IDEC DP^2 Series Vol. 7 No. 5

Household Access to Water and Education for Girls: The Case of Mountain Villages in Nepal

Ram P. Dhital
Institute of Engineering and Alternative Energy Promotion Center

Takahiro Ito
Graduate School of International Cooperation Studies
Kobe University

Shinji Kaneko
Graduate School for International Development and Cooperation (IDEC)
Hiroshima University

Satoru Komatsu
School of Global Humanities and Social Sciences
Nagasaki University

Yuichiro Yoshida
Graduate School for International Development and Cooperation (IDEC)
Hiroshima University

Department of Development Policy
Division of Development Science
Graduate School for International Development and Cooperation (IDEC)
Hiroshima University
1-5-1 Kagamiyama, Higashi-hiroshima
7398529 Japan

Household Access to Water and Education for Girls: The Case of Mountain Villages in Nepal☆

March 27, 2018

Ram P. Dhital,[*] Takahiro Ito,[†] Shinji Kaneko,[‡] Satoru Komatsu,[§] and Yuichiro Yoshida[¶]

***Abstract*:**

In many developing countries including Nepal, water fetching is traditionally conducted by women and girls. In a mountainous hinterland of Nepal without systematic water and electricity supply, it is inevitably laborious, and as a result, girls receive fewer educational opportunities than boys. This paper aims to identify the causal effect of household water accessibility on children's educational attainment measured by school attendance, grade repetition, and completion of primary and lower secondary schools in remote mountainous villages in Nepal. The estimation results evince that water hauling hinders girls from completing schooling, indicating that a one-hour increase in the time spent going to and from the water source will decrease the probability of female children completing primary school by 24.1 percentage points, while male children do not drop out, although they are more likely to repeat a grade. This implies that increased water accessibility—for example, by providing a solar water pumping system that the Nepali government promotes—improves household wellbeing, particularly girls' educational attainment, by reducing the burden of water collection.

JEL codes: I24, I25, O13

☆ This work was supported by JSPS KAKENHI (grant numbers JP25257102 and 16K21162).

[*] Director, Institute of Engineering and Alternative Energy Promotion Center, Lalitpur Sub Metropolitan City 14364, Nepal. E-mail: ram.dhital@aepc.gov.np

[†] Associate Professor, Graduate School of International Cooperation Studies, Kobe University, 2–1, Rokkodai-cho, Nada-ku, Kobe, Hyogo, 657-8501, Japan. E-mail: takahiro.ito@lion.kobe-u.ac.jp.

[‡] Professor, Graduate School for International Development and Cooperation, Hiroshima University, 1-5-1, Kagamiyama, Higashi-Hiroshima, Hiroshima, 739-8529, Japan. E-mail: kshinji@hiroshima-u.ac.jp.

[§] Associate Professor, School of Global Humanities and Social Sciences, Nagasaki University. 1-14 Bunkyo-machi, Nagasaki, 852-8521, Japan. E-mail: skomatsu@nagasaki-u.ac.jp.

[¶] Professor, Graduate School for International Development and Cooperation, Hiroshima University, 1-5-1 Kagamiyama, Higashi-Hiroshima, Hiroshima-ken, 739-8529, Japan. E-mail: yuichiro@hiroshimau.ac.jp.

1. Introduction

Being born a girl still persists as a primary cause for exclusion from one of the most basic human rights: education. This fetal inequality cripples the lives of millions of women and girls, especially in the developing hemisphere. To rectify gender disparities, sustained efforts have been made globally, as "gender equality" and "universal education" have been named in the Millennium Development Goals (MDGs) and subsequent Sustainable Development Goals (SDGs). However, the majority of the world's 750 million illiterate adults, or exactly 473 million, are female, and the situation has not improved in the last two decades (UNESCO, 2017). Although there has been some improvement for younger generations during the same period, girls are still disadvantaged in terms of school enrollment and literacy as compared with boys.

One of the biggest difficulties behind the scenes lies in the gender-based differences in family roles. Household chores such as cooking, cleaning, washing, and child-rearing are typical women's work in developing countries. In mountain villages in Nepal, the setting of this study, water fetching is added to the top of the list. People there live on the slope of a mountain, where the water source is typically at the bottom and electricity from the national grid does not reach; hence, the water supply is solely dependent upon labor. Like in many other developing countries, water fetching is traditionally conducted by women and girls in Nepal. One such consequence is reduced opportunities for girls to attend school because they spend a large portion of time engaged in water collecting activities for their households.

In this study, we examine the link between water accessibility and girls' (and boys') educational attainment, using original household data from remote mountain villages in Nepal. Access to improved water sanitation, together with gender equality and universal education, has always been an urgent issue in the international community, as indicated in the MDGs and SDGs. In fact, 2.4 billion people still lack access to improved water sanitation facilities worldwide (WHO/UNICEF, 2015). In Nepal, only 22 percent of the people in mountain villages have a private connection to a safe drinking water source (i.e., piped water). Furthermore, households with access only to unreliable water sources, such as river and spring water, account for 30% of the population (GON, 2011), indicating that they have to engage in strenuous water-carrying activities every day through the precipitous slope of a Nepali mountain. Previous studies have documented that children's educational attainment is associated with water accessibility due to water collecting activities (Nankhuni and Findeis, 2004; Koolwal and Walle, 2013; Nauges, 2017). Thus, establishment of a water supply system is doubly important from the perspective of developmental policy: it can improve children's educational attainment (and probably gender equality in education), as well as increase household welfare by reducing time spent fetching water. Water resource policies should be one of the first and foremost priorities in such marginal settlements with no electricity and running water. Thus, exploring the association between water accessibility and education is of consequence in its own right.

In addition, this study contributes to the literature twofold. The first contribution stems from the uniqueness of our primary dataset. We conducted a survey of 2,641 households in 45 wards (i.e., villages) without electricity and a water supply system in remote and isolated mountain villages in Nepal in 2014 and 2015. Remote and isolated settlements without basic utilities are of increasing significance to achieve the principle of "Leaving No One Behind," the core concept embodied in the SDGs. However, the body of empirical research on the role of basic public services in such settlements with a relatively large-scale dataset is surprisingly sparse. In South Asia, where progress in alleviating poverty has been steadily made, inaccessible mountain villages in Nepal—rural areas in one of the poorest countries—are the last hurdle to achieve the SDGs in this region. Therefore, this study provides an essential clue for resolving basic infrastructure development issues that impact human capital, especially women, in high poverty, rural areas.

The second contribution is the methodological rigor in which this study attempts to isolate the impact of household access to water on education. An analysis with observational data usually requires a couple of identification assumptions. The source of variation employed in this study for identifying the effect of water accessibility stems from geographical differences within villages that are considered to be exogenous, as will be explained later. We believe that our identification assumption is less assertive than those in prior studies. In addition, the assumption can be tested empirically to determine the validity of our identification strategy; exploiting a natural experimental setting with an empirically falsifiable assumption is one of the advantages of our empirical strategy.

The rest of the paper is structured as follows. In the next section, we explain our research design and the dataset used in the analysis. Section 3 presents estimation results, which show that water hauling hinders girls from completing schooling, indicating that a one-hour increase in the time spent going to and from the water source will lower the probability of female children's primary school completion by 24.1 percentage points, and although male children do not drop out of school, they are more likely to repeat a grade. Finally, Section 4 presents our conclusion.

2. Research Design

2.1. Empirical Framework

We estimate the effect of household water accessibility on children's educational outcomes based on the following equation:

$$E_{ijs} = \beta_b(w_j \times Boy_i) + \beta_g(w_j \times Girl_i) + X_{ij}\boldsymbol{\gamma} + \delta_s + u_{ijs} \quad (1)$$

where E_{ijs} is an educational outcome of child i in household j in ward (village) s, w_j is treatment, i.e., household access to water, which is included as interaction terms with gender dummies (Boy_i and $Girl_i$) to allow for the heterogeneous impact of water accessibility on educational attainment between genders, X_{ij} is a vector of individual and household characteristics, δ_s represents village fixed effects, u_{ijs} is an unobserved component, and β_b, β_g, and $\boldsymbol{\gamma}$ are the parameters to be estimated.

Just as other studies that evaluate the impact of a public good/service using household data across many communities, two typical concerns must be addressed to estimate the causal influence of water accessibility, β_b and β_g. First, provision of public goods and services may be potentially associated with a social or economic hierarchy; the higher a household is on the social/economic ladder in a community, the better the household's access to public goods/services is. If there exists such between-household heterogeneity within a community, which is not completely observed and is partly captured in u_{ijs}, it causes the estimated effect to be biased. Second, wealthier communities may have more public goods and services than poorer communities. Such heterogeneity between communities is usually controlled by community dummies as fixed effects (δ_s), but otherwise its existence results in misidentification of the causal effect.

Thus, the key to isolate the impact of household access to water is determining whether the confounding of heterogeneity within a community and between communities, if any, can be purged. One obvious way of addressing the issue is by randomization, i.e., to distribute public goods/services (or place facilities that provide them) in a random manner, both across communities and within a community. However, it is practically implausible to conduct such an experiment, particularly in a sustainable way, in many types of academic research, including our study, which is set in remote and isolated mountain villages without basic utilities. In facing this difficulty, almost all studies hinge on identification assumptions without a randomized experiment.[1] For example, several studies have addressed heterogeneity within a community by using the community-averaged variable of interest (or community characteristics as instruments), by assuming no confounders driven by between-community heterogeneity after controlling several community-level variables (see Ilahi and Grimard [2000], Koolwal and Walle [2013], and Nauges [2017] for water; and Grogan and Sadanand [2013] for electricity.)

On the other hand, some studies have addressed both heterogeneities by controlling entity- (unit of observation) or community-fixed effects using panel data (see, for example, Mangyo [2008], Gamper-Rabindran et al. [2010], and Zhang [2012] for water; and Khandker et al. [2013] for electricity). However, they rely on the assumption that changes in the treatment status in a community over years are exogenous (conditional on several variables they presume to be the key determinants of the change). Therefore, if the improvements are associated with (potential) demands for better child outcomes, as with other public policies in which the treatment status is often determined based on the

[1] There are a few exceptions that have utilized a randomized experiment. For example, Kremer et al. (2011) employed a randomized evaluation on the health impact of a water quality intervention in Kenya. Closely related to our study is the work by Devoto et al. (2012), who used a randomized experiment to study the impact of a private connection to the piped water system in urban Morocco. However, because their focus is on urban dwellers in Morocco, the time burden of water collection among their sampled households is considerably shorter than rural dwellers who are our main focus. Probably partly due to this, they found no causal link between in-home water connections and children's education outcomes.

present (or potential) status of the outcome, reverse causality may prevent the isolation of the causality. Dinkelman (2011), who explored the causal impact of electrification on employment growth, is the only exceptional study that utilized the panel fixed effects approach in combination with a natural experimental framework, where land gradient is instrumented for the implementation of an electricity project after eliminating entity-fixed effects.[2]

This study also utilizes a natural experimental setting to address the above-mentioned typical empirical issues. Specifically, we sever the correlation between within-community heterogeneity and treatment status (household access to water) by focusing on distance to natural water sources, while we eliminate the influence of between-community heterogeneity by controlling community-fixed effects. The location of natural water sources such as rivers, seepages, ponds, and springs are naturally (geographically) determined, and therefore, it seems plausible to assume that the proximity to them is exogenous to household characteristics, unlike the provision of public goods/services. We test its validity empirically based on several tests as will be shown in Sections 2.3 and 3.1, and thus exploit a natural experimental setting with an empirically falsifiable assumption as one of the advantages of our empirical strategy.

2.2. Data and Sample Features

Data used in the analysis are from a household survey conducted by the authors in remote mountain villages in Nepal in 2014 and 2015 with the cooperation of a government institution, Alternative Energy Promotion Center (AEPC), which is under the Ministry of

[2] Another strand of literature employs the propensity score matching (PSM) method, which hinges on the "selection-on-observables" assumption (see, for instance, Jalan and Ravallion, 2003 for piped water access; and Lokshin and Yemtsov, 2005 for infrastructure rehabilitation projects).

Science and Technology of Nepal.[3] In the main survey, based on cluster random sampling, we chose 45 wards (i.e., villages) randomly from 22,545 wards across the country that have neither a water supply system nor electricity, and 2,641 households in the 45 wards were interviewed.[4] The location of our survey sites is shown in Figure 1.

[Figure 1]

Table 1 summarizes the water accessibility of our sample households. Panel A of Table 1, which reports the mode of fetching water, shows that among 2,641 households in the entire sample, 1,653 households (62.5%) carry water directly from the primary or secondary source, or both. The remaining 988 households do not collect water directly from the source, and the majority of those households (935 households) uses a public well/storage. Panel B summarizes the water collection activities of households who carry water directly from the water source. On average, one water collection trip takes about 30 minutes, and households do four trips per day, carrying 1.4 bottles—each bottle with a capacity of 15 liter—each trip. As mentioned earlier, water collection is considered women's work in our study region: 75.6% of adult women and 21.0% of non-adult women (aged between 6 and 19 years) engage in water collection activities, while the percentages of adult and non-adult men are no more than 19.0 and 11.1, respectively.

[Table 1]

As explained in the previous section, we limit the sample used in the analysis to households that have no public/private well or storage and carry water directly from a natural water source. This is because the key source of variations in our empirical strategy comes from geographical differences, i.e., accessibility to natural water sources. Among

[3] AEPC has promoted renewable energy technologies, such as the solar photovoltaic and micro-hydro water pumping systems, in Nepal since 1996.
[4] See Appendix I for details of our household survey.

1,653 households that collect water directly from natural water sources, excluding 486 households that have no school-age children, 2,512 children live in 1,167 households.

Regarding the empirical variables in Equation (1), we use four different educational outcomes (E_{ijs}): current (or last attended) grade and the number of grade repetitions for the sample of children aged 6-16 years; and dummies for completing primary (5th grade) and lower secondary education (8th grade) for children aged 14-16 years. Household access to water (w_j) is measured by hours spent to make one round trip to the natural water source. If a household uses multiple natural water sources, we use the closest one in terms of time. As control variables (X_{ij}), we employ individual and household characteristics such as gender and age of the child, household size, dependency ratio, age/gender/education of household head, log of annual household income, and language- and social group-fixed effects. Also, we include dummies for survey month as another control. Table 2 reports the summary statistics of main variables used in the analysis.

[Table 2]

2.3. Validity of the Identification Assumption

The internal validity of our identification strategy hinges on the assumption that distance to natural water sources is independent from observed/unobserved determinants of educational attainment of children within a community. However, there could be several counter-arguments against our assumption. For example, if wealthier households live closer to natural water sources for some reason—e.g., by their residential selection—the proximity to natural water sources may reflect such households' affluence, which may be partly unobserved and affect children's educational attainment. Moreover, accessibility to natural water sources may reflect accessibility to other public facilities, such as schools, health centers, and village headquarters. If this is the case, the influence of household access to natural water sources is confounded with the influences of accessibility to other facilities.

To validate our identification strategy, we execute three different tests. First, we compare household characteristics between two groups, namely those who live closer to and farther from the natural water source than the average household in the community (the balancing test). Second, we examine correlation coefficients between the distance to the natural water source and several household characteristics (the correlation test).[5] Table 3 reports the results of these tests: the balancing test in Columns (1) to (5) shows that the difference is narrowly estimated to be zero for all characteristics, and the correlation test in Columns (6) and (7) also shows that correlation coefficients are all close to zero.

These results confirm that observed household characteristics are orthogonal to distance to the natural water source. For example, from the results on household income (row 1) and education of household head (rows 3 to 5), those who live farther from a water source are not relatively poorer or less educated. In addition, the lack of relationship between the year the house was built (rows 7 to 10) and water accessibility indicates that those who are far from the natural water source in a community are not new households that settled from outside the community or split from their parents' family. Thus, water accessibility does not symbolize social and economic status of the household in the community. More importantly, distance to the water source has nothing to do with

[5] Note that distance to the natural water source (measured in hours) here is demeaned from the community (ward) average.

distance to schools as shown in the last two rows. These results support our assumption that water accessibility is orthogonal to other unobserved household characteristics as well.

[Table 3]

Third, we employ the instrumental variables (IV) regression technique and determined the direction of the change in the coefficient. We use the median hours that households' neighbors spend going to and from the natural water source as the instrument. The definition of neighbors consists of the following criteria: (i) households within a 100-meter radius; (ii) if there is no household within a 100-meter radius, households within a 200-meter radius are used; and (iii) if there is no household within a 200-meter radius, we expanded the radius in 100-meter increments up to a 500-meter radius.[6] Different scenarios can potentially explain a possible change in the coefficient brought by the use of the IV estimation from the ordinary least squares (OLS) estimation. For instance, as discussed earlier, if the place of the natural water source or residence within a community is determined endogenously based on the demand for children's education, the effect of water accessibility is likely to be overestimated in magnitude.

At the same time, the reverse causality issue may also threaten our identification strategy. Because household access to water is measured by reported hours spent fetching water from the sources, households whose main water carrier is a child may report hours spent by children.[7] It is naturally expected that it will take children more time to carry water than adults, and therefore, not attending school and spending more time fetching

[6] 83 households (3.3%) have no "neighbors" probably due to errors in GPS data, and they are excluded from the IV estimations.

[7] The survey questionnaire contains several questions about the primary and secondary drinking water sources such as the type of sources, means of carrying, times spent to and from the sources, main carriers, the number of times and water bottles at one time, etc. Regarding the time spent to and from the sources, we cannot identify the carrier.

water could be simultaneous outcomes, because decisions on schooling and laboring are simultaneously made. This may also cause our OLS estimate to be biased upwardly in magnitude.[8]

Therefore, by comparing the size of the coefficients in OLS and IV estimations, we determined whether the endogeneity issues mentioned above exist. The results are reported in the next section.

3. Estimation Results

3.1. Effect on Scholastic Grade and Grade Repetition

Table 4 presents the OLS and IV estimation results for Equation (1) for the current (or last-attended) grade (Columns 1 and 3) and the number of grade repetitions (Columns 2 and 4). In these estimations, the full sample of children aged between 6 and 16 years is used.[9] The results in Column 1 show that the impact of water collection activity on the current grade is negative but insignificant for both boys and girls. The results in Column 2 show that the impact of water collection on the number of repetitions is positive with significance only for boys; the results indicate that a one-hour increase in the time spent going to and from the natural water source increases the number of grade repetitions by 0.066 for boys.

[Table 4]

We now turn to the results of the IV estimations shown in Columns (3) and (4) of Table 4. Note that the number of observations in the IV estimation is smaller than in

[8] On the other hand, employing a subjective measure of water accessibility (hours spent on water hauling) may raise the issue of measurement errors. If this is the case and the classical-measurement-error assumptions hold, IV estimation eliminates attenuation bias.

[9] Among all 2,512 children aged between 6 to 16 years, 30 and 14 children have missing information on their current (or last attended) grade and number of grade repetitions, respectively.

the OLS estimation because neighbor data for the first-stage IV regressions are unavailable for households with inaccurate GPS information or no neighbors. By comparing the results between the OLS and IV estimations, we found that the OLS estimates are smaller in magnitude, implying the existence of downward bias in absolute value. As already mentioned in Section 2.3, if water accessibility confounds unobserved household heterogeneity within a village or if the reverse causality due to simultaneous decisions on schooling and laboring matters, OLS estimates will be overstated. Therefore, these smaller OLS estimates in absolute value indicate that the endogeneity issues are not problematic, or negligible if they exist. Rather, the larger IV estimates (in magnitude) indicate elimination of attenuation bias arising from a measurement error in the water fetching time (i.e., treatment), which is present in the OLS estimation.

Regarding the insignificant result of repetition for girls in Column (4), it seems to imply that girls tend to withdraw from school when they have to repeat a grade. This conjecture is indeed supported by the results of the impact of water accessibility on completion of primary education as discussed in the next section.

3.2. Effect on the Completion of Education

Table 5 presents the OLS and IV estimation results of the impact of household access to water on completion of primary education (Columns 1 and 3) and lower-secondary education (Columns 2 and 4). The sample consists of children aged between 14 and 16 years.[10] The OLS estimation results in Column 1 show that a one-hour increase in the time spent going to and from the natural water source decreases the probability of

[10] Again, the number of observations in the IV estimations is smaller because neighbor data for the first-stage IV regressions are unavailable for some households because of wrong GPS information or no neighbors.

completing primary education for girls by as large as 15.3% points with a significance level of 5%. The estimated coefficient is negative but insignificant for boys; they may repeat a grade, but not drop out, and complete primary education. As shown in Column 2, our data could not detect a significant impact on the completion of lower-secondary school.

[Table 5]

These results are supported by the IV estimations as well. Column 3 of Table 5 shows that the girls with an additional hour of water hauling are 24.1 percentage points more likely to drop out of primary school. This accounts for about one third of the completion rate of primary school for girls (76.9%). Again, the results indicate the existence of attenuation bias in the OLS estimates, rather than an upward bias in magnitude, denying the possibility of endogeneity due to residential choice within a village and the reverse causality between schooling and household laboring.

3.3. Potential Causal Paths of the Estimated Impacts

To understand the mechanism behind the results better, we conduct the following set of additional causal paths analyses. Table 6 shows the impact of water hauling activity on engagement in miscellaneous household chores, including water hauling, child/elder/sick care, cleaning, and laundry, as well as all household chores. The results indicate that an increase in hours spent on water collection increases the likelihood that younger children (aged 6 to 9 years) are engaged in child/elder care, cleaning, and laundry for both boys and girls (weakly and significantly, as shown in Columns 4, 6, and 8). Column 10 shows that when the household is far from natural water sources, boys aged 14 to 16 years and girls aged 6 to 9 years are more likely to engage in domestic duties. The results indicate that when water accessibility is low, households may cope by increasing older boys' and younger girls' participation in household duties. This is potentially causing the result of increased grade repetition for boys with more water hauling activities. In addition, Girl dummy and its interaction with age categories are indeed all positive and many are significant; girls' participation in household chores is inherently high relative to boys. The results show that participation of younger girls aged 6 to 9 years increases this ratio even further when water accessibility is low, which is potentially causing girls' low completion of primary school.

[Table 6]

4. Conclusion

As in many other developing countries, water fetching is traditionally conducted by women and girls in Nepal. In its mountainous villages without systematic water and electricity supply, water collection becomes inevitably laborious, and as a result girls receive fewer educational opportunities than boys. This paper identified the adverse effect of water collecting activities by children, particularly girls, on their educational attainment measured in terms of school attendance, repetition, and completion of primary and lower secondary schools in remote mountainous villages in Nepal. The estimation results consistently show that the children's water collection activities never positively affect their school attendance and educational attainment. With increased water hauling activities, boys repeat grades more, yet they still tend to complete primary and lower-secondary education. However, girls in households who spend one more hour going to and from the water source have 24.1% lower probability of completing primary education. In other words, with more water hauling, girls do not repeat grades, but rather, they simply drop out. This implies that improvements in water accessibility—for example, by providing a solar water pumping system that the Nepali government promotes—improve household wellbeing, particularly girls' educational attainment, by reducing the burden of water collection.

References

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

Dinkelman, Taryn. 2011. The Effects of Rural Electrification on Employment: New Evidence from South Africa. *American Economic Review*, 101: 3078-3108.

Gamper-Rabindran, Shanti, Shakeeb Khan, and Christopher Timmins. 2010. The Impact of Piped Water Provision on Infant Mortality in Brazil: A Quantile Panel Data Approach. *Journal of Development Economics*, 92(2): 188-200.

GON (Government of Nepal). 2011. *Nepal Living Standards Survey 2010/11: Statistical Report Volume One*. Kathmandu: Central Bureau of Statistics, National Planning Commission Secretariat, GON.

Grogan, Louise and and Asha Sadanand. 2013. Rural Electrification and Employment in Poor Countries: Evidence from Nicaragua. *World Development*, 43: 252-265.

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(1): 45–75.

Jalan, Jyotsna and Martin Ravallion. 2003. Does Piped Water Reduce Diarrhea for Children in Rural India? *Journal of Econometrics*, 112(1): 153–73.

Khandker, Shahidur R. and Douglas F. Barnes, and Hussain A. Samad. 2013. Welfare Impacts of Rural Electrification: a Panel Data Analysis from Vietnam. *Economic Development and Cultural Change,* 61(3): 659–692

Koolwal, Gayatri, and Dominique Van de Walle. 2013. Access to Water, Women's Work, and Child Outcomes. *Economic Development and Cultural Change*, 61(2): 369-405.

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

Lokshin, Michael, and Ruslan Yemtsov. 2005. Has Rural Infrastructure Rehabilitation in Georgia Helped the Poor? *World Bank Economic Review*, 19(2): 311–333.

Mangyo, Eiji. 2008. The Effect of Water Accessibility on Child Health in China, *Journal of Health Economics*, 27(5): 1343-1356.

Nankhuni, Flora J. and Jill L. Findeis. 2004. Natural Resource-Collection Work and Children's Schooling in Malawi. *Agricultural Economics*, 31(2-3):123–134.

Nauges, Céline. Water Hauling and Girls' School Attendance: Some New Evidence from Ghana. *Environmental and Resource Economics* 66(1): 65-88.

UNESCO. 2017. Literacy Rates Continue to Rise from One Generation to the Next. Paris: UNESCO Institute for Statistics, UNESCO. Fact Sheet No. 45. Available: http://uis.unesco.org/sites/default/files/documents/fs45-literacy-rates-continue-rise-generation-to-next-en-2017.pdf. Accessed 16 March 2018.

WHO/UNICEF. 2015. *WHO/UNICEF Joint Monitoring Program for Water Supply and Sanitation*.

World Bank. 2017. *World Development Indicators 2016*. Washington, D.C.

Zhang, Jing. 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.

Figures and Tables

Figure 1: The location of our survey sites

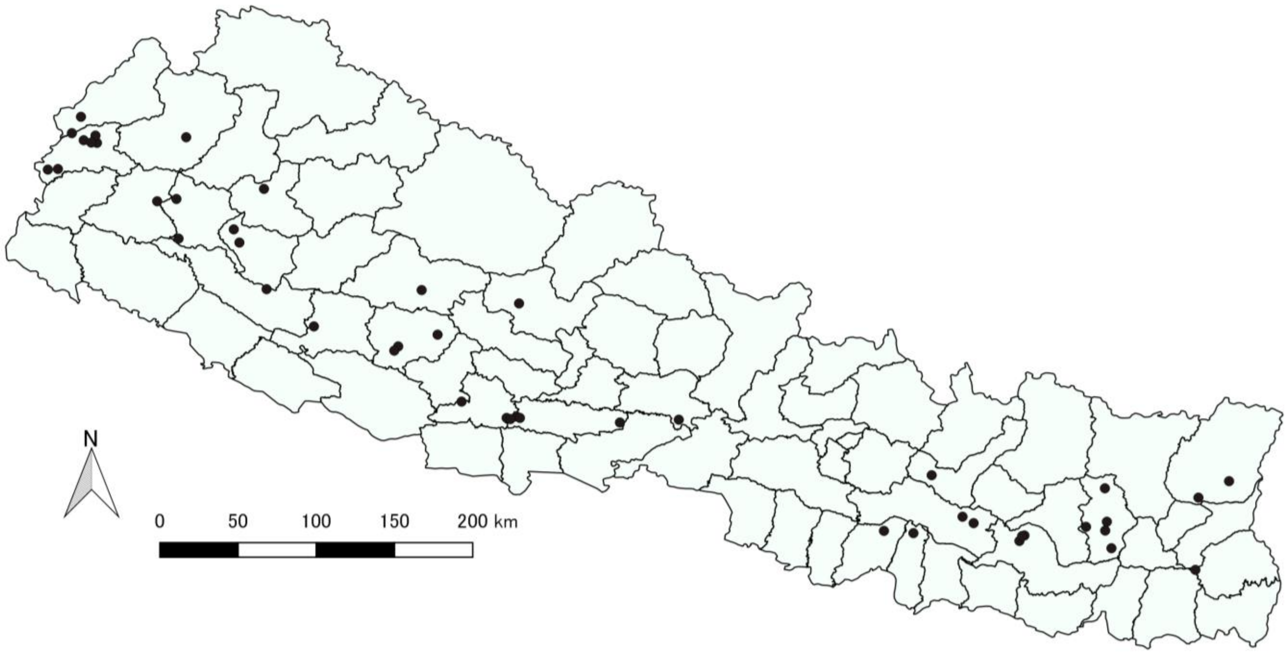

Table 1: Household access to water in the study region

A) Mode of water collection from primary and secondary water sources	
# of households who collect water from	
natural water source directly	1,653
private well	113
public well	454
private storage	73
public storage	1,012
other	10
B) Summary of water fetching activities (for households who carry water directly from the natural water source)	
Hours spent going to and from the source	0.523
# of water hauling trips per day	3.997
# of 15-liter bottles carried one time	1.355
Non-adult male (19 ≥ age ≥ 6) who engage in water fetching	0.111
Non-adult female (19 ≥ age ≥ 6) who engage in water fetching	0.210
Adult male (age ≥ 20) who engage in water fetching	0.190
Adult female (age ≥ 20) who engage in water fetching	0.756

Table 2: Summary statistics of main empirical variables

	NOBs	Mean	Std. dev.	Min	Max
A. Analysis for children aged 6 to 16					
Educational outcome					
Current (last attended) grade	2,482	4.431	2.779	0	12
No. of repetitions	2,498	0.097	0.322	0	3
Water accessibility					
Hours spent on water collection	2,498	0.525	0.393	0	5.967
Hours spent on water collection by neighbors	2,420	0.528	0.370	0	2.063
B. Analysis for children aged 14 to 16					
Educational outcome					
Completion of primary educ.	654	0.800	0.401	0	1
Completion of lower secondary educ.	654	0.321	0.467	0	1
Water accessibility					
Hours spent on water collection	654	0.514	0.370	0	2.083
Hours spent on water collection by neighbors	631	0.523	0.359	0	2.000

Table 3: Balancing and correlation tests for the identification assumption

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Balancing Test					Correlation Test	
	HHs closer to natural water source than the average		HHs farther to natural water source than the average		(2) - (4)	Correlation with distance to natural water source	
	Obs.	Mean (Std. Dev.)	Obs.	Mean (Std. Dev.)	Diff. [Std. Err.]	Obs.	Coef. {P-value}
HH total income (in 10,000 NPR)	860	10.046	714	11.076	-1.030	1,574	0.007
		(13.649)		(13.969)	[0.698]		{0.780}
HH size	896	4.711	757	4.749	-0.038	1,653	-0.013
		(2.156)		(2.419)	[0.113]		{0.592}
HH head completed primary educ.	896	0.228	757	0.210	0.018	1,653	-0.022
		(0.420)		(0.408)	[0.020]		{0.364}
HH head completed lower sec. educ.	896	0.089	757	0.106	-0.016	1,653	-0.008
		(0.285)		(0.308)	[0.015]		{0.761}
HH head completed sec. educ.	896	0.027	757	0.037	-0.010	1,653	0.007
		(0.162)		(0.189)	[0.009]		{0.793}
Age of HH head	886	44.143	751	45.254	-1.111	1,637	0.035
		(14.029)		(14.391)	[0.704]		{0.155}
House built within 5 years	891	0.164	753	0.163	0.001	1,644	-0.001
		(0.370)		(0.370)	[0.018]		{0.984}
House built within 10 years	891	0.347	753	0.335	0.012	1,644	0.022
		(0.476)		(0.472)	[0.023]		{0.378}
House built within 20 years	891	0.635	753	0.632	0.003	1,644	-0.011
		(0.482)		(0.483)	[0.024]		{0.652}
House built within 30 years	891	0.762	753	0.776	-0.013	1,644	-0.003
		(0.426)		(0.417)	[0.021]		{0.890}
Commute time to primary	468	27.788	397	29.461	-1.672	865	-0.018
school (min.)		(18.158)		(16.411)	[1.186]		{0.596}
Commute time to lower	290	39.210	226	42.403	-3.192	516	0.063
secondary school (min.)		(22.131)		(26.260)	[2.132]		{0.153}

Table 4: Impact on educational outcomes: current (last attended) grade and repetitions

	(1)	(2)	(3)	(4)
	OLS		IV	
Dep. var.:	Current (last attended) grade	No. of repetitions	Current (last attended) grade	No. of repetitions
Hours spent on water collection				
× Boy	-0.105	0.066**	-0.118	0.100**
	[0.127]	[0.028]	[0.314]	[0.048]
× Girl	-0.041	0.012	-0.073	0.042
	[0.136]	[0.032]	[0.267]	[0.045]
Fixed effects				
Age	Yes	Yes	Yes	Yes
Gender	Yes	Yes	Yes	Yes
Language	Yes	Yes	Yes	Yes
Caste	Yes	Yes	Yes	Yes
Survey month	Yes	Yes	Yes	Yes
Community (Ward)	Yes	Yes	Yes	Yes
First-stage F statistic				
			113.19***	112.47***
			348.55***	344.79***
Observations	2,482	2,498	2,405	2,420
R-squared	0.709	0.139	0.709	0.142

Table 5: Impact on educational outcomes: completion of primary and lower secondary schools

	(1)	(2)	(1)	(2)
	OLS		IV	
Dep. var.:	Completion of primary school	Completion of lower secondary	Completion of primary school	Completion of lower secondary
Hours spent on water collection				
× Boy	-0.066	0.131	-0.127	0.002
	[0.084]	[0.124]	[0.088]	[0.151]
× Girl	-0.153**	-0.032	-0.241**	-0.090
	[0.069]	[0.080]	[0.110]	[0.099]
Fixed effects				
Age	Yes	Yes	Yes	Yes
Gender	Yes	Yes	Yes	Yes
Language	Yes	Yes	Yes	Yes
Caste	Yes	Yes	Yes	Yes
Survey month	Yes	Yes	Yes	Yes
Community (Ward)	Yes	Yes	Yes	Yes
First-stage F statistic				
			121.62***	121.62***
			104.40***	104.40***
Observations	654	654	631	631
R-squared	0.266	0.280	0.260	0.277

Table 6: Impact on household chores (IV regression)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dep. var.:	Water hauling		Child/elder/sick care		Cleaning		Laundry		All household chores	
A. Sample mean										
Boys (obs. = 1,221)	0.092		0.023		0.047		0.054		0.157	
Girls (obs. = 1,199)	0.128		0.068		0.146		0.163		0.265	
B. Coefficient estimates										
Hours spent on water collection										
× Boy	-0.003		0.041		0.069		0.046		0.135**	
	[0.030]		[0.025]		[0.053]		[0.037]		[0.063]	
× Boy aged 6-9		-0.064*		0.069*		0.065		0.075*		0.083
		[0.035]		[0.036]		[0.055]		[0.041]		[0.067]
× Boy aged 10-13		0.012		-0.000		0.051		0.027		0.134
		[0.049]		[0.029]		[0.065]		[0.050]		[0.085]
× Boy aged 14-16		0.056		0.043		0.089		0.019		0.227**
		[0.081]		[0.034]		[0.066]		[0.052]		[0.108]
× Girl	-0.022		0.030		0.054		-0.033		0.033	
	[0.033]		[0.023]		[0.051]		[0.036]		[0.059]	
× Girl aged 6-9		-0.032		0.046*		0.112**		0.091**		0.164**
		[0.047]		[0.026]		[0.050]		[0.039]		[0.081]
× Girl aged 10-13		-0.050		-0.007		-0.010		-0.082*		-0.073
		[0.040]		[0.029]		[0.063]		[0.045]		[0.074]
× Girl aged 14-16		0.023		0.066		0.080		-0.110		0.053
		[0.076]		[0.064]		[0.098]		[0.095]		[0.107]
Girl dummy	0.054**		0.053***		0.118**		0.156***		0.174***	
	[0.022]		[0.019]		[0.056]		[0.039]		[0.055]	
Girl aged 6-9		-0.011		0.042**		0.022		0.017		0.016
		[0.026]		[0.019]		[0.035]		[0.020]		[0.036]
Girl aged 10-13		0.072*		0.057*		0.147**		0.171***		0.242***
		[0.037]		[0.031]		[0.075]		[0.062]		[0.088]
Girl aged 14-16		0.119***		0.055		0.197**		0.312***		0.280***
		[0.045]		[0.037]		[0.081]		[0.070]		[0.079]
Observations	2,420	2,420	2,420	2,420	2,420	2,420	2,420	2,420	2,420	2,420
R-squared	0.190	0.194	0.127	0.128	0.209	0.217	0.245	0.266	0.279	0.283

Appendix I: Nepal Marginal Settlements Survey: Household 2014/15

Survey Outline

In this appendix, we describe our original survey, Nepal Marginal Settlements Survey: Household 2014/15 (Nepal MSS:H-2014/15), which was conducted in remote and isolated mountain villages. In conducting the main survey, we had to start by constructing a village database, which provides village-level information regarding basic utilities, because our target is rural villages without electricity or water supply system and the government has no such database (see next section in this appendix for the construction of the village database).

We selected 45 wards (villages) randomly from 1,146 wards with no electricity or water supply system in the village database. Because some errors were contained in the village database due to several data limitations as explained below, we were able to survey 31 wards from the original list of 45 wards. These wards were surveyed from October to December 2014. The remaining 14 wards had access to electricity or an improved water supply system by the time of the survey. These non-eligible 14 wards were replaced by another 14 wards randomly chosen from the list, and they were surveyed from February to July 2015. The reason why the "second phase" of the main survey became prolonged is that most of the target villages are located in extremely remote areas. Hence, it took a long time and we incurred high costs confirming through visits that the selected villages (wards) were actually eligible. In addition to the replacement and confirmation tasks, several other factors prolonged the survey period. These factors are: (1) because our survey team is small, we could not implement the survey at one time and each team had to visit several different villages; (2) the survey had to be suspended for two months during the winter season due to snow and ice cover as most of our target villages are in high mountain terrains; and (3) the great earthquake of Nepal occurred in April 2015, which damaged our target villages and important access roads to our target villages.

Data Quality Management

Survey investigators were local NGO staff who had worked with AEPC for several years. Before the main survey, we conducted three pilot surveys in 2014 to train the investigators. The first pilot survey was conducted in Nagadaha Village District Center (VDC) in Ramechhap District in March, the second was conducted in Chatrebajh VDC and Puranogaun VDC in Kavrepalanchok District in May, and the third was conducted in a village in Tanahu District in August. Through these training sessions, we examined 16 candidates, and 10 investigators were selected for the main survey. As a result of the pilot surveys, we revised the questionnaire several times.

The contracts with investigators were carefully designed to enhance data quality. Remuneration for the survey work consisted of a base salary based on a piecework rate (per household), and a bonus based on data quality in terms of the paucity of inconsistent answers and invalid blanks. The expected base salary in a day was established to be more than twice the average daily rate for an entry-level government officer. Furthermore, we stipulated explicitly in the contract form that the renewal of individual contracts was dependent up on performance.

Village Database on the Status of Electrification and Water Supply

To construct a village database, we utilized published data and unofficial government databases provided by AEPC, which included information on the water supply improvement projects of the Department of Water Supply and Sewage (DWSS), electrification status of primary schools from the Ministry of Education, and electrification status of Village Development Committees (VD Committees) from District Development Committees (DDC) of target districts.

Regarding electrification status, villages were classified as electrified if the villagers had access to the national grid. Out of all 75 districts in Nepal, villages in 12 districts—Banke, Bardiya, Bara, Dadeldhura, Dang, Dhanusa, Jhapa, Kailali, Kanchanpur, Kathmandu, Lalitpur, and Saptariare—most of which are located in southern plain land areas (Terai) or capital zones, have achieved almost 100% electricity coverage. Regarding villages in the remaining 63 districts, we collected electrification status of VD Committees from the DDCs of each district. Although the information from the DDCs is the best available information at the time it is gathered, it is not necessarily the most updated information. Although the information provided is in various forms, all VD Committees in 63 districts can be classified into at least three categories in terms of household electrification rate: fully electrified, partially electrified, and non-electrified. Therefore, we removed the fully electrified VD Committees from the list and kept the partially electrified and non-electrified VD Committees on the list. As a VD Committee typically consists of 9 wards in a rural part of Nepal, we also used a primary school database with electrification status because almost every village has a primary school. Because schools are the priority for electrification, we assumed that when a school in a village has no electricity facilities, the village is regarded as being non-electrified, and thus it was kept on the list.

Regarding the status of improved water supply, we used a government database that contains 36,417 water and sanitation projects in Nepal and classifies the condition of water supply systems into five categories: (1) well-functioning, (2) rehabilitation, (3) reconstruction, (4) major repair, and (5) minor repair. We classified villages as having a water supply system, i.e., water-supplied villages, if the conditions were (1) well-functioning or (5) minor repair.

As a result, 1,146 out of 22,545 wards (villages) were classified as eligible wards, i.e., villages without electricity and running water. It is important to note that even if a village is classified as an electrified and water-supplied village, not all households in the village enjoy electricity and water.