Freshwater availability and water fetching distance affect child health in Sub-Saharan Africa

Amy J. Pickering[1,2]* *and Jennifer Davis*[2,3]*

1 Emmett Interdisciplinary Program in Environment and Resources, School of Earth Sciences, Stanford University
2 Woods Institute for the Environment, Stanford University
3 Environment and Water Studies, Civil and Environmental Engineering, Stanford University

*Address correspondence to either author. Email: amyjanel@gmail.com; phone: (510) 410-2666; Email: jennadavis@stanford.edu; phone: (650) 725-9170.

Supporting Information: Pages S1-S10, Tables S1-S3, Figure S1

Methods

Demographic and Health Survey (DHS) data. The DHS program is supported by the US Agency for International Development (USAID) and aims to generate nationally representative data on population, health, and nutrition for low- and middle-income countries. A subset of households were located in regions with missing GPS coordinates (N=2392), and respondents that were classified as "not *dejure* residents" did not answer questions about walk time to water source or water source type (N=3732). Such observations with missing data were excluded from the multivariate modeling analyses.

Data sources. The DHS data are available at www.measuredhs.com. The freshwater runoff data used in this analysis are available at http://www.grdc.sr.unh.edu/.

Water source type. In the case where a household source reported as "well or spring," the source was assumed to be a well (N=3103). When a household reported using a "protected source," the source was categorized to be a well (N=535). The category "gravity flow water," reported only in Uganda, was classified as "other" (N=55).

Sanitation type. Households reporting use of a bucket toilet (N=206), hanging toilet (N=1103), or "other" form of sanitation (N=469) were classified as lacking access to a sanitation facility in the stratified analysis.

Walk time. Respondents in the DHS survey are asked how long it takes to walk to their households' main drinking water source. This measurement is used as an indicator of water access; however, it cannot be used to estimate how much time the household spends collecting water in total, because the number of trips to the water source is not recorded. The range in reported walk time was 0 to 900 minutes, with 99% of the walk times under 180 minutes (3 hours). Although high values such as 600 (10 hours) or 900 (12 hours) minutes of round-trip walk time may be erroneous, we did not feel justified in omitting or adjusting these observations. However, we did re-estimate all models using a trimmed walk time to

source variable (see Table S3, column G). This trimmed walk time variable was created by treating all values over 360 minutes (6 hours) as missing data (n=154). The mean of this trimmed walk time variable is 24 minutes (SD 35).

Development of dietary diversity index. A dietary diversity index (DDI) was created as a proxy for nutritional intake of children.[1] Mothers reported whether or not each child in the household had consumed a variety of different types of foods in the 24 hours prior to interview. An index value for each child in the dataset was created by summing the number of different food types s/he had consumed, using the following categories: 1) bread, rice, noodles, potatoes, cassava, tubers, porridge, or cereal; 2) beans, peas, lentils, or nuts; 3) milk other than breast milk, cheese, or yogurt; 4) meat, poultry, fish, eggs, and other animal products; 5) vitamin A-rich fruits and vegetables, *e.g.*, pumpkin, red or yellow yams, squash, carrots, red sweet potatoes, dark green leafy vegetables, mango, or papaya; 6) other fruits and vegetables (or fruit juices); and 7) foods made with oil, fat, or butter. Table S3, column D shows results from IV models including the DDI.

Precipitation. Gridded mean monthly precipitation is used to control for seasonality in an alternative IV model specification presented in Table S3, column C. Mean monthly precipitation assigned to households is 113 mm (median 95, SD=103, N=196675). Precipitation data were downloaded from WorldClim – Global Climate Data (www.worldclim.org) as described in Hijmans *et al.* (2005).[2]

Instrument diagnostics. The possibility of weak identification by the instrumental variable was rejected using the Kleibergen-Paap Wald rk F statistic, whose values were higher than critical values proposed by Stock and Yogo (2005) for a single endogenous regressor (F>200).[3-6]

Discussion

Households with water on plot. It was found that households with on-plot water supplies were more likely to live in areas with less runoff as compared to households walking very short distances to water sources. This is most likely because 95% of households with water

on their plots had access to piped water (51%) or a well (44%), both of which require infrastructure. Since a walk time to water source of 0 minutes was used for households with water on plot in the instrumental variable models, the unique relationship with runoff specific to these households may have weakened the first-stage regression. Exclusion of these households is complicated by the possibility that runoff indirectly affected selection of these households into having their own water source. Regardless, the magnitudes of the estimated effect of walk time on child health outcomes are only slightly reduced when households with water on plot are excluded (see Table S3, column A).

Figure S1. Conceptual model of causal pathways by which water access may affect child health.

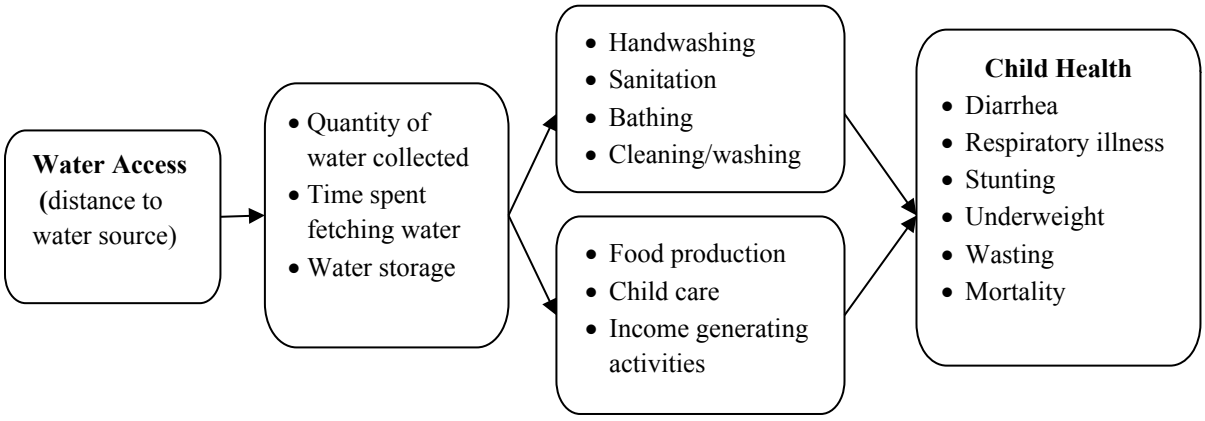

Table S1. First stage regression of round-trip walk time to source.

Variable	*β*		Std. Err.	t
Runoff	-0.0095	**	0.0002	-53.2
Piped water	-18.3696	**	0.2913	-63.1
Well	-11.4049	**	0.2383	-47.9
Spring	-0.7080	†	0.3258	-2.2
Rainwater	-33.3177	**	1.0954	-30.4
Vendor	-4.0007	*	1.1869	-3.4
Other	-4.7179	†	1.0728	-4.4
Child age (months)	0.0056		0.0046	1.2
Gender [a]	0.0286		0.1640	0.2
Urban [b]	-7.5486	**	0.2202	-34.3
Constant	38.8701	**	0.2634	147.6

[a] *Male = 1, female = 0*

[b] *Urban = 1, rural = 0*

***P<0.001, *P<0.01, † P <0.05*

Reference water source category is surface water.

N=183267, F=1251, Adj. R-squared = 0.06

Table S2. Re-estimated results following stratification by access to sanitation. Average marginal effect of 1-min change in walk time to water source on health outcomes.

Health Outcome	**A. Household practices open defecation**		**B. Household has access to a sanitation facility**	
	ME (SE)	***N***	**ME (SE)**	***N***
Diarrhea	0.0036 **	*59148*	0.0058 **	*101353*
	(0.0004)		(0.0005)	
Fever	0.0027 **	*59054*	0.0032 **	*101253*
	(0.0005)		(0.0007)	
Cough	0.0006	*59090*	0.0047 **	*101259*
	(0.0006)		(0.0006)	
WAZ	-0.0291 **	*38307*	-0.0458 **	*63554*
	(0.0029)		(0.0067)	
HAZ	-0.0137 **	*40844*	-0.0231 **	*65575*
	(0.0026)		(0.0046)	
WHZ	-0.0300 **	*38307*	-0.0341 **	*63554*
	(0.0031)		(0.0058)	
Mortality	0.0003	*67583*	0.0013 **	*114895*
	(0.0002)		(0.0003)	

***P<0.001, *P<0.01, † P <0.05*

Table S3. Average marginal effect (ME) estimates from alternative IV model specifications.
Column A presents estimates after excluding households with their own water source on the premises. Column B models include a household wealth index variable. Column C models include mean monthly precipitation as a control for seasonality. Column D includes a dietary diversity index (DDI) variable as proxy for child nutritional intake. Column E models include the number of years of education of the respondent. Column F models include a dummy variable indicating whether the household participates in agriculture as their main form of employment. Column G shows model estimates when the walk time to water source variable is trimmed to exclude outliers over 6 hours. Column H models use a natural log transformed walk time variable. All models control for child age in months, child gender, urban *versus* rural, and water source type.

Health	**A. Exclude own source**		**B. Wealth index**		**C. Precipitation**		**D. Nutrition (DDI)**	
Outcome	**ME (SE)**	***N***	**ME (SE)**	***N***	**ME (SE)**	***N***	**ME (SE)**	***N***
Diarrhea	0.0042 **	*132250*	0.0047 **	*160734*	0.0036 **	*160734*	0.0039 **	*66058*
	(0.0003)		(0.0003)		(0.0004)		(0.0004)	
Fever	0.0031 **	*132073*	0.0031 **	*160539*	0.0007	*160539*	0.0004	*65994*
	(0.0004)		(0.0004)		(0.0006)		(0.0005)	
Cough	0.0031 **	*132110*	0.0028 **	*160579*	-0.0011	*160579*	0.0014 *	*65972*
	(0.0004)		(0.0004)		(0.0006)		(0.0006)	
WAZ	-0.0315 **	*82752*	-0.0358 **	*102013*	-0.0326 **	*102013*	-0.0153 **	*41231*
	(0.0027)		(0.0033)		(0.0031)		(0.0026)	
HAZ	-0.0169 **	*86843*	-0.0151 **	*106573*	-0.0154 **	*106573*	-0.0078 *	*44914*
	(0.0023)		(0.0025)		(0.0028)		(0.0028)	
WHZ	-0.0250 **	*82752*	-0.0339 **	*102013*	-0.0307 **	*102013*	-0.0120 **	*41231*
	(0.0026)		(0.0033)		(0.0032)		(0.0026)	
Mortality	0.0005 *	*150812*	0.0007 *	*183267*	n.a.		n.a.	
	(0.0002)		(0.0002)					

***P<0.001, *P<0.01, † P <0.05*

Table S3 (continued)

Health	**E. Education**		**F. Agriculture**		**G. Trimmed walk time**		**H. Log (walk time)**	
Outcome	**ME (SE)**	***N***	**ME (SE)**	***N***	**ME (SE)**	***N***	**ME (SE)**	***N***
Diarrhea	0.0044 **	*160613*	0.0049 **	*150238*	0.0049 **	*160635*	0.1424**	*160734*
	(0.0003)		(0.0003)		(0.0003)		(0.0084)	
Fever	0.0027 **	*160418*	0.0034 **	*150042*	0.0031 **	*160440*	0.0922**	*160539*
	(0.0004)		(0.0004)		(0.0004)		(0.0130)	
Cough	0.0027 **	*160458*	0.0029 **	*150090*	0.0027 **	*160480*	0.0809**	*160579*
	(0.0004)		(0.0004)		(0.0005)		(0.0140)	
WAZ	-0.0302 **	*101942*	-0.0379 **	*94377*	-0.0394 **	*101969*	-1.1053**	*102013*
	(0.0027)		(0.0033)		(0.0033)		(0.1326)	
HAZ	-0.0139 **	*106501*	-0.0180 **	*98873*	-0.0186 **	*106529*	-0.4981**	*106573*
	(0.0024)		(0.0026)		(0.0026)		(0.0768)	
WHZ	-0.0277 **	*101942*	-0.0343 **	*94377*	-0.0358 **	*101969*	-1.0032**	*102013*
	(0.0027)		(0.0032)		(0.0032)		(0.1257)	
Mortality	0.0005 *	*183124*	0.0008 **	*171021*	0.0008 **	*183156*	0.0247**	*183267*
	(0.0002)		(0.0002)		(0.0002)		(0.0067)	

***P<0.001, *P<0.01, † P <0.05*

References

1. Arimond, M.; Ruel, M. T., Dietary Diversity Is Associated with Child Nutritional Status: Evidence from 11 Demographic and Health Surveys. *The Journal of Nutrition* **2004,** *134* (10), 2579-2585.
2. Hijmans, R. J.; Cameron, S. E.; Parra, J. L.; Jones, P. G.; Jarvis, A., Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **2005,** *25* (15), 1965-1978.
3. Staiger, D.; Stock, J. H., Instrumental variables regression with weak instruments. *Econometrica* **1997,** *65*, 557-586.
4. Stock, J. H.; Yogo, M., Testing for weak instruments in linear IV regression. *In Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg, ed. D. W. K. Andrews and J. H. Stock* **2005**, 80–108.
5. Angrist, J. D.; Pischke, J. S., *Mostly harmless econometrics: An empiricist's companion*. Princeton Univ Pr: 2008.
6. Kleibergen, F.; Paap, R., Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics* **2006,** *133* (1), 97-126.