AIDDATA
A Research Lab at William & Mary

WORKING PAPER 96

May 2020

A micro-based approach to evaluate the effect of water supply on health in Uganda

Raymond Boadi Frempong
University of Bayreuth

Lucas Kitzmüller
Harvard Kennedy School

David Stadelmann
University of Bayreuth and CREMA

Abstract

This paper investigates the impact of improved water provision on individual health outcomes in rural Uganda. We merge household and individual panel datasets with sub-county level administrative data on water supply projects. Our approach allows us to estimate fixed-effect panel data models which use temporal and spatial variation at the sub-county level as identifying variation. We find evidence of small effects from more installations of improved water supply on its water usage, health outcomes of household members, and water collection times. Increasing the sub-county rate of improved water sources per capita leads to a reduction in the likelihood of individuals suffering from symptoms of illness associated with inadequate water supply. We argue that our micro-based approach provides a more externally-valid and highly cost-effective means of evaluating scalable development projects.

JEL codes: *O12; I15; I31*

Keywords: *Evaluation; development projects; drinking water; sanitation; child mortality*

Author Information

Raymond Boadi Frempong
Cluster of Excellence "Africa Multiple", University of Bayreuth, Germany
raymond.frempong@uni-bayreuth.de

Lucas Kitzmüller
Harvard Kennedy School, United States.
lucaskitzmueller@hks.harvard.edu

David Stadelmann
Cluster of Excellence "Africa Multiple", University of Bayreuth, Germany.
CREMA – Center for Research in Economics, Management, and the Arts, Switzerland
david.stadelmann@uni-bayreuth.de

Acknowledgements

The authors would like to thank Gustavo Torrens for excellent feedback and inspiring discussions. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2052/1 – 390713894.

I. INTRODUCTION

In 2015, an estimated 663 million people worldwide lacked access to safe drinking water. The situation is especially severe in sub-Saharan Africa, with people living in this region accounting for almost half of the world's population who do not have access to an improved[1] water source (UNICEF\WHO, 2015b). Achieving "universal and equitable access to safe and affordable drinking water for all by 2030" is among the United Nation's Sustainable Development Goals. Governments and international donors support campaigns and water infrastructure projects to improve water supply. Such projects offer potential welfare improvements, including the reduction of water-borne diseases and the reduction in water collection times by women or children.

Water supply interventions have been extensively studied (see Wolf et al. (2014) for a review). Apart from methodological and implementation challenges, the cost of randomisation has been a significant hindrance in conducting large-scale impact evaluation studies of water supply infrastructure projects. With the notable exceptions of Opryszko et al. (2010) and Devoto, Duflo, Dupas, Parienté, and Pons (2012), studies on water supply interventions use data which result from contentious choices regarding control groups, use small sample sizes, or are limited in geographical and temporal scope. These challenges may affect the external validity of the literature's findings.

We suggest a micro-based approach to evaluate the impact of water supply interventions on the use of improved water and the health outcomes of such use. We match an existing water supply programme data with geo-coded household panel surveys. This approach allows us to account for fixed-effects in our analyses so that we can exploit the spatial and temporal variation in programme intensity to control for unobserved differences across households and geographical regions. In total, we investigate 88,728 water-related projects in Uganda across 256 sub-counties, with up to 4,748 observations at the household level, for the years 2009 to 2011.

There are at least five main advantages to our micro-based approach with the individual fixed-effects that may also apply for settings outside of Uganda: (1) it allows us to study the impact of water supply investments on a large, often nationally-representative sample of households over a longer period, rather than limiting the analysis to a single cross-section; (2) the panel structure allows controlling for *unobservable* time-invariant factors; (3) given the richness of most household surveys, it is possible to control for many confounding time-variant factors; (4) it allows for reasonable identification in cases where randomised control trials are not feasible or are too costly; (5) due to their low frequency and the variety of topics addressed in household surveys, estimates of treatment effects are less likely to suffer from a *Hawthorne effect* (i.e., a bias that occurs if the act of being surveyed itself causes behavioural change (Zwane et al., 2011)).

Importantly, using our methodology is much less costly compared to a randomised study of a similar sample size because we can rely on existing datasets which can be

---

[1] Improved drinking water sources are defined as those that "by nature of their construction, are likely to be protected from outside contamination, in particular from contamination with faecal matter" (WHO/UNICEF Joint Monitoring Programme, 2015b). Examples are deep boreholes, public taps, protected springs, and shallow wells.

merged on geo-codes. Moreover, such studies can be conducted within a larger geographical area and across urban and rural areas, as is the case in this study. Our evaluation approach is informed by Rosenzweig and Wolpin (1986), Pitt, Rosenzweig, and Gibbons (1993), Gertler and Molyneaux (1994) and, more recently, Zhang (2012).[2]

Aside from the methodological innovations, we provide evidence on the importance of proximity to improved water sources, in addition to the common socio-cultural factors in the literature (UNESCO, 2002; Wasonga, Okowa, and Kioli, 2016). Once controlling for household (or individual) fixed-effect, it is not, *apriori*, clear that the provision of water sources in the proximity of the household would automatically increase improved water usage, decrease disease incidence, and reduce the amount of time spent on fetching water. Employing our setting, we find some evidence that an improved water supply increases the likelihood of improved water usage, in that it may reduce the presence of water-related disease symptoms and may reduce water collection times.

The remainder of this paper is structured as follows: Section II provides a review of the literature; we discuss our data methodology in Section III; Section IV provides results for the impact of improved water provision on usage, health outcomes and water collection times; and Section V offers concluding remarks.

II. RELATED LITERATURE

Following the early works of Esrey, Potash, Roberts, and Shiff (1991), Esrey, Feachem, and Hughes (1985), and Blum and Feachem (1983), researchers have intensively studied the health impact of water and sanitation interventions (e.g., Aiello, Coulborn, Perez, & Larson, 2008; Amrose, Burt, & Ray, 2015; Cairncross et al., 2010; Clasen, Schmidt, Rabie, Roberts, & Cairncross, 2007; Ejemot-Nwadiaro, Ehiri, Arikpo, Meremikwu, and Critchley, 2015; Loevinsohn et al., 2015; Norman, Pedley, and Takkouche, 2010; Taylor, Kahawita, Cairncross, and Ensink, 2015). Published studies tend to find relevant health effects of water projects in low- and middle-income countries.

Many researchers have pointed out the methodological challenges that are associated with the approaches employed in numerous studies in this field (e.g., Blum & Feachem, 1983; Dupas & Miguel, 2016; Fan & Mahal, 2011; Strunz et al., 2014). Often, the use of cross-sectional data does not clearly identify the effects of water projects, and relevant control groups are hard to find (e.g., Majuru, Michael Mokoena, Jagals, & Hunter, 2011; Nanan, White, Azam, Afsar, and Hozhabri, 2003; Tonglet, Isu, Mpese, Dramaix, and Hennart, 1992). Moreover, studies may offer limited scope due to their sample sizes, as large samples are expensive to gather (e.g., Aziz et al., 1990; Blum et al., 1990; Gasana, Morin, Ndikuyeze, & Kamoso, 2002; Hasan et al., 1989; Huttly et al., 1990; Madrigal, Alpízar, & Schlüter, 2011; Majuru et al., 2011; Ryder et al., 1985; Tonglet et al., 1992; Zeng-Sui et al., 1990). Similarly, geographical coverage may be limited, which can lead to concerns regarding external validity.

---

[2] We relate to the literature using difference-in-difference estimators to study the impact of improved water provision (e.g. Giliani, Gonzalez-Rozada, and Shargrodsky (2009) and Gross, Günther, and Schipper (2018)).

Observational and randomised evaluations may be associated with the *Hawthorne effect* (i.e., study participants may alter their behaviour due to their awareness of being observed). Indeed, Zwane et al. (2011) show that Kenyan households that are frequently surveyed about their children's health status behave differently.[3] Frequent surveys affect the estimated treatment effect of ongoing water quality interventions (Kremer, Leino, Miguel, & Zwane, 2011). The *Hawthorne effect* has important implications because frequent data collection is typical in epidemiology and, consequently, the effect may bias evaluations of water supply intervention. A solution is to collect more data, but passively, through administrate data, 'big data' approaches, or 'pure control' groups (see Dupas and Miguel, (2016)). However, the latter, to our knowledge, has not been applied in the literature on water supply interventions in Sub-Sahara Africa.

The literature has also applied econometric techniques such as propensity score matching (PSM) or similar methods on cross-sectional, nationally-representative household survey data (Begum, Ahmed, & Sen, 2011; Capuno, Tan, & Fabella, 2015; Fan & Mahal, 2011; Fink, Günther, & Hill, 2011; Jalan & Ravallion, 2003; Khanna, 2008; Roushdy, Sieverding, & Radwan, 2013). Such approaches have the advantages of using larger sample sizes, being less likely to suffer from bias from the *Hawthorne effect,* and employing noteworthy identification strategies.[4] PSM is most useful when only observed characteristics are believed to affect treatment assignment and outcomes. This is unlikely to be the case with water supply projects. As Pitt et al. (1993) argue, coverage and timing of water-related interventions may be responsive to unobserved characteristics of the targeted populations, which may influence outcomes of interest and introduce bias. For example, to maximise their impact, water supply interventions may be targeted at areas with low coverage of basic health care while the local availability of health care is likely to also affect households' health outcomes. PSM methods that do not account for the location-specific provision of health care will lead to biased estimates of the health effects of water supply interventions.

An alternative is to employ fixed-effects methods to study programmes by governments in developing countries. Rosenzweig and Wolpin (1986), Pitt et al. (1993), and Gertler and Molyneaux (1994) have employed this technique in different studies including planning, health, and schooling programmes in the Philippines and Indonesia.[5] Duflo (2001) used a similar approach to evaluate schooling and labour market consequences of school construction. We contribute to this literature by merging available geo-coded household panel data with geo-coded data on water supply projects in Uganda. Few studies are closely related to what we do in this paper. Zhang (2012) reconstructs programme data from a household panel survey to evaluate the impact of introducing village-level access to water plants on health in rural China. Using a difference-in-difference while controlling for the province- and year-fixed effects, Dolan et al. (2019) show that the long-lasting insecticide

---

[3] The authors identify a *Hawthorne effect* by showing that more frequent health surveys lead to statistically significant effects on the usage of chlorine disinfectant, reported child diarrhoea, and measured water quality.

[4] In the case of PSM, the key assumption is *conditional independence* (i.e., given a set of observable covariates that are not affected by treatment, potential outcomes are independent of treatment assignment (Gertler, Martinez, Premand, Rawlings, and Vermeersch (2016)).

[5] Pitt et al. (1993) describe the comparative advantage of using panel instead of cross-sectional data in detail.

nets distribution programme leads a decline in under-5 mortality risk among children in the Democratic Republic of Congo.[6]

III. DATA AND IDENTIFICATION STRATEGY

Data and Merging of Geo-Coded Households to Water Projects

Following the reform of the Ugandan water sector in the mid-1990s, the government intends to provide access to safe water for 80 percent of the population by 2020 (Ministry of Water, Lands and Environment, 1999; National Planning Authority, 2015). From 1990 to 2015, access to improved water increased from 39 to 75 percent in rural areas (UNICEF\WHO, 2015a).

We draw on panel data from two sources and merge them based on geo-codes of households: (1) the Uganda Water Supply Atlas (UWSA) for data on improved water sources; and (2) the Uganda National Panel Survey (UNPS) for data on household and individual-level outcomes.

The UWSA is published by the Ugandan Ministry of Water and Environment through its Directorate of Water Development, and gives an overview of the water supply situation in Uganda. The online version of the Atlas contains data on location, type, year of construction, functionality, etc. of all improved water sources in Uganda according to the Ministry (i.e., including sources developed by NGOs).7 The UWSA is based on nationwide data collections carried out in 2009/10 and 2015, which involved physical visits to all improved water sources. The core water source data is available via an OGC Web Mapping Service (WMS) and Web Feature Service (WFS). We directly accessed the location and the type of improved water from the website. However, data on the year of construction was not available, and we generated this variable from the various annual sub-county level reports for the financial years 2009/10 to 2015/16 that present aggregated data from the UWSA. For each year and sub-county, we obtained data on the number and type of improved water sources, their functionality, and population figures. These reports are accessible via an internal interface of the UWSA website. Since missing data was not coded consistently, we assumed missing data for rural sub-counties which consistently report zero functional or non-functional improved water sources (2.5 percent of all rural sub-county annual observations).

We obtain household and individual-level information from the Uganda National Panel Survey (UNPS). This is a multi-purpose data set which contains detailed information on households and their members. We use three rounds of data (from 2009 to 2011) which are available at the World Bank Living Standards Measurements Survey's website. Aside from the socio-economic variables, the data contain geographic information and geo-coordinates of the sampled households. This enabled us to match the households to the water supply

---

[6] Combining data from the Demographic Health Survey from several countries and geocoded aid datasets on World Bank and Chinese development finance, Cruzatti, Drehery, and Matzat (2019) show that Chinese aid and concessional World Bank aids increase infant mortality.

[7] The data set has been used, among others, by Foster (2013) in a quantitative study to investigate predictors of the functionality of community-managed hand pumps.

data based on the sub-county and year of the survey. A typical sub-county has a population size of fewer than 40,000 inhabitants. We merged household for three waves of the UNPS to the water supply data provided by the UWSA. As the same households were observed several times, our approach allows us to account for fixed-effects for households or individuals.

Overall, the various waters sources met about 68 percent of the water needs of Ugandans, and this increased to about 72 percent in the 2011/2012 sample. On average, a household's use of improved water as the main source of water for drinking and domestic activities increased from 58 percent in 2009/2010 to 60 percent in 2011/2012.

Variables and Measurement

In our analysis, we employ three dependent variables.

(1) The first dependent variable reflects a household's actual usage of improved water as the primary source of drinking water (WaterUsage). This variable is measured with a question in the UNPS that asks, "What is the main source of water for drinking for your household?" Applying the UNICEF/WHO definition, we define a household as using an improved water source if its main source of domestic water is from a private connection to a public pipeline, public taps, a borehole, or a protected spring. All other responses are classified as non-improved. Thus, we generate a binary variable whose value is one if the household used any of those sources and zero if otherwise.

(2) To identify the effect of improved water provision on health, we employ a dichotomous variable which indicates whether an individual suffered from illness related to acute or chronic diarrhoea and vomiting (Ill) during the reference period. We select these illnesses because they represent symptoms of the consumption of contaminated water. Our second dependent variable is at the individual level. Its value is one if a member of the household suffered from acute or chronic diarrhoea and vomiting and zero if otherwise.

(3) In keeping with the development literature on water and household time use (Boone, Glick, & Sahn, 2011; Gross et al., 2018), we provide analyses of the effect of improved water provision on time spent on water collection. Physical access has been cited as one of the main problems related to water. As has commonly been done in the existing literature, we use the total time spent on water collection by the household (Time) as a dependent variable. An increase in the installed capacity per capita should ultimately reduce the amount of time a household spends fetching water.

Table 1: Variable description and summary statistics

Variable	2009/2010	2010/2011	2011/2012	2009-2012
Variables in the water usage and collection time models (improved water usage and collection time models)				
Improved water supply capacity per capita, *WaterSupply*	0.68	0.69	0.72	0.69
The Household drinks improved water, *WaterUsage*	0.58	0.59	0.60	0.59
Total collection time (minutes), *Time*	53.97	45.31	39.17	46.3
Number of people in the household	6.66	7.33	7.84	7.27
Ratio of adults to household members	0.46	0.46	0.47	0.47
Age of household head	47.16	47.04	47.48	47.22
The household head has been to school	0.78	0.82	0.79	0.79
The household has a sanitation facility	0.53	0.50	0.50	0.51
Dwelling wall is durable	0.32	0.37	0.39	.36
Log of household expenditure	10.72	10.52	10.55	10.6
Number of rooms in the house	2.98	3.01	3.05	3.01
The community has access to a road	0.99	0.97	0.98	0.98
Variables in the diarrhoea model (diarrhoea **model)**				
The individual suffered from diarrhoea	0.07	0.05	0.04	0.05
Household drinks improved water,	0.57	0.58	0.59	0.58
Age of the individual	24.01	24.12	24.22	24.12
The individual has been to school	0.83	0.84	0.84	0.84
Age of household head	47.96	48.06	48.36	48.13
The household head has been to school	0.80	0.84	0.81	0.82
Number of people in the household	8.17	8.70	9.30	8.72
Log of household expenditure	10.69	10.5	10.5	10.57
The household has a sanitation facility	0.60	0.55	0.56	0.57
Number of rooms in the house	3.38	3.36	3.37	3.37
Dwelling wall is durable	0.35	0.39	0.41	0.38
The type of toilet in the household				
Bush/uncovered	0.19	0.22	0.26	0.22
Pit latrine	0.78	0.75	0.71	0.75
VIP latrine	0.02	0.03	0.02	0.02
Flush	0.01	0.00	0.01	0.01
The community has access to a road	0.99	0.97	0.98	0.98
The community has a health facility	0.94	0.78	0.54	0.76

Notes: Statistics (means) are unweighted. Variables for individuals are aggregated to the household level before being average.

Sources: Uganda National Panel Survey (2009-2012) and the Uganda Water Supply Atlas.

For our main independent variable, we measure the provision of improved water by the improved water supply capacity per capita (WaterSupply). This variable takes into account the fact that different types of water projects may have different capacities. The improved water supply rate was calculated by multiplying each improved water source in the sub-county with the number of people it was capable of serving.[8] We then divide the total number of people that could be served in the sub-county by the population size to obtain the improved water supply rate. It is worth mentioning that the improved water supply rate does not measure actual access to water, but describes the installed improved water supply capacity per capita that takes into consideration the different capacities of the various water projects (i.e., if improved water sources are not close to citizens, we still account for them when calculating the improved water supply rate). Finally, for the classification of improved community water sources, the definition of the UNICEF/WHO Joint Monitoring Programme was applied. At baseline, the average improved water supply rate was 0.58. Yearly changes in the improved water supply rate can be negative if water sources break down or the population grows faster than the number of people that can be served by improved community water sources.

The use of a multi-purpose household dataset such as the Uganda National Panel Survey (UNPS) allows us to account for many control variables in our empirical setting. These include whether the household head has been to school, household size, household expenditure, etc. All variables and descriptive statistics for the years of analysis are presented in Table 1.

Identification Strategies

Our empirical identification strategy is straightforward. The improved water supply capacity per capita (WaterSupply) is measured in the sub-county where the household was located at the time of survey. We observe the households several times in the panel, and household-specific (as well as community-specific) control variables are available. For time-invariant unobservables, we employ household fixed-effects (or individual fixed-effects where appropriate) and account for interview wave fixed-effects. Thus, we estimate the following three specifications for each of the dependent variables, WaterUsage, Ill, and Time:

$$WaterUsage_{ht} = \psi_1 WaterSupply_{ht} + \beta_h + \mathbf{C_{ht}}\,\boldsymbol{\psi}_\mathbf{2} + \phi_t + \epsilon_{ht} \tag{1}$$

$$Ill_{it} = \theta_1 WaterSupply_{it} + \gamma_i + \mathbf{C_{it}}\,\boldsymbol{\theta}_\mathbf{2} + \phi_t + \eta_{it} \tag{2}$$

$$Time_{ht} = \pi_1 WaterSupply_{ht} + \delta_h + \mathbf{C_{ht}}\,\boldsymbol{\pi}_\mathbf{2} + \phi_t + \xi_{ht} \tag{3}$$

WaterUsage is the indicator variable measuring whether the household h uses an improved water source or not. Ill is the indicator variable that indicates whether an individual

---

[8] Appendix A1 provides details of the capacities of the various water sources.

i in the household suffered from water-related illness or not. Time is the total household time spent on water collection. Since the dependent variables of (1) and (2) are dummies, we estimate them using a Logit estimator,9 while equation (3) is estimated with OLS. β_h, γ_i and δ_h represent household fixed-effects (equations 1 and 3) and individual fixed-effects (equation 2). C represents a design matrix which includes household and individual characteristics when appropriate (see Table 1) as well as characteristics of the community where the household is located, such as access to a road and health facility. ϕ accounts for time fixed-effects.

Our main independent variable of interest is WaterSupply. The coefficients ψ_1 , θ_1 and π_1 capture the impact of the improved water supply capacity per capita. As we account for household (or individual) fixed-effects, we can investigate how a change in WaterSupply affects the same household (or the same individual) over time (i.e., whether an increase in the improved water supply capacity per capita leads to increased improved water usage, fewer water-related illness, and lower water collection times). Household (or individual) fixed-effects go a substantial way towards eliminating endogeneity issues in comparison to the existing literature. Households in sub-counties with low levels of improved water coverage may vary from those in sub-counties with high levels of coverage in many ways that also matter for the outcomes of interest. If these differences remain constant throughout the sample period, however, they are controlled for by the inclusion of fixed-effects. Similarly, the ‘bottom-up’ approach to local water development planning may fuel concerns about the endogeneity of infrastructure placement (e.g., some communities develop faster than others and request or receive more improved water sources).10 Again, fixed-effects account for pre-existing infrastructure projects. The UNPS also allows us to account for time-variant characteristics of households so that we can further alleviate omitted variable bias. Finally, our data permits us to account for infrastructure developments over the sample period that are related to changes in roads and health facilities, which are included in the design matrix C.11

Our fixed-effects approach comes with challenges: (1) parameter estimates of treatment effects may still be affected by other time-variant unobserved factors that are correlated with treatment, such as heterogeneity in sub-country economic development; (2) the reliance on household surveys limits our ability to address all research questions of interest; (3) since water programme data are drawn from administrative records, the quality of the final data and sample depend on the completeness of the information provided. Bias from measurement error may be increased by transformations to estimate the model resulting in less efficient estimates (Angrist & Krueger, 1999). Keeping these challenges in mind, we believe that our micro-based approach with fixed-effect offers a worthwhile method for evaluating development programmes ex-post.

---

[9] Linear probability models yield qualitatively similar results.

[10] Local development may also affect health outcomes, and thus, if not controlled for, bias estimates of the effects of improved water sources on health outcomes.

[11] There is the possibility for local politicians to influence the allocation of water and other infrastructural (Quin et al., 2011). Data from the Ugandan Water Supply Atlas provides some evidence that, for the sub-counties during the sample period, allocation regulations were generally followed. The data suggest that rural sub-counties that had a safe water coverage below the national average (and should, therefore, have qualified for more resources) saw an increase in the rate of improved water sources per capita in the following year.

IV. THE IMPACT OF IMPROVED WATER SUPPLY

Effects on the Usage of Improved Water

Table 2 provides results for the impact of improved water supply capacity per capita on a household's actual use of improved water for domestic chores and drinking. Models (1) and (2) provide random effect models and (3) and (4) present conditional household fixed-effects estimates. In model (1), the household's usage of improved water is regressed only on the provision of improved water, without any additional control variables or household fixed-effects. In model (2), other control variables, comprising household and community time-variant characteristics and time fixed-effects, are added. Model (3) estimates the use of improved water as a function of improved water provision and household fixed-effects, while all covariates are added to the household fixed-effects in model (4). Thus, model (4) fully implements our identification strategy and is our preferred specification.

Table 2: Impacts of improved water supply on the usage of improved water

	(1) Logit-RE	(2) Logit-RE	(3) Logit-FE	(4) Logit-FE
Household drinks improved water, WaterUsage	0.158	0.163	0.379***	0.480**
	(0.101)	(0.103)	(0.040)	(0.222)
Household and community time-variant controls	No	Yes	No	Yes
Time fixed-effects	No	Yes	No	Yes
Household fixed-effects	No	No	Yes	Yes
Observations	4748	4748	1008	1008
Wald χ^2	4.038	36.934	3.407	25.081
	[0.044]	[0.000]	[0.065]	[0.014]
Pseudo R^2			0.009	0.056

Notes: Logit models are estimated, and marginal effects are reported. Household and community time-variant controls include all household level controls shown in Table 1. Estimated standard errors in parentheses are clustered at the sub-county level. * p < .1, ** p < .05, *** p < .01. [#] denotes p-value of statistic. Sample differences between Logit-RE and Logit-FE are due to the strict reliance of the conditional fixed effect logit model on variations in the outcome variable. The results remain qualitatively the same when we use the linearity probability model.

All results show a positive effect of improved water supply on the probability of improved water usage. While the results of the random effect models are statistically insignificant at conventional significance levels, models (3) and (4) suggest that an increase in WaterSupply has a statistically significant and positive effect on WaterUsage. Note that the change in significance level and the increase in the estimated marginal effect when moving from random to fixed-effects models suggests that omitted time-invariant household characteristics may have led to a downward bias of the impact of improved water supply on wager usage in specifications (1) and (2). Quantitatively, model (4) indicates that the probability that a household's use of improved water increases by about 45 percentage points for a unit increase in the installed improved water capacity per capita.

Effect on Health

Given that citizens make use of improved water if the supply of it increases, we investigate whether improved water supply affects people's health (equation 2). Table 3 presents the result for the dependent variable Ill, which captures water-related disease symptom of individuals in the household in the last thirty days before the survey (e.g., diarrhoea). Thus, our estimation is performed for individuals and we use individual fixed-effects (instead of household fixed-effects).

Table 3: Impacts of improved water supply on water-related disease symptoms (diarrhoea)

	(1) Logit-RE	(2) Logit-RE	(3) Logit-FE	(4) Logit-FE
Improved water supply capacity per capita	0.001	0.003	-0.315**	-0.071
	(0.005)	(0.005)	(0.123)	(0.125)
Individual, community and toilet type time-variant controls	No	Yes	No	Yes
Time fixed-effects	No	Yes	No	Yes
Individual fixed-effects	No	No	Yes	Yes
Observations	23580	23580	2504	2504
Wald χ^2	0.015	345.584	2.527	884.139
	[0.002]	[0.066]	[0.002]	[0.066]
Pseudo R^2			0.002	0.066

Notes: Logit models are estimated, and marginal effects are reported. Individual and community time-variant controls include all individual level controls shown in Table 1. Estimated standard errors in parentheses are clustered at the sub-county level. * p < .1, ** p < .05, *** p < .01. [#] denotes p-value of statistics. Sample differences between Logit-RE and Logit-FE are due to the strict reliance of the conditional fixed effect logit model on variations in the outcome variable. The results remain qualitatively the same when we use the linearity probability model.

There are no statistically significant effects when we only employ random effect models (models 1 and 2). However, once we implement the fixed-effects approach and thereby account for time-invariant characteristics of individuals, we find that an increase in the improved water supply rate significantly reduces a household member's likelihood of suffering from water-related symptoms of illness (model 3). Once we add household, community and individual time-variant characteristics (model 4), statistical significance is not reached, and the effect size becomes substantially smaller while remaining negative.

A potential a priori expectation was a negative and significant effect of improved water supply on water-related illness symptoms. However, we attribute the observed non-significance to the fact that diarrhoeal symptoms may be associated with different kinds of sicknesses, some of which are not water-related. Unfortunately, data constraints limit our ability to include variables that could explain these symptoms. Hence, omitted variables, especially time-variant ones, could confound the coefficient of improved water supply rate, even with our micro-based approach using individual fixed-effects. Even so, the effect signs provide evidence of a relationship between water provision and the disease burden.

Effect on Water Collection Time

Finally, we investigated the impact of improved water supply on total water collection times (Table 4).

Table 4: Impact of improved water supply on total water collection time (minutes)

	(1) Total collection time	(2) Total collection time	(3) Total collection time	(4) Total collection time
Improved water supply capacity per capita	-10.645**	-7.636*	-39.625*	-20.577
	(4.408)	(4.126)	(21.455)	(19.377)
Household and community time-variant controls	No	Yes	No	Yes
Time fixed-effects	No	Yes	No	Yes
Household fixed-effects	No	No	Yes	Yes
Observations	4576	4576	4576	4576
R^2	0.005	0.037	0.005	0.019

Notes: OLS models are estimated. Household and community time-variant controls include all household level controls shown in Table 1. Estimated standard errors in parentheses are clustered at the sub-county level. * p < .1, ** p < .05, *** p < .01

The results show a consistent negative effect of improved water installation on the number of minutes spent on water collections. The effect is statistically significant for all estimations apart from model (4), which uses the full set of fixed-effects and household as well as community time-variant controls. Quantitatively, the results point to a substantial reduction in water fetching times, of between 20 to 40 minutes.

Methodologically, the results in Table 4 also serve as a robustness check on the findings and conclusions in Tables 2 and 3, as additional water supply should not negatively affect collection times even if citizens choose to use other sources. More importantly, the results provide indirect evidence of the relevance of water supply on issues such as household time allocation (especially for women and children), women's empowerment, and poverty reduction (see Ivens, 2008; Ray, 2007). In a rural African context, we would, for instance, expect a reduction in water collection time to translate into more time spent on studies/leisure for children, and more time spent on the farm or commercial activities for women (Jasper, Le, & Bartram, 2012; Koolwal & van de Walle, 2013).

V. CONCLUDING REMARKS

A significant proportion of development funds are devoted to water, sanitation, and hygiene projects. Development funds need to satisfy competing needs, and therefore assessing the impacts of diverse projects is relevant. The existing literature has investigated potential impacts of improved water provision on health and household welfare. While much of the literature shows a positive impact from water provision, methodological challenges in the design of the evaluations or the lack of funds for large-scale randomised control trials affect the extent to which policymakers can rely on the results.

We suggest a micro-based and cost-effective approach which entails evaluating past development projects and employing newly merged datasets to investigate the impact of improved water supply on water usage, health, and collection times. The empirical results indicated modest but positive effects of improved water provision on a household's usage of this water. We also find that an increase in the coverage of improved water may decrease water-related illness symptoms in household members. Furthermore, water collection times are substantially reduced as the supply of improved water increases.

The major contribution of our study, in addition to empirical findings, is our micro-based evaluation approach. It employs available datasets but merges them for impact evaluations (i.e., we merge administrative information on water supply with household socio-economic panel data). Methodologically, the merging of data allows us to employ household and individual fixed-effects when investigating potential impacts of development projects. Our approach systematically reduces endogeneity issues. Socio-economic data is collected in many countries at regular intervals and samples are large.

Moreover, a Hawthorne effect is less likely for socio-economic surveys. Our approach also allows us to evaluate a large number of past projects at the same time, and it is, thus, scalable. As new data need not be collected, it is highly cost-effective. There are, however, limitations as socio-economic data may not be detailed enough for specific research purposes. Moreover, it is challenging to account for infrastructure outlays that may affect households alongside the specific projects that are investigated. Nevertheless, our approach likely has higher external validity than is typically achieved with randomised control trials. In any case, it offers a viable complement to standard methods of programme evaluation in developing countries, which can be conducted at low cost using pre-existing administrative data and micro-surveys, years after development projects have occurred.

VI. REFERENCES

Aiello, A. E., Coulborn, R. M., Perez, V., & Larson, E. L. (2008). Effect of hand hygiene on infectious disease risk in the community setting: A meta-analysis. *American Journal of Public Health*, *98*(8), 1372–1381.

Amrose, S., Burt, Z., & Ray, I. (2015). Safe Drinking Water for Low-Income Regions. *Annual Review of Environment and Resources*, *40*(1), 203–231.

Angrist, J. D., & Krueger, A. B. (1999). Empirical Strategies in Labor Economics. In *Handbook of Labor Economics. Handbook of labor economics* (Vol. 3, pp. 1277–1366). Amsterdam: Elsevier.

Aziz, K. M., Hoque, B. A., Hasan, K. Z., Patwary, M. Y., Huttly, S. R. A., Rahaman, M. M., & Feachem, R. G. (1990). Reduction in diarrhoeal diseases in children in rural Bangladesh by environmental and behavioural modifications. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *84*(3), 433–438.

Begum, S., Ahmed, M., & Sen, B. (2011). Do Water and Sanitation Interventions Reduce Childhood Diarrhoea? New Evidence from Bangladesh. *Bangladesh Development Studies*, *34*(3), 1–30.

Blum, D., Emeh, R. N., Huttly, S. R. A., Dosunmu-Ogunbi, O., Okeke, N., Ajala, M., . . . Feachem, R. G. (1990). The Imo State (Nigeria) Drinking Water Supply and Sanitation Project, 1. Description of the project, evaluation methods, and impact on intervening variables. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *84*(2), 309–315.

Blum, D., & Feachem, R. G. (1983). Measuring the impact of water supply and sanitation investments on diarrhoeal diseases: Problems of methodology. *International Journal of Epidemiology*, *12*(3), 357–365.

Cairncross, S., Hunt, C., Boisson, S., Bostoen, K., Curtis, V., Fung, I. C. H., & Schmidt, W.-P. (2010). Water, sanitation and hygiene for the prevention of diarrhoea. *International Journal of Epidemiology*, *39 Suppl 1*, i193-205.

Capuno, J. J., Tan, C. A. R., & Fabella, V. M. (2015). Do piped water and flush toilets prevent child diarrhea in rural Philippines? *Asia-Pacific Journal of Public Health*, *27*(2), NP2122-32.

Clasen, T., Schmidt, W.-P., Rabie, T., Roberts, I., & Cairncross, S. (2007). Interventions to improve water quality for preventing diarrhoea: Systematic review and meta-analysis. *BMJ (Clinical Research Ed.)*, *334*(7597), 782.

Cruzatti, J. C., Drehery, A., & Matzat, J. (2019). *Aid and Health at the Regional Level Evidence from over 70,000 locations over the 1995-2014, mimeo, University of Heidelberg.*

Devoto, F., Duflo, E., Dupas, P., Parienté, W., & Pons, V. (2012). Happiness on Tap: Piped Water Adoption in Urban Morocco. *American Economic Journal: Economic Policy*, *4*(4), 68–99.

Dolan, C. B., BenYishay, A., Grépin, K. A., Tanner, J. C., Kimmel, A. D., Wheeler, D. C., & McCord, G. C. (2019). The impact of an insecticide treated bednet campaign on all-cause child mortality: A geospatial impact evaluation from the Democratic Republic of Congo. *PloS One*, *14*(2), e0212890.

Duflo, E. (2001). Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment, *91*(4), 795–813.

Dupas, P., & Miguel, E. (2016). Impacts and Determinants of Health Levels in Low-Income Countries. *National Bureau of Economic Research*.

Ejemot-Nwadiaro, R. I., Ehiri, J. E., Arikpo, D., Meremikwu, M. M., & Critchley, J. A. (2015). Hand washing promotion for preventing diarrhoea. *The Cochrane Database of Systematic Reviews*. (9), CD004265.

Esrey, S. A., Feachem, R. G., & Hughes, J. M. (1985). Interventions for the control of diarrhoeal diseases among young children: Improving water supplies and excreta disposal facilities *Bulletin of the World Health Organization*, *63*(4), 757–772.

Esrey, S. A., Potash, J. B., Roberts, L., & Shiff, C. (1991). Effects of improved water supply and sanitation on ascariasis, diarrhoea, dracunculiasis, hookworm infection, schistosomiasis, and trachoma. *Bulletin of the World Health Organization*, *69*(5), 609–621.

Fan, V. Y.-M., & Mahal, A. (2011). What prevents child diarrhoea? The impacts of water supply, toilets, and hand-washing in rural India. *Journal of Development Effectiveness*, *3*(3), 340–370.

Fink, G., Günther, I., & Hill, K. (2011). The effect of water and sanitation on child health: Evidence from the demographic and health surveys 1986-2007. *International Journal of Epidemiology*, *40*(5), 1196–1204.

Foster, T. (2013). Predictors of sustainability for community-managed handpumps in sub-Saharan Africa: Evidence from Liberia, Sierra Leone, and Uganda. *Environmental Science & Technology*, *47*(21), 12037–12046.

Gasana, J., Morin, J., Ndikuyeze, A., & Kamoso, P. (2002). Impact of Water Supply and Sanitation on Diarrheal Morbidity among Young Children in the Socioeconomic and Cultural Context of Rwanda (Africa). *Environmental Research*, *90*(2), 76–88.

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. J. (2016). *Impact Evaluation in Practice, Second Edition*: The World Bank.

Gertler, P. J., & Molyneaux, J. W. (1994). How Economic Development and Family Planning Programs Combined to Reduce Indonesian Fertility. *Demography*, *31*(1), 33.

Giliani, S., Gonzalez-Rozada, M., & Shargrodsky, E. (2009). Water Expansions in Shantytowns: Health and Savings. *Economica*, *76*(304), 607–622.

Gross, E., Günther, I., & Schipper, Y. (2018). Women Are Walking and Waiting for Water: The Time Value of Public Water Supply. *Economic Development and Cultural Change*, *66*(3), 489–517.

Hasan, K. Z., Briend, A., Aziz, K. M., Hoque, B. A., Patwary, M. Y., & Huttly, S. R. A. (1989). Lack of impact of a water and sanitation intervention on the nutritional status of children in rural Bangladesh. *European Journal of Clinical Nutrition*, *43*(12), 837–843.

Huttly, S. R. A., Blum, D., Kirkwood, B. R., Emeh, R. N., Okeke, N., Ajala, M., . . . Feachem, R. G. (1990). The Imo State (Nigeria) Drinking Water Supply and Sanitation Project, 2. Impact on dracunculiasis, diarrhoea and nutritional status. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *84*(2), 316–321.

Ivens, S. (2008). Does Increased Water Access Empower Women? *Development*, *51*(1), 63–67.

Jalan, J., & Ravallion, M. (2003). Does piped water reduce diarrhea for children in rural India? *Journal of Econometrics*, *112*(1), 153–173.

Jasper, C., Le, T.-T., & Bartram, J. (2012). Water and sanitation in schools: A systematic review of the health and educational outcomes. *International Journal of Environmental Research and Public Health*, *9*(8), 2772–2787.

Khanna, G. (2008). *The impact on child health from access to water and sanitation and other socioeconomic factors* (HEI Working Paper No. 02/2008).

Koolwal, G., & van de Walle, D. (2013). Access to Water, Women's Work, and Child Outcomes. *Economic Development and Cultural Change*, *61*(2), 369–405.

Kremer, M., Leino, J., Miguel, E., & Zwane, A. P. (2011). Spring Cleaning: Rural Water Impacts, Valuation, and Property Rights Institutions. *The Quarterly Journal of Economics*, *126*(1), 145–205.

Loevinsohn, M., Mehta, L., Cuming, K., Nicol, A., Cumming, O., & Ensink, J. H. J. (2015). The cost of a knowledge silo: A systematic re-review of water, sanitation and hygiene interventions. *Health Policy and Planning*, *30*(5), 660–674.

Madrigal, R., Alpízar, F., & Schlüter, A. (2011). Determinants of Performance of Community-Based Drinking Water Organizations. *World Development*, *39*(9), 1663–1675.

Majuru, B., Michael Mokoena, M., Jagals, P., & Hunter, P. R. (2011). Health impact of small-community water supply reliability. *International Journal of Hygiene and Environmental Health*, *214*(2), 162–166.

Ministry of Water, Lands and Environment (1999). *A national water policy.*

Nanan, D., White, F., Azam, I., Afsar, H., & Hozhabri, S. (2003). Evaluation of a water, sanitation, and hygiene education intervention on diarrhoea in northern Pakistan. *Bulletin of the World Health Organization*, *81*(3), 160–165.

National Planning Authority (2015). *Second National Development Plan (NDPII) 2015/16-2019/20.*

Norman, G., Pedley, S., & Takkouche, B. (2010). Effects of sewerage on diarrhoea and enteric infections: a systematic review and meta-analysis. *The Lancet Infectious Diseases*, *10*(8), 536–544.

Opryszko, M. C., Majeed, S. W., Hansen, P. M., Myers, J. A., Baba, D., Thompson, R. E., & Burnham, G. (2010). Water and hygiene interventions to reduce diarrhoea in rural Afghanistan: A randomized controlled study. *Journal of Water and Health*, *8*(4), 687–702.

Pitt, M. M., Rosenzweig, M. R., & Gibbons, D. M. (1993). The determinants and consequences of the placement of government programs in Indonesia. *The World Bank Economic Review*, *7*(3 (September 1993)), 319–348.

Ray, I. (2007). Women, Water, and Development. *Annual Review of Environment and Resources*, *32*(1), 421–449.

Rosenzweig, M. R., & Wolpin, K. L. (1986). Evaluating the effects of optimally distributed public programs: child health and family planning interventions. *American Economic Review*, *76*(3), 470–482.

Roushdy, R., Sieverding, M., & Radwan, H. (2013). *The impact of water supply and sanitation on child health: Evidence from Egypt.*

Ryder, R. W., Reeves, W. C., Singh, N., Hall, C. B., Kapikian, A. Z., Gomez, B., & Sack, R. B. (1985). The childhood health effects of an improved water supply system on a remote Panamanian island. *The American Journal of Tropical Medicine and Hygiene*, *34*(5), 921–924.

Strunz, E. C., Addiss, D. G., Stocks, M. E., Ogden, S., Utzinger, J., & Freeman, M. C. (2014). Water, sanitation, hygiene, and soil-transmitted helminth infection: A systematic review and meta-analysis. *PLoS Medicine*, *11*(3), e1001620.

Taylor, D. L., Kahawita, T. M., Cairncross, S., & Ensink, J. H. J. (2015). The Impact of Water, Sanitation and Hygiene Interventions to Control Cholera: A Systematic Review. *PloS One*, *10*(8), e0135676.

Tonglet, R., Isu, K., Mpese, M., Dramaix, M., & Hennart, P. (1992). Can improvements in water supply reduce childhood diarrhoea? *Health policy and planning*, *7*(3), 260–268.

UNESCO (2002). *Universal Declaration on Cultural Diversity*. Paris.

UNICEF\WHO (2015a). *Estimates on the Use of Water Sources and Sanitation Facilities. Uganda*. Uganda.

UNICEF\WHO (2015b). *Key Facts from JMP 2015 Report.*

Wasonga, J., Okowa, M., & Kioli, F. (2016). Sociocultural Determinants to Adoption of Safe Water, Sanitation, and Hygiene Practices in Nyakach, Kisumu County, Kenya: A Descriptive Qualitative Study. *Journal of Anthropology*, *2016*(23-24), 1–5.

Wolf, J., Prüss-Ustün, A., Cumming, O., Bartram, J., Bonjour, S., Cairncross, S., . . . Higgins, J. P. T. (2014). Assessing the impact of drinking water and sanitation on diarrhoeal disease in low- and middle-income settings: Systematic review and meta-regression. *Tropical Medicine & International Health: TM & IH*, *19*(8), 928–942.

Zeng-Sui, W., Shepard, D. S., Yun-Cheng, Z., Cash, R. A., Ren-Jie, Z., Zhen-Xing, Z., & Fu-Min, S. (1990). Reduction of enteric infectious disease in rural China by providing deep-well tap water. *Health Policy*, *14*(2), 155–156.

Zhang, J. (2012). The impact of water quality on health: Evidence from the drinking water infrastructure program in rural China. *Journal of Health Economics*, *31*(1), 122–134.

Zwane, A. P., Zinman, J., van Dusen, E., Pariente, W., Null, C., Miguel, E., . . . Banerjee, A. (2011). Being surveyed can change later behavior and related parameter estimates. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(5), 1821–1826.

APPENDIX

Table A1: Improved Community Water Source User Number Assumptions

Water source type	Users
Protected Spring	200
Shallow Well	300
Deep Borehole	300
Kiosk / Public Tap	150
Rainwater Harvesting Tank <10,000 l	3
Rainwater Harvesting Tank >10,000 l	6
Institutional connection	100
Yard Tap	24