The Effects of Access to Clean Water and Sanitation on Household Healthcare Expenditure: Evidence from the Philippines

Tianrui Liu

tliu1@macalester.edu

Macalester College

Abstract

Lack of access to clean water and basic sanitation is an issue around the world, especially in developing countries like the Philippines. Insufficient access to clean water has contributed to an important share of the total burden of disease worldwide (Bartram, 2008). A large portion of the healthcare spending from people with lower socioeconomic status are induced by diseases that can be prevented by available access to safe water. If households are freed from the catastrophic health expenditure induced by preventable water-related disease, fewer people will be pushed below the poverty line, which decreases the portion of tax increment financing spent on poverty assistance programs, and allows governmental resources to be redirected to other areas. This paper studies the effects of water access on household healthcare expenditure in the Philippines. The study uses panel data and employs Ordinary Least Square, Fixed Effect, and Instrumental Variable specifications. The final result of this study shows that cleanliness of water and toilet facility are significantly correlated with healthcare expenditure. Therefore, access to clean water and sanitation can lessen the household's burden in health expenditure.

Keywords: Sanitation, Clean Water Access, Healthcare, Health Expenditure

I. Introduction

Lack of access to clean water and basic sanitation is an existing issue around the world, especially in developing countries. Globally, about 844 million people lack access to clean water, 263 million people have to spend over 30 minutes per round trip to collect water, and 2.3 billion people live without access to basic sanitation (Guppy and Anderson, 2017). Unfortunately, the issue of unsafe sanitation is largely limited to low and lower-middle income countries (Ritchie & Roser). In developing countries, over half of the primary schools don't have access to water and sanitation facilities (UNICEF and WHO, 2012). In the Philippines, there are about 7.5 million people without access to water supply facilities and 24 million without access to sanitation (The World Bank, 2015).

Research has directly linked poor water quality with health risks. Contamination by human or animal feces is the most regular and pervasive health risk associated with unsafe drinking water, which may cause waterborne diseases, such as cholera, dysentery, and diarrhea (Yongsi, 2010). Diarrhea is the second leading cause of death among children under the age of five, killing 2,195 children everyday worldwide (Centers for Disease Control and Prevention). And it's been proven that the risk of diarrheal illness can be significantly reduced by access of safe drinking water (Fewtrell et. al, 2005). According to Bartram (2010), around 2.4 million deaths around the world could be prevented annually if everyone had good sanitation and clean drinking water.

As health risks increase, out-of-pocket (OOP) healthcare payments dramatically rise, which can become catastrophic health expenditure for households in developing countries. On average, OOP payments for people in developing countries represent around 40% of health spending and pose a significant barrier to access and utilization of health services (WHO, 2017).

According to the WHO, about 150 million people suffer financial crisis annually in paying for their healthcare, while 100 million are pushed below the poverty line (WHO, 2010). High OOP spending has several negative effects on people with lower socioeconomic status. A high OOP level leads to increased mortality among the marginalized groups (Plumper and Neumayer, 2012), reduces the likelihood of school attendance (Capuno et al., 2009), and accounts for an increase in poverty (Gupta, 2009). However, a large portion of the healthcare spending from people with lower socioeconomic status are induced by diseases that can be prevented by available access to safe water (Memirie et. al, 2017; Sarker st. al, 2018). If households are freed from the catastrophic health expenditure induced by preventable water-related disease, fewer people will be pushed below the poverty line, which decreases the portion of tax increment financing spent on poverty assistance programs, and allows governmental resources to be redirected to other areas.

Previous research has studied the effects of intervention in hygiene, sanitation, water supply or quality. A well-implemented cost-effective intervention on poor conditions is able to reduce diarrhea disease prevalence by up to a third (Bartram, 2010). Eder, Schooley, et al. estimate the effect of the $24 Development Assistance Program (DAP) water and sanitation interventions in Bolivia, an initiative made to promote the construction of adequate infrastructure for community water supply and household sanitation. After the six-year period of this program, households in the intervention community were 30% more likely to give good ratings to their water infrastructure status. These communities were able to improve their health status through improvements in water infrastructure, thus reducing the overall health burden. Research by Nandi et al. (2016) indicates that, by scaling up the coverage of piped water and improved sanitation among Indian households to a near-universal 95% level, there will be an estimated saving of $357,788 in OOP diarrhea treatment expenditure. Kremer, et al. study the effects of improvement in source water quality by randomly evaluating a spring water protection project conducted by a nongovernmental organization (NGO) in Kenya. Their result shows that spring protection reduced fecal contamination of source water by approximately 66%, and reduced child diarrhea by a quarter.

Two similar previous studies provide good modeling techniques to frame this research. Nandi et al. (2016) estimate the health benefits (in terms of reduced cases of diarrheal) and economic benefits (measured by OOP expenditure) of scaling up the coverage of piped water and sanitation. They use a formula to assume the combined efficacy for households with access to water and sanitation that take into account the interdependency between the effects of water and sanitation on diarrheal diseases. Watson (2005) studies the effects of federal sanitation interventions on reduction in infant mortality. Because these interventions take place across different geographic locations in different years, this research uses year fixed effects and controls for time-variant characteristics correlated with health, such as the number of hospitals, population growth, and a county-group-specific linear time trend. The final results of the study show that in sample areas, sanitation interventions are responsible for about 16 percent of the overall decline in the Indian infant mortality rates (Watson, 2005). This term paper employs a similar strategy, where a fixed effect is used to control for region invariant and time invariant factors respectively.

This paper studies the effects of water access on household healthcare expenditure in Philippine. I analyze the relationship between households' access to improved water sources and the effects on their healthcare expenditure. An improvement in clean water access may lessen the burden on households in terms of decrease in healthcare expenditure or it may barely have any effect on healthcare expenditure.

This study brings a unique perspective. So far, there is no research that studies the direct effects of clean water or sanitation on household health expenditure in developing countries over time. Most of the economic research focuses on macroeconomic aspects of improved water sources and sanitation and fails to link these factors with household healthcare expenditure. Watson et. al study the cumulative effects of sanitations on households by aggregating up to county group effects. Even though Nandi et. al also study the impact of water and sanitation on household level, they only apply cross-sectional data, thus fail to account for invariant factors across time. This study focuses on a household- and community- level analysis of the effects of clean water access and sanitation on health expenditure using panel data. The research object of this study is households in the Philippines. This is because of two major reasons: first, the Philippines is a typical developing country that suffers from clean water deficiency; and second, DHS data in the Philippines contains all of the dependent and independent variables of interest across five inconsecutive years.

The following paper is organized as such: section II provides the economic theories behind the objective of this paper as well as presenting the empirical approach; section III describes the household-level data in the Philippines acquired by the Demographic and Health Surveys (DHS) program; section IV demonstrates a series of empirical specifications in additional to discussing the regression results; section V will conclude the paper while providing some policy implications.

II. <u>Theory</u>

A. Economic Theory

It has been well-established that several global diseases are associated with the lack of clean water, including diarrhea, malnutrition, and intestinal nematode infections (Prüss-Üstün A, Bos R, Gore F, Bartram J., 2008). Without improved sanitation – a facility that safely separates human waste from human contact – people have no choice but to use inadequate communal latrines or to practice open defecation, and exposed fecal matter will be transferred back into people's food and water resources, which makes people susceptible to disease transmission (UN-Water). As a result, access to clean water and sanitation should lead to better health outcomes.

The relationship between health outcomes and healthcare expenditures can be explained by several Health economics theories. Assuming that an individual's total amount of time in a day(θ) can be divided into productive time(T^P) and sick time(T^S), and they can dedicate their productive time into three parts: time to work(T^W), time to have fun(T^Z), and time to invest into their health(T^H), we get equation 1:

$$T^P = \theta - T^S = T^W + T^Z + T^H \qquad [1]$$

If we assume that their utility level is determined by their health(H_t) and some other consumption goods(Z_t) that contribute to leisure, we get equation 2:

$$U_t = U(H_t, Z_t) \qquad [2]$$

Where H is the level of health and Z is the representative composite good.

In any given period t, the health level(H_t) of the individual is determined by the health level from the last period(H_{t-1}) in addition to the time(T_t) and money(M_t) they spent on health. The level of leisure(Z_t) depends on the time(T_t) and market inputs(J_t) spent on composite goods in this period. These relationships are captured by equation 3 and 4.

$$H_t = H(H_{t-1}, T_t, M_t) \qquad [3]$$

$$Z_t = Z(T_t, J_t) \qquad [4]$$

Based on the Grossman model, when an individual's health is at a low level, devoting time and resources that contribute to increasing the health level will result in more available productive time (Bhattacgarya, 2013). Because the individual can spend more time both at work and at play, they will reach a higher utility level. However, this relationship isn't totally linear. The production possibility frontier(PPF) for H and Z looks like figure 1, where the dashed line is a what a normal PPF looks like. Notice that from point A to point C is the so-called "free-lunch zone": since the person's health is still low, they are on the steep portion of the ill-avoidance function, so an hour spent increasing health yields more than an hour reduction in sick time, thus contributing to higher utility level. However, when the individual's health level becomes substantially high, further increases in health will not produce enough extra productive time to offset the time he must dedicate to improving H.

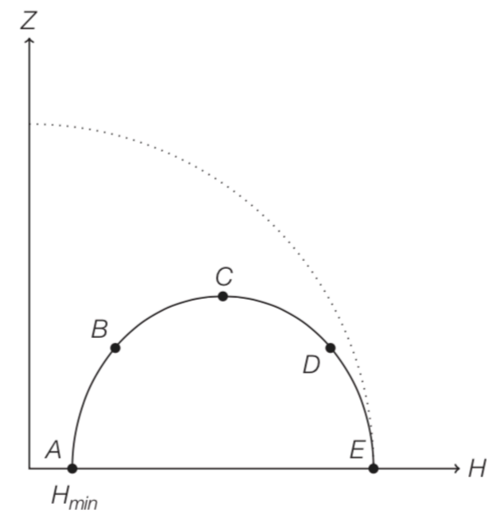

Figure 1: Production Possibility Frontier for H(Health) and Z(Leisure)

Based on the model, when the household's health level is low as a result of lack of access to clean water and sanitation, any extra money spent on improving health would boost up their utility level. This relation is represented by the area from point A to point C in the graph.

However, after their health level surpasses a certain point, any further increase in health expenditure would take up the resources that would have been spent on composite good and thus lower the utility level. This relation is represented by the area from point C to point E. Therefore, if person M who has better access to clean water lands on the curve BC whereas person N who lacks such an access lands on the curve AB, the theory suggests that person N would be to spend more health expenditure in order to reach their optimal utility level. If person M lands on the curve CD whereas person N lands on the curve AC, the theory suggests that person N would have to spend some health expenditure to reach their optimal level and person M shouldn't even bother to improve their health level (*H*). However, if both persons have sufficient access to clean water, thus landing on a high health level between point C and E, their expenditure shouldn't vary that much as both of them wouldn't bother to further improve their health. As such, the theories in combined suggest that, when access to clean water and sanitation is deficient, increase in it will lead to increased health outcomes and lower health expenditure; but after the access becomes sufficient, increase in it wouldn't make a huge difference on health expenditure.

B. Empirical Theory

Given the theoretical predictions, the models are as follows:

$$M_{it} = \beta_0 + \beta_1 Access_{it} + \beta_2 Controls_{it} + \beta_3 Municality_i + \beta_4 Year + \varepsilon_{it} \qquad [5]$$

$$M_{it} = \beta_0 + \beta_1 Cleanness + \beta_2 Controls_{it} + \beta_3 Municality_i + \beta_4 Year + \varepsilon_{it} \qquad [6]$$

$$M_{it} = \beta_0 + \beta_1 Sanitation_{it} + \beta_2 Controls_{it} + \beta_3 Municality_i + \beta_4 Year + \varepsilon_{it} \qquad [7]$$

Where t= 1993, 1998, 2003, 2008, 2013, and 2017, and i= 117 municipalities in the Philippines, the smallest geographical units used in the survey, and M is the healthcare expenditure. Because the water access points may not be randomly distributed - it is likely the case that richer people live with cleaner water sources - I put municipality and year as multiple dummy variables in this regression to account for endogeneity across time and across locations.

I use several variables to measure healthcare expenditure, access to water, cleanness of water, and sanitation. Total annual hospital bills are used to measure healthcare expenditure. Two dummy variables that measure if the household has toilet flush or toilet facility are used to measure sanitation. I use the average time spent to acquire water, and a dummy variable that describes if the household's water source is interrupted in the past month to measure access to water. I use a dummy variable that measures the household's source of water as an indication of how clean the water is. Controls include the number of family members in the household, the wealth level of the household, whether the household lives in urban or rural area, and highest year of education obtained by the household.

After breaking up the variables and adding control variables, the regressions become:

$$M_{it} = \beta_0 + \beta_1 Dtoilflush_{it} + \beta_2 Dtoilfacility_{it} + \beta_3 NumMember_{it} + \beta_4 WealthCombined_{it} + \beta_5 TypeResidence_{it} + \beta_6 HighestYearsEdu_{it} + \beta_7 Municality_i + \beta_8 Year + \varepsilon_{it} \quad [8]$$

$$M_{it} = \beta_0 + \beta_1 TimeAccessWater_{it} + \beta_2 NumMember_{it} + \beta_3 WealthCombined_{it} + \beta_4 TypeResidence_{it} + \beta_5 HighestYearsEdu_{it} + \beta_6 Municality_i + \beta_7 Year + \varepsilon_{it} \quad [9]$$

$$M_{it} = \beta_0 + \beta_1 Dcleandrinwater_{it} + \beta_2 NumMember_{it} + \beta_3 WealthCombined_{it} + \beta_4 TypeResidence_{it} + \beta_5 HighestYearsEdu_{it} + \beta_6 Municality_i + \beta_7 Year + \varepsilon_{it} \quad [10]$$

In equation 8, β_1 indicates the average increase in health expenditure if toilet flush is available; β_2 indicates the average increase in health expenditure if toilet facility is available; β_3 is the average increase in health expenditure as the number of family members increases by 1; β_4 is the average increase in health expenditure as the wealth combined increases by 1; β_5 is the average increase in health expenditure if the household lives in urban area; β_6 is the average increase in health expenditure as the household's highest year of education increases by 1; β_7 and β_8 represents differences in the intercept depending on where the household is located and year. Based on the economic theories, I expect β_1 and β_2 to have negative effects on the dependent variable, and β_3, β_4, β_5, β_6 to have positive effects on the dependent variable.

For equation 9 and 10, the control variables are the same as equation 8. β_1 in equation 9 is the average increase in health expenditure as the time to access water increases by 1 minute; β_1 in equation 10 indicates the average increase in health expenditure as the drinking water becomes clean. I expect β_1 in equation 9 to have positive impact on the dependent variable and the β_1 in equation 10 to have negative impact on the dependent variable.

Alternative specification of the regression model includes changing the control variables and including instrumental variables. Other control variables that likely have an impact on the dependent variable are: whether the household owns a car, whether the household feels like getting money for the treatment is a problem, and whether the household feels like the distance to a health facility is a problem. In order to further reduce the omitted variable bias, alternative specification also includes an instrumental variable that impacts the dependent variable through its correlation with independent variables in the primary regression. The instrumental variable is a dummy that measures if the household owns either a car, a motorcycle, or a bicycle. The instrumental variable doesn't have any impact on the health expenditure itself; however, it definitely impacts the time spent accessing water.

Several issues may arise in my empirical approach. Because I use the fixed-effects model that accounts for municipal level time invariant factors and time level municipal invariant factors, the estimated coefficients cannot be biased because of omitted time-invariant or municipal-invariant characteristics. However, I may still encounter omitted variable biases - there are variables that vary across time or across locations that are not captured by the model. For instance, if the municipality implements a new policy that increases the residents' insurances OOP coverage, it will probably become a source of change in health expenditure. Also, there are things that vary across time and locations that possibly impact my variables, such as individual municipality laws. Simultaneous causality problem may also occur, as higher healthcare expenditure may also cause longer time to access water. It's also possible that the variance of health expenditure is not consistent across different time to access water. Therefore, heteroskedasticity may occur.

III. Data Description

The datasets used for this paper, "Philippines - National Demographic and Health Survey", was acquired from the Demographic and Health Surveys (DHS) program. The datasets contain results from standard DHS surveys in Philippines in the years of 1993, 1998, 2003, 2008, 2013, and 2017. Survey areas of interest include alcohol consumption, health expenditures, health insurance, and household characteristics that measure access to clean water. The sampling scheme provides data representative of the country as a whole, for urban and rural areas separately, and for each of the country's administrative regions. The household selection process for the surveys is subject to random samplings. In situations where a housing unit contained one to three households, all households were interviewed. In 2017, a total of 31,791 households were selected for the sample, of which 27,496 were successfully interviewed. In other years, similar sample sizes of about 25,000 households in each year were selected to be interviewed.

Table 1, 2, 3, 4, and 5 in appendix A contain summary statistics for the variables of interest. Table 1 shows that there are several large outliers even after excluding the top 5% variables.

Figure 2, 3, and 4 in appendix A feature visualizations of the data.

Figure 2 shows a clear difference in health expenditure between household with and without toilet facility. However, the graph shows a counterintuitive result: households without toilet facility tends to have lower hospital bill. This is likely because households without toilet facility have more limited budget constraints and therefore lower hospital bill.

Figure 3 shows that people living in urban area generally spend less time accessing water. The two curves generally have the same pattern, though the density of households in urban area is more centralized on values equal to 0. This pattern is consistent with my expectation.

Figure 4 shows that, on average, households with higher wealth index spend more on healthcare, which is consistent with my expectation.

IV. Results

Results of the primary regression in figure 5 shows that almost all of the independent variable of interests as well as control variables are statistically significant with time and municipality fixed effects. However, the results of model 1 are problematic. They indicate that time to access water isn't statistically significant, and the signs of several variables in this model aren't consistent with my expectations. For instance, the model predicts that number of family members is negatively correlated with healthcare expenditure and living in urban area has a positive impact on healthcare expenditure. There are several possible explanations for these unexpected results. One possibility is that the outliers in the independent variable of interest is driving the coefficient to be odd. Also, notice that the observation number for this model is substantially lower than the other two models, indicating a low respond rate in this variable. Another possibility is that there is an omitted variable impacting both the time to access water and healthcare expenditure. If a municipal government launches a program that subsidize family with low wealth index, so that some households are able to spend more on hospital bills by using the money, this change is likely not captured by this model.

Model 2 and model 3 yield promising results. All of the variables, except for the shared toilet dummy variable, are statistically significant at 1% confidence level. The current exchange rate between Filipino peso and USD is 0.02. The average family income in the Philippines is about 22 thousand pesos monthly (Philippines Statistics Authority, 2020). These models predict that having a toilet flush decreases the household's healthcare expenditure by an average of 4230 pesos, or about 20% of the average monthly family income; having clean drink water decreases the household's healthcare expenditure by an average of 1393 pesos, or about 6% of the average monthly family income; as the household's family member increases by 1, healthcare expenditure increases by an average of 581 to 494 pesos in models 2 and 3 respectively; as the wealth index increases by 1, healthcare expenditure increases by an average of 4699 to 4791 pesos in models 2 and 3 respectively; one more year in highest year of education obtained by the household's members increases the healthcare expenditure by an average of 953 to 866 pesos. Notice that the coefficients of the same variables in these two models are approximately the same. The mean VIFs (Variance Inflation Factor) of all of the regressions are under 2, indicating that there isn't any multicollinearity. The Chi-square of all of the regressions are substantially small, indicating that heteroskedasticity is likely not an issue.

The first alternation specification in figure 6 changes the set of control variables in the primary regression model. It includes variables that measure whether the household has any vehicles, finds getting money for treatment or the distance to health facilities as a problem, as well as whether the household has pregnant member. The results of these specification is highly consistent with the primary regression model. The variables that are statistically significant in primary regression are still statistically significant in this alternative specification. The values of the coefficients are also close to the coefficients in the primary model.

Similar to primary regression, model 1 in this specification also yields problematic results. Time to access water is still not significantly correlated with the independent variable, and the signs of number of family members and type of residence are not consistent with my expectation. In order to improve the performance of model 1, an instrumental variable, whether the household owns a vehicle or a bicycle, is introduced to simulate time to access water. The variable “time to access water” is endogenous and thus is likely influenced by omitted variables. As figure 7 in appendix B shows, the instrumental variable is not significantly correlated with the dependent variable while it is significantly correlated with time to access water. Because the instrumental variable only affects the healthcare expenditure through time to access water, it helps reduce the omitted variable bias by reducing the endogeneity of the model. However, as model 4 shows, the time to access water is still not statistically significant, even though the sign of the coefficient becomes reasonable. There are several explanations for this problem. First, it’s likely that omitted variable bias still occurs as one instrumental variable cannot eliminate all. Second, the low observation number may cause higher standard error, thus making the variable of interest less significant.

Model 2 and 3 predict that having toilet flush decreases health expenditure by an average of 4510 pesos; having clean drink water decreases the household health expenditure by an average of 1427 pesos; as the number of family member increases by 1, the average health expenditure increases by 605 and 521; an increase in wealth index increases the health expenditure by an average of 4548 and 4670 pesos. All of these estimated vales are close to what the primary regression shows. Even though two of the added control variables, distance to health facility and currently pregnant member, are not statistically significant, because of how similar the results of this specification is to the primary regression, the primary regression survives the robustness check.

The control variable that measures whether the household finds it difficult to get money for the treatment is highly significant in all of the models. If the household has that problem, the models predict that the household would spend an average of about 1828 more pesos on total hospital bills. This is actually consistent with my expectation as families with lower socioeconomic status are also families who likely suffer from deficiencies in clean water. The economic theory would predict that these families have to spend more to improve their health level to a sufficient level that results in the largest utility. However, the wealth coefficient violates the theory, as rich people spend more on healthcare. This is probably because households with lower socioeconomic status have lower budget constraints. Therefore, even if they need to spend more on health expenditure, they are not able to do so.

In order to eliminate the problem caused by budget constraint, the primary regression is implemented three times on the poorest, medium, and richest group respectively. The results in figure 8, 9, and 10 in appendix B are highly consistent with the economic theory, as the independent variables of interest are all statistically significant in the poorest and medium group, except for the variable that measures if households have toilet flush. However, in the richest group, none of these variables are significant. This shows that within the lower-socioeconomic group, households suffering from deficiency in access to clean water and sanitation have to spend more to improve their health level; whereas in the richest group, the health expenditure is not correlated with access to clean water or sanitation at all.

V. <u>Conclusion</u>

In general, the empirical results verify the predications of the economic theories. Put simply, access to clean water and sanitation does affect the household level healthcare expenditure. As the primary regression suggests, when access to clean water and sanitation is deficient, household expenditure on health increases by an average of 5623 pesos in combined, which is 26% of the average monthly family income in the Philippines. The hypothesis is verified in instances where cleanness of water and toilet facility are significantly correlated with health expenditure. However, the impact of time to access water on health expenditure isn't completely clear, suggesting that further research has to be conducted to deem its significance.

Other limitations of this study include omitted variable bias, selection bias, simultaneous causality issues, and potentially imperfect regression specification. As discussed in previous sessions, omitted variable bias is likely a problem that exists in all of the regression models, given that randomized controlled trial (RCT) is not conducted for this study. These omitted variables include things that change across time and municipalities, such as individual laws of each municipality. Selection bias also likely occur due to a significant reduction in the sample sizes. The original data contains about 25,000 samples in each year; however, after merging datasets and dropping outliers and unrecorded values, the eventual sample sizes of my regressions are about 12,000. In each year, about 5% of the selected participants didn't respond to the interview, which is also a source for selection bias. It's also possible that some lower socioeconomic families live in areas that are hard for the interviewers to reach, such as mountainous areas, so that some of the variables are not recorded. Simultaneous causality is also a problem because healthcare expenditure can influence access to clean water and sanitation through omitted variables, which worsen the omitted variable bias. For instance, if a household cares more about their living standard, so that they are willing to spend nearly all of their income in improving their living standard, they likely spend more on both health and sanitation. Finally, the regression specifications used in this study may not be perfect. Some variables that better describe the dependent and independent variables, such as the self-paid medical bill and a dummy that describes if water source is interrupted in the last month, don't exist in all of the surveys across different years. These variables are not able to be included in the regressions.

The results of this study indicate that the Philippines government should put more efforts into improving water sources as it can significantly decrease household healthcare expenditure, which saves lower socioeconomic families from catastrophic health expenditure. As of 2018, over 26 million people of the 105 million Filipino population lived in poverty(Opportunity International, 2018). Despite substantial efforts from governmental anti-poverty programs, such as the Comprehensive Agrarian Reform, *Lingap Para sa Mahirap*, and the Social Reform Agenda, the Millennium Development Goal milestone of reduction in poverty has been a slow process (*Bayudan-Dacuycuy, 2014*). However, if the government could turn its attention into an ignored source of poverty - lack of access to clean water and sanitation - they will likely have better progress in poverty reduction.

Reference

Bartram, J., & Cairncross, S. (2010). Hygiene, sanitation, and water: Forgotten foundations of health. Plos Medicine, 7(11), 1-9. doi:10.1371/journal.pmed.1000367

Bayudan-Dacuycuy, Connie & Lim, Joseph. (2014). Chronic and Transient Poverty and Vulnerability to Poverty in the Philippines: Evidence Using a Simple Spells Approach. Social Indicators Research. 118. 10.1007/s11205-013-0409-5.

Centers for Disease Control and Prevention. Global Diarrhea Burden.

Eder, C., Schooley, J., Fullerton, J., & Murguia, J. (2012). Assessing impact and sustainability of health, water, and sanitation interventions in Bolivia six years post-project. Revista Panamericana De Salud Publica, 32(1), 43-48.

Fewtrell L, Kaufmann RB, Kay D, Enanoria W, Haller L, & Colforl JM Jr. (2005). Water, sanitation, and hygiene interventions to reduce diarrhoea in less developed countries: a systematic review and meta-analysis.

Guppy and Anderson (2015). Global Water Crisis: The Facts. United Nations University. Institute for Water, Environment and Health. Retrieved from: https://inweh.unu.edu/wp-content/uploads/2017/11/Global-Water-Crisis-The-Facts.pdf.

Gupta, I. 2009. Out-of-pocket expenditures and poverty: Estimates from NSS 61st round. Draft paper. New Delhi: Institute of Economic Growth.

Hutton G, Haller L (2004) Evaluation of the costs and benefits of water and sanitation improvements at the global level. Geneva: World Health Organization. WHO/SDE/WSH/ 04.04.

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

Memirie, S. T., Metaferia, Z. S., Norheim, O. F., Levin, C. E., Verguet, S., & Johansson, K. A. (2017). Household expenditures on pneumonia and diarrhoea treatment in Ethiopia: a facility-based study. BMJ global health, 2(1), e000166. https://doi.org/10.1136/ bmjgh-2016-000166

Nandi, Megiddo, Ashok, Verma, Laxminarayan. 2016. Reduced burden of childhood diarrheal diseases through increased access to water and sanitation in India: A modeling analysis. Social Science & Medicine, Volume 180, May 2017, Pages 193-195

Opportunity.org. 2018. Retrieved from https://opportunity.org/what-we-do/where-we-work/ philippines-facts-about-poverty

Plumper, T.and E. Neumayer (2012) Health spending, out-of-pocket contributions, and mortality rates. Public Administration 91(2):403–418.

Progress on Drinking Water and Sanitation: 2014 Update. (2014). WHO & UNICEF. Retrieved from http://www.unicef.org/gambia/Progress_on_drinking_water_and_sanitation_2014_update.pdf

Prüss-Üstün, A., Bos, R., Gore, F., & Bartram, J. (2008). Safer water, better health.

Ritchie Hannah, Roser Max. (2019). Clean Water. Published online at OurWorldInSata.org. Retrieved from: 'https://ourworldindata.org/water-access'

Sarker, A. R., Sultana, M., Mahumud, R. A., Ali, N., Huda, T. M., Salim Uzzaman, M., Haider, S., Rahman, H., Islam, Z., Khan, J., Van Der Meer, R., & Morton, A. (2018). Economic costs of hospitalized diarrheal disease in Bangladesh: a societal perspective. Global health research and policy, 3, 1.

Water Supply and Sanitation in the Philippines. The World Bank (2015). Retrieved from https://www.wsp.org/sites/wsp/files/publications/WSP-Philippines-WSS-Turning-Finance-into-Service-for-the-Future.pdf

Watson. 2006. Public Health Investments and the Infant Mortality Gap: Evidence from Federal Sanitation Interventions on U.S. Indian Reservations. Journal of Public Economics. September 2006, Pages 1537-1560.

World Health Organization and the International Bank for Reconstruction and Development / The World Bank. Tracking universal health coverage: 2017 global monitoring report. WHO and the World Bank. 2017. http://www.who.int/healthinfo/universal_health_coverage/report/2017/en. Accessed 11 Mar 2018. [Ref list]

World Health Organization and UNICEF Joint Monitoring Programme. (2017). Progress on Drinking Water and Sanitation, 2017 Update.

World Health Organization. (2010). Health systems financing: the path to universal coverage.

Yongsi H. B. (2010). Suffering for water, suffering from water: access to drinking-water and associated health risks in Cameroon. Journal of health, population, and nutrition, 28(5), 424–435. https://doi.org/10.3329/jhpn.v28i5.6150

Appendix A: Tables

Table 1: A detailed summary of time to get to water source.

	e(count)	e(sum_w)	e(mean)	e(Var)	e(sd)	e(min)	e(max)	e(sum)
TimeAccess~r	**53105**	**53105**	**12.78411**	**1004.283**	**31.69042**	**0**	**900**	**678900**

Table 2: A tabulated summary of the dummy variable that measure if the household considers distance to health facility as a problem .

DProb_dist_h	e(b)	e(pct)	e(cumpct)
big_problem	**37022**	**33.5557**	**33.5557**
not_a_big_~m	**73308**	**66.4443**	**100**
Total	**110330**	**100**	

Table 3: A tabulated summary of the dummy variable that measures if the household considers getting money for treatment as a problem.

DProb_treatm	e(b)	e(pct)	e(cumpct)
big_problem	**59889**	**54.2817**	**54.2817**
not_a_big_~m	**50441**	**45.7183**	**100**
Total	**110330**	**100**	

Table 4: A tabulated summary of the dummy that measures sanitation.

DSharedToile	e(b)	e(pct)	e(cumpct)
no	**76587**	**77.14398**	**77.14398**
yes	**22691**	**22.85602**	**100**
Total	**99278**	**100**	

Table 5: A tabulated summary of the household wealth index combined.

WealthCombin	e(b)	e(pct)	e(cumpct)
poorest	**35099**	**31.81274**	**31.81274**
poorer	**25865**	**23.44331**	**55.25605**
middle	**20117**	**18.23348**	**73.48953**
richer	**16121**	**14.61162**	**88.10115**
richest	**13128**	**11.89885**	**100**
Total	**110330**	**100**	

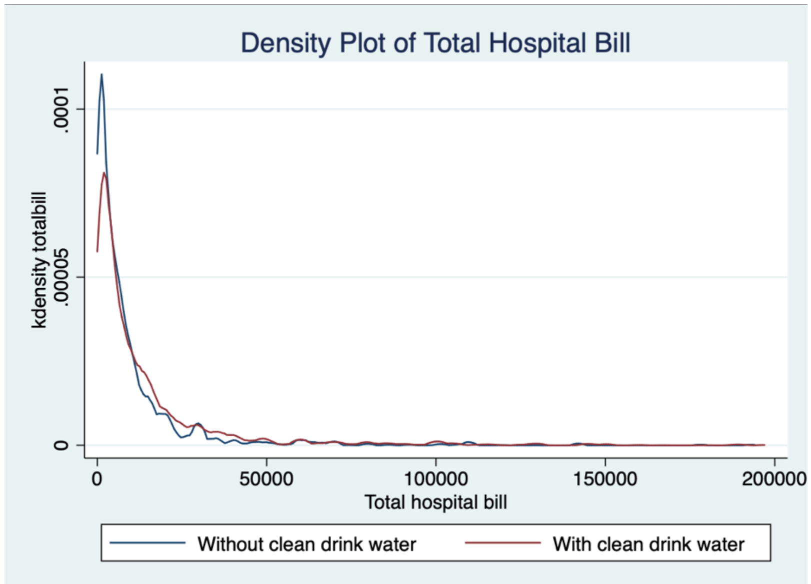

Figure 2: A density plot of total hospital bill with and without toilet facility.

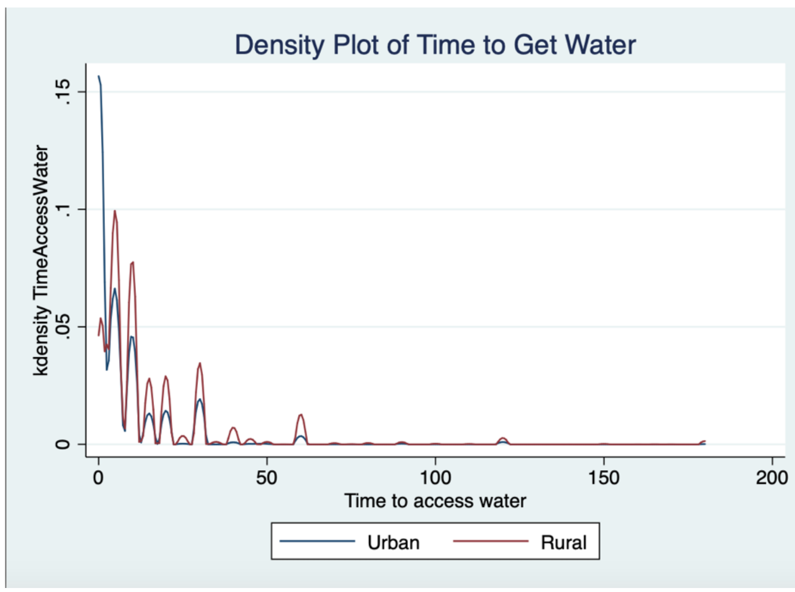

Figure 3: A density plot of total time to access water grouped by type of residence.

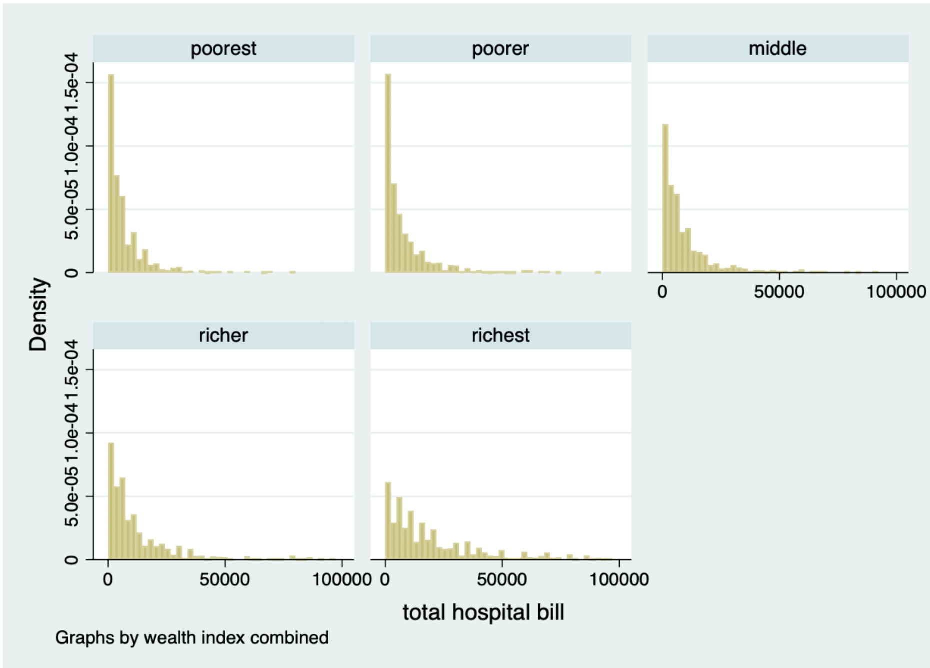

Figure 2: A density plot of total hospital bill by wealth index.

Appendix B: Regression outputs

Primary Regression

VARIABLES	(1) Access	(2) Sanitation	(3) Cleanness
Dtoilflush		-4,230***	
		(1,263)	
DSharedToilet		48.70	
		(785.0)	
Num_HHmembers	-137.5	581.0***	494.2***
	(123.3)	(120.9)	(110.5)
WealthCombined	4,448***	4,699***	4,791***
	(262.8)	(275.4)	(249.2)
TypeResidence	714.0	-1,511**	-1,660**
	(768.5)	(705.6)	(670.9)
HighestYearEdu	628.7***	953.2***	866.2***
	(138.6)	(125.2)	(116.0)
TimeAccessWater	-13.10		
	(13.79)		
Dcleandrinwater			-1,393***
			(674.1)
Constant	-2,142	519.5	-1,886
	(2,115)	(2,340)	(1,944)
Observations	5,783	11,181	12,120
R-squared	0.088	0.071	0.073

Figure 5: a table that shows the results of primary regression. Note that year and municipality dummy variables are manually deleted from the table in order to see the coefficients that matter.

Alternative Specifications

VARIABLES	(1) Access	(2) Sanitation	(3) Cleanness	(4) IV: transit
Dtoilflush		-4,510***		
		(1,274)		
DSharedToilet		-1.495		
		(785.3)		
Num_HHmembers	-99.50	605.6***	522.2***	-196.6
	(124.0)	(121.5)	(111.0)	(165.5)
WealthCombined	4,477***	4,547***	4,670***	5,325***
	(305.3)	(305.5)	(279.7)	(1,076)
TypeResidence	745.9	-1,372*	-1,563**	-3,030
	(776.3)	(714.8)	(678.0)	(4,488)
HighestYearEdu	607.4***	917.2***	832.1***	411.2
	(139.3)	(125.7)	(116.5)	(305.9)
Dtransit	-737.1	-251.5	-275.2	
	(721.0)	(699.3)	(647.2)	
DProb_treatmentmoney	1,940***	1,718**	1,832***	
	(681.7)	(673.4)	(628.8)	
DProb_dist_healtfacility	-333.5	903.9	450.8	
	(721.4)	(725.6)	(669.8)	
CurrentlyPregnant	42.41	-2,169	-918.2	
	(1,660)	(1,721)	(1,592)	
TimeAccessWater	-13.18			722.3
	(13.81)			(862.3)
Dcleandrinwater			-1,429**	
			(674.4)	
Constant	-4,592*	-3,016	-5,134**	-4,064
	(2,535)	(2,668)	(2,291)	(3,425)
Observations	5,783	11,181	12,120	5,783
R-squared	0.089	0.073	0.074	

Figure 6: a table that shows the results of alternative specifications. Note that year and municipality dummy variables are manually deleted from the table in order to see the coefficients that matter.

Instrumental Variable Specifications

VARIABLES	(1) if IV correlates with x	(2) if IV correlates with y
Dtransit	-3.733***	-249.8
	(0.341)	(647.0)
Num_HHmembers		526.6***
		(110.9)
WealthCombined		4,532***
		(267.4)
TypeResidence		-1,323**
		(660.8)
HighestYearEdu		823.6***
		(116.3)
DProb_treatmentmoney		1,960***
		(595.4)
Constant	15.02***	-5,630***
	(0.213)	(2,065)
Observations	34,021	12,120
R-squared	0.004	0.073

Figure 7: a table that tests if Dtransit is a good instrumental variable. Note that year and municipality dummy variables are manually deleted from the table in order to see the coefficients that matter.

Regression on the poorest group

VARIABLES	(1) Access	(2) Sanitation	(3) Cleanness
Dtoilflush		-734.3	
		(705.7)	
DSharedToilet		2,636***	
		(635.5)	
Num_HHmembers	74.74	398.6***	237.5**
	(132.7)	(122.1)	(98.24)
TypeResidence	-2,023**	-1,170	-188.7
	(988.4)	(849.8)	(702.4)
HighestYearEdu	379.0***	528.3***	507.6***
	(135.4)	(118.2)	(100.0)
TimeAccessWater	-17.15		
	(16.41)		
Dcleandrinwater			-1,636***
			(541.2)
Constant	6,679*	2.580	1,179
	(3,714)	(3,047)	(2,569)
Observations	1,956	2,598	3,186
R-squared	0.085	0.089	0.064

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 8: a table that shows the regression on poorest group. Note that year and municipality dummy variables are manually deleted from the table in order to see the coefficients that matter.

Regression on the medium group

VARIABLES	(1) Access	(2) Sanitation	(3) Cleanness
Dtoilflush		-11,628**	
		(4,678)	
DSharedToilet		-2,144*	
		(1,127)	
Num_HHmembers	-0.645	296.4*	328.1*
	(310.2)	(166.7)	(169.6)
TypeResidence	4,637**	-1,176	-2,113**
	(1,802)	(944.7)	(990.5)
HighestYearEdu	1,286***	696.5***	717.7***
	(359.9)	(191.2)	(194.7)
TimeAccessWater	-58.48*		
	(33.80)		
Dcleandrinwater			-1,261
			(1,018)
Constant	-2,877	20,368***	9,434***
	(4,094)	(5,314)	(2,605)
Observations	950	2,347	2,408
R-squared	0.070	0.042	0.040

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 9: a table that shows the regression on medium group. Note that year and municipality dummy variables are manually deleted from the table in order to see the coefficients that matter.

Regression on the richest group

VARIABLES	(1) Access	(2) Sanitation	(3) Cleanness
DSharedToilet		-8,053	
		(8,316)	
Num_HHmembers	1,421**	2,848***	2,899***
	(596.7)	(638.1)	(626.3)
TypeResidence	-8,503**	-9,603***	-9,671***
	(3,595)	(3,642)	(3,660)
HighestYearEdu	2,107**	1,542*	1,623*
	(820.6)	(849.1)	(831.8)
TimeAccessWater	-23.81		
	(53.78)		
Dcleandrinwater			-570.1
			(7,680)
Constant	19,876***	23,267***	22,756**
	(7,084)	(8,207)	(11,429)
Observations	852	1,657	1,681
R-squared	0.064	0.089	0.090

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 10: a table that shows the regression on richest group. Note that year and municipality dummy variables are manually deleted from the table in order to see the coefficients that matter.