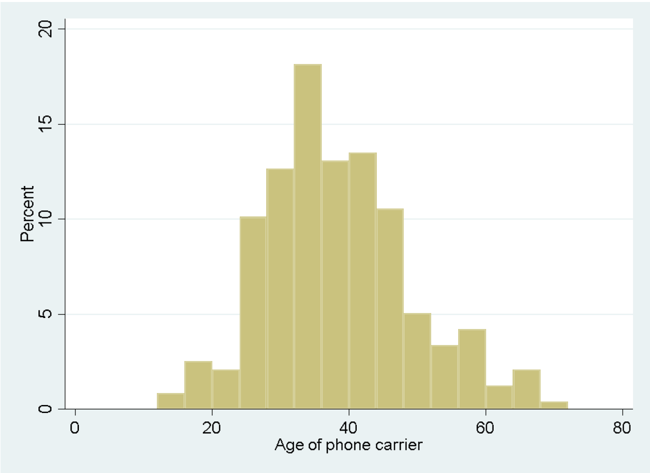

Figure A 2. ESM Survey diagnostics

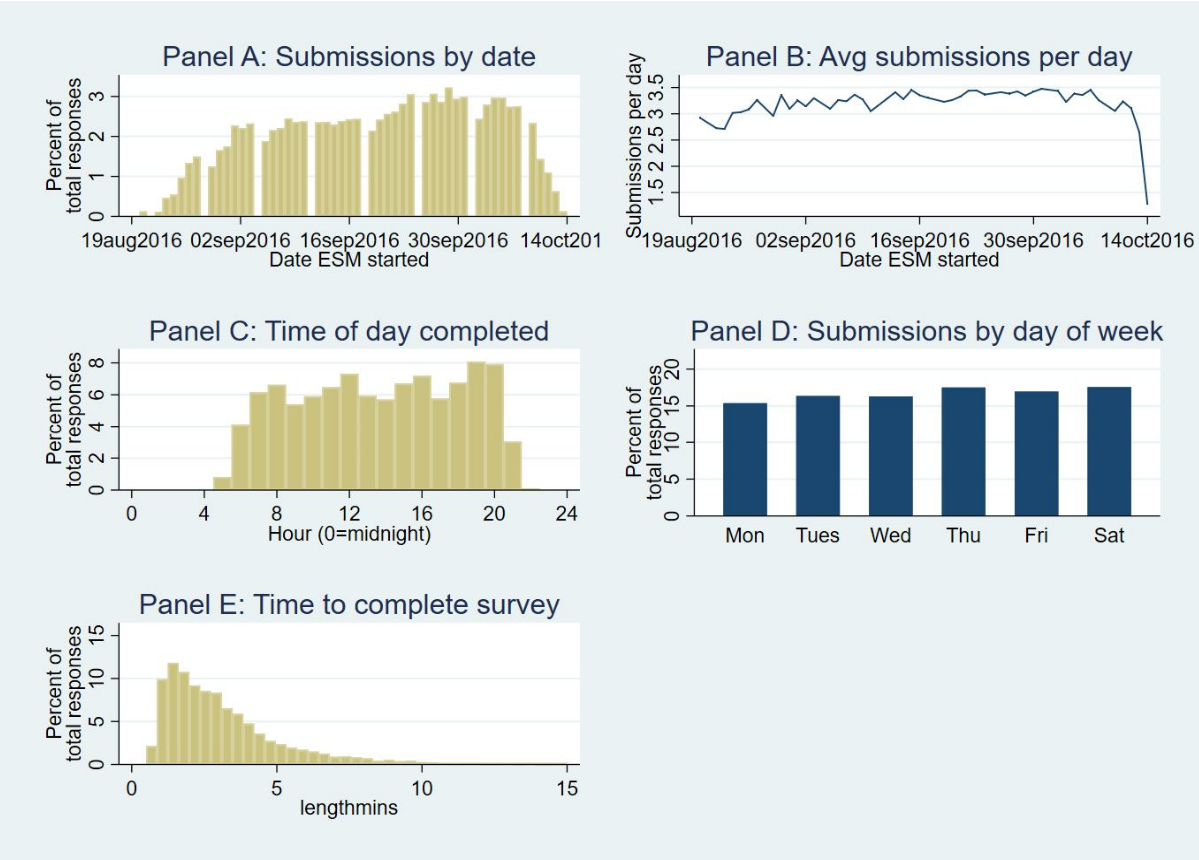

Table A 1. Balance table: characteristics of main water collector

Variable	Control	Treatment	Difference
Female	.94	.91	-.03
	(.23)	(.28)	(.034)
Age	38	38	-.8
	(11)	(11)	(1.4)
Read with difficulty	.18	.15	-.033
	(.39)	(.36)	(.048)
Illiterate	.11	.078	-.028
	(.31)	(.27)	(.038)
Years of schooling	6.9	7.2	.26
	(3.5)	(2.8)	(.42)
Work on household's farm	.99	.97	-.027
	(.091)	(.18)	(.019)
Work for wages past 2 weeks	.37	.43	.063
	(.48)	(.5)	(.064)
Work in own private business	.11	.096	-.011
	(.31)	(.3)	(.039)
Days working for wages, past 2 weeks	4.8	4.5	-.29
	(3.2)	(3.4)	(.69)
Average daily wage (ksh), past 2 weeks	220	256	36
	(103)	(128)	(24)
Total mins spent collecting water yesterday	170	176	5.5
	(107)	(108)	(15)
Subset of collection time spent waiting (mins)	54	57	3.3
	(40)	(53)	(6.7)
Liters of water collected yesterday	86	94	8.3
	(48)	(92)	(10)
Observations	122	115	237

Table A 2. Balance table: household-level variables

Variable	(1) Control	(2) Treatment	(3) Difference
Reported monthly income (Ksh)	8373	9160	787
	(4223)	(5508)	(646)
Earned wages past 2 weeks, all members	1638	1919	281
	(6090)	(5882)	(779)
Wealth index (PCA)	-.17	.18	.36
	(1.6)	(1.9)	(.24)
Dirt floor	.85	.71	-.14
	(.36)	(.45)	(.053)***
Num. buildings	2.7	2.9	.25
	(.94)	(1.1)	(.13)*
Acres of land owned	.91	1.8	.87
	(1.1)	(1.8)	(.2)***
Has Electricity connection	.11	.061	-.046
	(.31)	(.24)	(.036)
Total water storage on compound (Liters)	312	378	66
	(382)	(571)	(63)
Percent who treat water before drinking	.48	.32	-.15
	(.5)	(.47)	(.063)**
Liters delivered by vendors in past 7 days	208	220	11
	(158)	(198)	(43)
Observations	122	115	237

Table A 3. Time Use Categories and Descriptions

Label (English)	**Label (Kimeru)**	**Description (English)**	**Description (Kimeru)**
Meetings	micemanio	Time spent attending community meetings, going to church, or attending funerals	Igiita riria utumaira gwita micemanione ya ntura, gwita kanicene kana gwita mathikone
Collecting firewood	Kuuna nku	Time spent collecting firewood. This includes the time spent walking from the house to the area where you collect and back.	Igiita riria utumaira kuuna nku, ugitaranagia, kagita karia utumaira kuuma njaa gweta naaria uunaa nku na gucoka njaa
Sleeping	Kumama	Time spent sleeping	Kagiita karia utumaira kumama
Other work outside the home	Ngugi ingi ome ya njaa	Time spent on other kinds of work or business outside the home, besides farming or caring for livestock.	Igiita riria utumaira kurita ngugi ingi kana biachara ome ya njaa iti kurima kana kumenyera nyomoo cia njaa
Reading and studying	kuthoma	Time spent reading books, newspapers, magazines or, for children, doing homework	Igiita riria utumaira kuthoma, mauku, gazeti, kana aana kubwithia ngugi iria baei cia cukuru
Games and hobbies	Michetho an matu yaria wendete	Time spent on games and fun activities or hobbies	igiita riria utumaira guchetha michetho na mantu yaria wendete/ yaria yakugwiragia.
Not much of anything	guti uu kuthithagia	Time spent doing not much of anything	Igiita riria utumaira utiu ukuthithia
Bathing	Kuthamba	Time spent bathing and washing your own body. If you wash a source away from the household, include the time spent walking there and back.	Igiita riria utumaria kuthamba, akethirwa uthambaira kuraja na njaa, utaranie kagita karia utumaira gwita na gucoka
School	Cukuru	For children, time spent at school	Kagiita karia aana batumaira cukuru
Caring for children and others	Kumenyera aana na bangi	Time spent caring for children, including breastfeeding, bathing children, dressing children, and helping them with their homework; and time spent caring for elders, the sick and the physically challenged who need your support	Igiita riria utumaira kumenyera twana, ugitaranagia gwonkia kubathambia, kubekira nguo, na kubatetheria kurita gungi cia cukuru, na igiita riria utumagira kumenyera antu bakuuru, aajie, baria bataukiri ni icunci bia mwiri, akiri, na ibakwenda, utenthio bwaku

Radio or TV	Kameme na TV	Time spent watching television or listening to the radio	Kagiita karia utumagira kwona TV kana kuthiukira kameme
Socializing	Kurianira na antu bangi	Time spent socializing and talking with friends and relatives, and time spent eating meals, including breakfast, lunch, and dinner. Do not include time spent preparing food.	Igiita riria utumaira, kurianiria na kwaranina na acore na antu benu, na kagita karia utumaria kuria biakuria, witaranagia biakuria bia rukiiri thaa mugwanja na biogoro, utigutarania kagita karia utumirite kuthuranira biakuria biu
Collecting water	Gutaa ruuji	Time spent collecting water, including the time spent walking from the house to the water source, the time spent waiting to fill the container, and the time spent walking home	Igiita riria utumagira gutaa ruuji, ugitaranagia kagita karia utumagira kuuma njaa gwita naaria utaaga ruuji, na kagita karia wetagira kujuria kiria ugutaa nakio na kagita karia utumagira gucoka njaa.
Private	Mantu ya witho	This category is for time you spend doing something else that is private for you and you do not want to tell us about. That is OK.	Kagiita karia utumaira kubuithia mantu ya witho kana mantu jaria utikwenda kwariria
Market	Thoko	Time spent traveling to the market and back, as well as time spent at the market	Igiita riria utumaira gwita thoko na gucoka, amwa na igiita riria utumaira thokone.
Livestock	Nyomoo cia njaa	Time spent taking animals to graze or drink	Igiita riria utumaira kurithia nyomoo cia njaa kana kunyuithia ruuji
Farming	Urimi	Time spent for farming work including plowing, sowing, weeding, harvesting, or hoeing. This include working on your own farm, or working on someone else's farm, either paid or unpaid.	Igiita riria utumaira urimine, ugitaranagia, gucimba, kuanda, kurimira, kana guketha. Ugitaranagia kurita ngugi muundene jwaku, kurita ngugi muundene jwa munti ungi ukiriawa kana utikuriwa
Household chores	Gwita ngugi cia njaa	Time spent on household chores, like preparing meals or tea, pounding grain or shelling beans, sweeping, washing dishes and utensils, washing clothes and tidying.	Igiita riria utumaira kurita ngugi cia njaa, ta kuthuranira biakuria na kuruga, gutira into ja mpempe, kana kuura mungau, kwegera, kuthambia into, kuura nguo na kutheria

Appendix B: Sensitivity analysis using different definitions of treatment

As discussed in the main text, the research team struggled to get water consistently delivered to households over the first two weeks of treatment. Our preferred specification in the paper uses detailed information collected by field staff about which locations received vended water as planned and which did not. We drop observations from treated households on days when they should have received water but did not, and keep observations from those first two weeks for treated households who did receive vended water.

In this appendix, we explore two alternative definitions of treatment. In the "**conservative**" definition, we drop all records from treated households in those first two weeks. Because most of the problems happened in the first two days of water vending as logistical issues were worked out, a second "**simple**" definition drops all observations from all treated households only for those first two days.

Table A 4. Sensitivity analysis: impact of treatment on **affect** under three definitions of treatment (OLS)

Treatment definition	**Happy**	**Sociable**	**Energetic**	**Safe**	**Some pain**	**Wish Do Else**
Main (results in main body of paper)	2.65** (1.07)	1.61 (1.60)	5.63*** (1.67)	4.44*** (1.29)	-0.079*** (0.021)	-0.075*** (0.020)
Conservative	3.41** (1.41)	1.86 (1.75)	5.97*** (1.82)	4.51*** (1.55)	-0.087*** (0.028)	-0.089*** (0.023)
Simple	2.89** (1.12)	1.49 (1.86)	5.81*** (1.85)	5.06*** (1.54)	-0.075*** (0.024)	-0.094*** (0.021)
Mean in control group	79.1	63.1	34.5	70.7	0.24	0.17

Notes: Each cell represents a separate fixed-effects linear panel data model (Stata: xtreg, fe) regression with the dependent variable of the affect measure. Robust standard errors (shown in parentheses) clustered at the respondent level in parentheses. All models include controls for whether the respondent was alone and time of day (morning, afternoon, evening). All regressions based on responses from 195 water carriers. Because of item non-response (e.g. a respondent answered the question about happiness but not about energy), the exact number of ESM records vary by regression. For the main definition of treatment, they range from 8,425 to 8,466. For the "simple" treatment definition, they vary from 8,435 to 8,476. For the "conservative" definition they range from 6,059 to 6,084.

Table A 5. Sensitivity analysis: Impact of treatment on **time use** under three definitions of treatment

	Main treatment definition (results in main paper)	**Conservative treatment definition**	**Simple treatment definition**
Water collection	-0.104***(0.013)	-0.122***(0.015)	-0.123***(0.014)
Household work	0.0508***(0.017)	0.0537***(0.019)	0.0453**(0.018)
Leisure	0.0203**(0.0080)	0.0298***(0.0098)	0.0175**(0.0087)
Social	0.00255(0.0089)	0.00633(0.011)	0.0118 (0.0099)
Own farm	0.0252*(0.013)	0.0321**(0.014)	0.0326***(0.012)
Meetings	-0.0068(0.0083)	0.00278(0.011)	-0.0136(0.0090)
Caring for others	0.00553(0.0080)	-0.000361(0.0087)	0.00681(0.0084)
Going to market	0.0035 (0.0068)	-0.000781(0.0088)	0.00492(0.0069)
Paid work	0.00148(0.0076)	-0.00526(0.0070)	-0.000928(0.0068)
Livestock	0.0112 (0.0075)	0.00886(0.0096)	0.0177*(0.0093)
Washing, bathing	-0.00326(0.0060)	-0.00128(0.0075)	-0.000271(0.0066)
Wood	0.00268(0.0054)	-0.00206(0.0057)	0.00139(0.0058)
Reading/studying, school	-0.00529(0.0043)	-0.00438(0.0060)	-0.0000616(0.0045)
Private	-0.00218(0.0048)	0.004 (0.0059)	0.00236(0.0048)
Informal, unpaid work	-0.00185(0.0032)	-0.00147(0.0037)	-0.00306(0.0034)

Notes: All results based on ESM records from 195 groups. For "main" treatment, n=8,561 ESM records. For conservative treatment, n=6,167 records, and for "simple" treatment n=8,571 records. Each row represents a separate fixed-effects linear panel data model (Stata: xtreg, fe) regression with the dependent variable of the time use category. Coefficient reported is for the treatment dummy variable. Robust standard errors shown in parentheses, clustered at the respondent (phone). "Leisure" includes sleeping, playing sports/hobbies, listening to TV/radio, socializing or doing "not much of anything". "Paid work" includes paid work in the formal or informal sector, including paid work on someone else's farm, and time spent in one's own private business. For the main definition of treatment, they range from 8,425 to 8,571. For the "simple" treatment definition, they vary from 8,435 to 8,571. For the "conservative" definition they range from 6,059 to 6,167* = 90%, **95%, ***99%.

Table A 6. Impact of treatment on school attendance and time spent studying by school-aged children under three treatment definitions

PANEL A: Main treatment definition				
	(A) School-reported attendance	**(B)** School-reported attendance	**(C)** Self-reported school attendance	**(D)** Self-reported school attendance
Treatment	0.036 *** (0.013)	0.065*** (0.025)	0.141*** (0.027)	0.125*** (0.035)
Trt*Girl		-0.009 (0.032)		0.042 (0.038)
Trt*Age10-15		-0.061** (0.027)		-0.012 (0.038)
Trt*Age16-19		0.070 (0.059)		0.0247 (0.083)
N (groups)	6,483 (241)	6,483 (241)	3,831 (382)	3,831 (382)
PANEL B: Conservative treatment definition				
Treatment	0.048 (0.017)***	0.089*** (0.317)	0.228*** (0.038)	0.203*** (0.046)
Trt*Girl		-0.17 (0.042)		0.0669 (0.045)
Trt*Age10-15		-0.072** (0.034)		-0.00998 (0.050)
Trt*Age16-19		0.045 (0.070)		0.0233 (0.087)
N (groups)	4,593 (241)	4,593 (241)	2,671 (372)	2,671 (372)
PANEL C: Simple treatment definition				
Treatment	0.044*** (0.0144)	0.076*** (0.0281)	0.194*** (0.034)	0.168*** (0.042)
Trt*Girl		-0.013 (3.54)		0.0916** (0.042)
Trt*Age10-15		-0.053* (0.0295)		-0.0254 (0.047)
Trt*Age16-19		0.009 (0.059)		-0.0139 (0.088)
N (groups)	6,483 (241)		3,836 (382)	3,836 (382)
Mean in control group	0.92	0.92	0.82	55 minutes

Notes: Columns A and B use school-reported attendance; columns C and D use data as self-reported by households. Robust standard errors clustered at the respondent/household level for all models. * = 90%, **95%, ***99%.

Table A 7. Impact of treatment on parent-reported time spent studying by school-aged children under three definitions of treatment (household-fixed effects OLS)

PANEL A: Main treatment definition (results in paper)		
	(A) Minutes spent studying	**(B)** Minutes spent studying
Treatment	7.85**(3.07)	9.73**(4.84)
Trt*Girl		-5.07 (4.65)
Trt*Age10-15		-2.35 (4.57)
Trt*Age16-19		11.47 (10.45)
N (groups)	3,707 (381)	3,707 (381)
PANEL B: Conservative treatment definition		
Treatment	4.121 (3.85)	3.818 (6.58)
Trt*Girl		-7.448 (6.17)
Trt*Age10-15		1.647 (6.48)
Trt*Age16-19		18.86 (13.5)
N (groups)	2,565 (368)	2,565 (368)
PANEL C: **Simple treatment definition**		
Treatment	5.955* (3.60)	11.34* (6.34)
Trt*Girl		-9.305*(5.24)
Trt*Age10-15		-3.739 (6.26)
Trt*Age16-19		2.324 (10.2)
N (groups)	3,710 (384)	3,710 (384)
Mean in control group	55 minutes	55 minutes

Notes: Minutes studying as self-reported by parents on a mobile phone survey conducted at the end of each school day. Robust standard errors clustered at the respondent/household level for all models. * = 90%, **95%, ***99%.

Table A 8. Impact of treatment on chores performed by school-aged children

	Water collection	No Chores	Wood collection	Cleaning	Cooking	Farm/ Livestock
Treatment main	-0.256***	0.0425	0.029	0.087***	0.044*	0.011
	(0.032)	(0.030)	(0.038)	(0.021)	(0.023)	(0.017)
Conservative	-0.299***	0.0603	0.0565	0.101***	0.0641***	0.0287
	(0.037)	(0.041)	(0.052)	(0.028)	(0.022)	(0.024)
Simple	-0.264***	0.0394	0.048	0.0952***	0.0438**	0.011
	(0.033)	(0.031)	(0.035)	(0.021)	(0.022)	(0.020)

Notes: Main regressions based on 3,891 survey responses for 384 school-aged children. Conservative treatment definition uses 2,719 ESM responses from 375 children. Simple definition uses 3,897 records for 384 children. Robust standard errors are clustered at the household level. * = 90%, **95%, ***99%.

Appendix C: Additional details on cleaning ESM records and sensitivity analysis

In this section we detail the steps taken to clean the ESM records. We faced several challenges and we provide this detail both for research transparency and to help guide future researchers who might be interested in implementing ESM using ODK. We begin by describing the challenges and then detail the procedure for dropping or flagging records.

First, there were a small number (29) ESM records in our data that could not be matched back to a specific phone or household using the phone's unique deviceid or SIM card serial number, which was stored in the ODK record.

Second, enumerators had concerns about the quality of data from 8 households, for example because of eyesight problems or mental health concerns.

Third, the research team themselves generated ESM records when we installed the ODK app on new phones for the first time, generating a test record to be sure the app was working. We similarly generated test ESM records when phones malfunctioned and needed to be returned to the team to have software re-installed. We also had enumerators run through the app as they explained the ESM program to respondents, again generating "false" ESM data. Unfortunately, we did not *ex ante* design a simple way to distinguish these test records from "real" ESM records in the program.

Fourth, in some cases the server assigned incorrect dates or times to database records for reasons we do not understand. For example, one record was time-stamped as beginning in October 2037 and ending in September 2016.

Fifth, as discussed in the main text, we drop ESM records where we had reason to believe another household member had taken the phone and was filling out the forms. This includes records with an incorrect "secret animal", which we use as a password.

We drop records from our analysis for the five reasons described above.

A sixth problem is that there are cases where the timing and length of ESM records suggest that the survey was not completed in the manner expected. Rather than drop these records, we flag them as potentially problematic. Ideally, the data record should show that an ESM took about 3-5 minutes, and then 3-4 hours (or overnight) before they are prompted to complete the next ESM survey. There are three behaviors that would complicate this picture.

A) the subject completes the survey but thinks it did not go through, and quickly completes another one. Start times will not be exactly the same, but very close together. Nearly all answers will be the same. Length of both entries will be short. One should keep the first record, though it should not matter since the records have the same main fields.

B) The subject starts an ESM, but then leaves the app to take a call or do something else. The entry is left uncompleted. The Randomly-Remind Me prompt goes off hours later to remind them to do the next entry, and they see they have the older entry uncompleted. They complete it, then start the second one. The second one will appear to have been completed just after the last one has been completed. But the first entry will have a very long completion time. We cannot know how much of the first entry was completed at the correct time and how much was completed later, based on recall. So, the first entry should likely be deleted, but the second entry is likely OK.

C) The subject ignores the RRM requests and completes many ESM records at once, analogous to the stories from Cziksentmihalyi's earlier pager and logbook problem. This would be characterized by several ESM records in quick succession.

The table below details our data cleaning procedures. Although our anonymized data is available in the SND depository, the Stata code to implement these procedures runs on the de-anonymized data. The code itself is, however, available on request.

Description	**Records dropped**	**Total ESM records**
1. Initial total records	--	23,943
2. Drop all records from 8 households (3 control, 5 treatment) where enumerators had concerns about the quality of data from household	146	23,797
3. Drop record from Feb 2016 during testing in the U.S.	1	23,796
4. Drop records with implausible dates for both the "start" and "end" timestamp.	639	23,157
5. Drop records likely to be during programming or training		
a. Timestamp between 11pm and 5am. Phones were programmed at night, and users would never have been prompted to complete at those times.	430	22,727
b. Start date is on or before the date the phone was deployed to the household	794	21,933
c. Form started or completed on a Sunday. Users would not have been prompted to complete, and research team used Sundays to program phones.	635	21,298
6. Drop if record submitted after study ended in October 2016	72	21,226
7. Drop records with incorrect passwords		
a. Drop *all* records from 12 phones where the password was incorrect in >90 % of records.	1,264	19,962
b. Drop all remaining records with incorrect password	465	19,497

8. Drop records with the exact same start time (to the second) on the same phone (server problem)	4	19,493
9. **Flag, but not drop**, records that have long completion times, were completed in close proximity to other records, or where more than four per day were completed.		
a. Flag records with a length to complete > 15 minutes	Flag 3,548	
b. Also flag records that began within 60 minutes of the last ESM completed	Addt'l 3,208	
c. Also flag records from days where more than four ESMs were recorded.	Addt'l 2,773	
d. Also lag all records from 27 phones where more than 75% of its ESM records are flagged in a), b) or c) above.	Addt'l 405	
TOTAL RECORDS UNFLAGGED (main dataset) TOTAL RECORDS FLAGGED (adds for sensitivity sample) TOTAL RECORDS for "full" dataset	**N=9,959** 9,934 **N=19,493**	

Sensitivity analysis

Table A 9. Sensitivity analysis: impact of receiving vended water on **affect** keeping flagged ESM observations (OLS)

	Happy	Sociable	Energetic	Safe	Some pain	Wish Do Else
Treatment (results in main paper)	2.65** (1.07)	1.61 (1.60)	5.63*** (1.67)	4.44*** (1.29)	-0.079*** (0.021)	-0.075*** (0.020)
Treatment ("full" dataset, keeping flagged records)	3.09*** (0.94)	2.23* (1.35)	3.72** (1.47)	5.08*** (1.22)	-0.086*** (0.016)	-0.073*** (0.016)
N=	17,332	17,271	17,279	17,380	17,408	17,277
Groups =	222	222	222	222	222	222
R^2	0.004	0.036	0.026	0.0063	0.0078	0.0082
Mean in control group	78.1	64.1	32.5	69.8	0.27	0.16

Notes: Robust standard errors clustered at the respondent level. Controls for the time of day (morning, afternoon or evening) when the ESM record was completed and whether the respondent was alone are included but not reported. Uses main (preferred) definition of treatment status (see main text). Results without dropping control households are based on roughly n=8,450 ESM records from n=195 respondents (see main text)

Table A 10. Sensitivity analysis: impact of receiving vended water on **time use** keeping flagged ESM observations (OLS)

	Results in main paper	**“Full” ESM data, keeping flagged records**
Water collection	-0.104***(0.013)	-0.102***(0.010)
Household work	0.0508***(0.017)	0.0417*** (0.014)
Leisure	0.0203**(0.0080)	0.014 (0.0086)
Social	0.00255(0.0089)	0.0146* (0.0083)
Own farm	0.0252*(0.013)	0.0195** (0.0094)
Meetings	-0.0068(0.0083)	-0.0134* (0.0068)
Caring for others	0.00553(0.0080)	0.00991 (0.0072)
Going to market	0.0035 (0.0068)	0.00144 (0.0049)
Paid work	0.00148(0.0076)	0.00457 (0.010)
Livestock	0.0112 (0.0075)	0.00987* (0.0058)
Washing, bathing	-0.00326(0.0060)	0.00839* (0.0046)
Wood	0.00268(0.0054)	-0.00347 (0.0035)
Reading/studying, school	-0.00529(0.0043)	-0.00266 (0.0034)
Private	-0.00218(0.0048)	0.000807 (0.0031)
Informal, unpaid work	-0.00185(0.0032)	-0.00185 (0.0028)

Notes: Each row represents a separate fixed-effects linear panel data model (Stata: xtreg, fe) regression with the dependent variable of the time use category. Main results based on n=8,561 ESM records from n=195 subjects. “Full” results based on n=17,740 records from n=222 subjects. Coefficient reported is for the treatment dummy variable. Robust standard errors shown in parentheses, clustered at the respondent (phone). “Leisure” includes sleeping, playing sports/hobbies, listening to TV/radio, socializing or doing “not much of anything”. “Paid work” includes paid work in the formal or informal sector, including paid work on someone else’s farm, and time spent in one’s own private business. Uses “main” definition of treatment (see text) * = 90%, **95%, ***99%.

Appendix D: Additional details on sampling

We targeted a total sample of 250 households in the four sub-locations. We began by targeting households which had participated in 2013 survey (chosen by a random transect approach in 2013) and met the following criteria:

- The main source of water for the household was not a source in the compound (i.e. private well) nor primarily vended water.
- Households should have school aged children (6-18 years) who do not attend a boarding school i.e. they go home every day after school and must be within the area of study.

Of the 387 households interviewed in 2013, 180 household met these criteria. All these households were visited. If a household was not at home during the first visit, the enumerators asked around about the household and tried a second time to contact the household. In total, 85 were recruited in the study from the 2013 sample frame, or 47% of the total 180.

This prompted recruitment of new households to reach the target sample size. The replacement was to be done within the same four sub-locations with the same recruitment criteria above. We constructed a census by asking the area sub-chiefs and village elders to list all the households within their sub-locations which met the criteria and were not among the households interviewed in 2013. This census included 218 households in the Nairiri location, 78 in the Machako location, 194 in Mutionjuri and 20 in Kianjai.

A systematic sampling was applied to the listed households where every second household in the list was recruited. Households would sometimes fall within the same compound (i.e. a shared common gate). In these cases, we randomly chose one of the households to be interviewed through choosing straws.

In total, we contacted 264 households who had not been interviewed in 2013. Of these, 187 households (71%) were recruited, and roughly in proportion to the census lists drawn up by the local leaders (see Table 1). Of the 77 households who were contacted but not recruited:

- Eleven respondents (4% of the total 264) were dropped because the main water collector was physically challenged (i.e. hearing and sight problems) in a way that would have made it impossible to complete the time use surveys, had insurmountable difficulties using a smart phone; had substance abuse problems; or did not, in fact, have school-aged children at home.
- 48 households (18%) refused to be interviewed, either because the household head would not allow the spouse or main water collector to participate, the parents of the main water collector refused to allow the child to be involved in the study, or because of other reasons. A significant fraction of refusals were due to rumors that the researchers were members of the Illuminati or were devil-worshipers.
- 15 households (6%) could not be contacted, though only after one attempt.

Sample frame

Sublocation	No. of households listed	No. of households recruited
Nairiri	218	95 (44%)
Machako	78	25 (32%)
Mutionjuri	194	46 (24%)
Kianjai	20	7 (35%)
Total	510	173