Summer 7-14-2018

Water and healthcare access disparities: Impacts on health, wealth and education.

Justin Whetten

Justin Whetten

Candidate

Economics

Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Matías Fontenla, Chairperson

Kira Villa

David van der Goes

Jami Nunez

Water and healthcare access disparities: Impacts on health, wealth and education.

by

Justin Whetten

B.A. Economics, University of New Mexico, 2012
M.A. Economics, University of New Mexico, 2015

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Econoimcs

The University of New Mexico
Albuquerque, New Mexico

July, 2018

Water and healthcare access disparities: Impacts on health, wealth and education.

By

Justin Whetten

B.A., Economics, University of New Mexico, 2012
M.A., Economics, University of New Mexico, 2015
P.h.D., Economics, University of New Mexico, 2018

ABSTRACT

Rural communities in the United States and developing countries face a common problem of access. Lack of clean water or medical specialists can be solved by current technology, but there is a lack of resources and understanding of the problems. Rural communities in developing countries have a lack of access to clean water. Expansion of current water infrastructure and sanitation facilities could be done, but these are large costly projects. The lack of clean water, however, has dire negative effects on child mortality, household morbidity and overall household earning potential. In the United States, rural communities face a similar problem with access to a medical specialist. These specialists are in high demand and limited quantity, and a rural hospital does not have the required resources to justify the employment of said specialist. The lack of these specialists leads to higher overall medical costs and much worse patient health outcomes. This dissertation investigates the potential positive effect when these disparities are reduced or removed. We will use a combination of methods (instrumental variables, difference and difference, synthetic control, etc..) to look at how access to clean water can improve household wealth ware indicators. We will compare two different modeling approaches to determine the best methods for modeling cost and health outcomes from reducing access to care disparities for stroke patients. We find in all instances that a reduction in disparities leads to better outcomes for the individual, household, and community as a whole.

Table of Contents

Chapter 1

A Markov Model to Estimate Stroke Morbidity, Mortality, and Cost

Advances in the medical treatment of stroke and current technology, as well as the systems for delivering care, have caused stroke mortality rates to decrease (Lackland et al., 2014). However, the increasing proportion of stroke survivors is associated with a growing number of severely disabled patients. These patients are much more likely to have persistent motor symptoms. This affects their everyday life as they have a lower functional independence (Nichols-Larsen et al., 2005). As a result, stroke is still the leading cause of long-term disability (Go et al., 2013). Despite stroke being a significant burden on society and health care systems, only a few studies have evaluated the total cost of stroke from a social perspective. The average cost during the first year varies widely among studies, ranging from $2,860 to $43,652, due to differences in methodology, resources used for patients, and unit costs between countries (Carod-Artal et al., 1998; Spieler et al., 2004; Hervas-Angulo et al., 2006; Navarrete-Navarro et al., 2007; Fattore et al., 2012; Persson et al., 2012; Morris et al., 2016). Furthermore, few studies include all types of costs. There is a need to quantify the costs and address routine clinical practices (specialized neurological care, stroke units, reperfusion treatments, etc.) when modeling comprehensive stroke care. This study estimates the cost of comprehensive stroke and quantifies health outcomes for current clinical practices. Using this approach grants better understanding of the current level of care and allows for a complete comparison of emerging practices and treatments.

Previous studies either do not estimate morbidity, mortality, life/care duration or cost in one package or use very simple cross-sectional algorithms. Current models neither explicitly estimate care duration nor use time as an independent predictor. However, quality of health is a significant predictor of medical attention (Hawton et al., 2013; Hallberg et al., 2016; MacDougall et al., 2008) and cost (Jäkel et al., 2013). A Markov model bridges this gap. Markov models are a series of discrete, mutually exclusive events. These events called Markov States represent noteworthy phases of an illness (Briggs and Sculpher, 1998). Illness progression is modeled by transitions between relevant health states (Sonnenberg and Beck, 1993). For a current health state, the likelihood of the next event in illness progression is the transition probability of moving from the current health state to another. Transitions occur at equal fixed intervals (Markov cycles); only one transition occurs per cycle. Markov methods are suited to health care evaluation because they explicitly represent the timing of noteworthy events (Sonnenberg and Beck, 1993; Gandjour and Stock, 2007; Kirsch, 2015). Markov models predicting prevalence, duration, morbidity, and end points of illnesses are used to estimate the cost-effectiveness of nationwide programs (Gandjour and Stock, 2007; Kirsch, 2015), budgetary impact of new health treatments (Mar et al., 2008), and future morbidity and mortality costs (Gallet, 2017; Müllhaupt, 2015; Wong et al., 2000). We propose a Markov model with two methods to explicitly predict morbidity, mortality, and resulting public health costs of acute ischemic stroke (AIS) to an affected community. The affected community combines numerous perspectives, including individual patients and their families; society; third-party payers; federal, state and local governments; and health providers. The marginal method explicitly estimates temporal stroke behavior and calculates separate time-dependent costs for each patient. The aggregate method predicts ultimate disease morbidity and mortality for patients and calculates average patient costs. We estimate morbidity, mortality, and cost variability with Monte Carlo simulations.

METHODS

Markov model

Stroke morbidity is modeled and portrayed by a Markov Chain with five states (m): three transitive, one absorbing, and one tunnel (see Figure 1). The model starts with all patients in the stroke tunnel state (m=0). Depending on the outcome of the stroke, they progress to one of the transitive states (m=1-3) or the absorbing state, which is death (m=4). Transitive Markov states are in-home without third-party care (m=1), in-home with third-party care (m=2), and long-term care facility (m=3). Patients suffering a stroke with resulting minimal to no disability can be discharged to home without third-party care. Patients that had a stroke with resulting moderate disability require follow-up physician visits and rehabilitation. Long-term care would be for patients that had a stroke with resulting severe disability that requires constant nursing care.

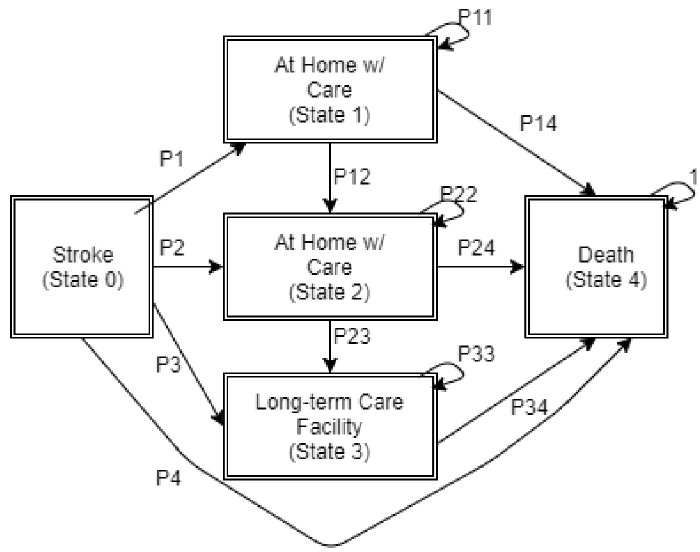

Figure 1. Markov chain for progression of stroke

Patients begin in State 0 and progress through the Markov chain until death. Progression follows one of seven mutually exclusive paths (k). Three paths travel through State 1: State 1 directly to State 4 (k=1), State 1 through State 2 to State 4 (k=2), and State 1 through States 2 and 3 to State 4 (k=3). From State 0 two paths pass through State 2: State 2 directly to State 4 (k=4), State 2 through State 3 to State 4 (k=5). From State 0 one path moves through State 3 and passes directly to State 4 (k=6). The last path travels directly from State 0 to State 4 (k=7). These are those who died from initial stroke complications. Each conceivable path k is assigned a sojourn probability (P_k). The sojourn probability is a function of the probabilities of state changes during a single cycle for the states that lie along pathway k. The resulting sojourn probabilities and transition probability matrix are shown in Table I.

Table I. Transition and path probability definitions

		Transition probabilities p_{ij}				Path probabilities P_k	
		Ending state				Path k	Probability
		[1]	[2]	[3]	[4]	1	$\frac{p_1 p_{14}}{(1-p_{11})}$
Starting State	[1]	p_{11}	p_{12}	0	p_{14}	2	$\frac{p_{12} p_{24}}{(1-p_{11})(1-p_{22})}$
	[2]	0	p_{22}	p_{23}	p_{24}	3	$\frac{p_{12} p_{23} p_{34}}{(1-p_{11})(1-p_{22})(1-p_{33})}$
	[3]	0	0	p_{33}	p_{34}	4	$\frac{p_2 p_{24}}{(1-p_{22})}$
	[4]	0	0	0	1	5	$\frac{p_2 p_{23} p_{34}}{(1-p_{22})(1-p_{33})}$
						6	$\frac{p_3 p_{34}}{(1-p_{33})}$
						7	p_4

Total patient number (N) is a model input. For potential/future study areas, N is estimated with existing incidence/prevalence models. For past studies, N is the reported number of patients. The expected number of patients following pathway k is found from equation 1.

$$E\,[N|k] = \bar{N}_k = NP_k \tag{1}$$

NP_k gives us the expected number of patients for each path, resulting in knowing their morbidity and mortality. By summing the per-patient cost of pathway k (C_k) over the six possible pathways, the expected total cost (TC) is

$$E[TC|N] = \sum_{k=1}^{7} \bar{N}_k \bar{C}_k = N \sum_{k=1}^{7} P_k \bar{C}_k \tag{2}$$

C_k is the total cost of all health services and time spent ill. These costs are obtained from published national survey data, societal economic cost models, and previous stroke cost estimates. Because all paths lead to death, the added cost of loss of life is not added. Cost descriptions for each path are shown in Table II. Medical costs for State 1 (S_1) are expenses paid by the patient out-of-pocket. Medical costs for States 2 and 3 (P_0, P_1, H_0, and H_1) are per-patient costs without a definite payer; the model does not estimate payer status.

Table II. Path cost definitions

Path k	States on path k	Cost Along path k (C_k)
1	1,4	$C_1 = T_1\ S_1$
2	1,2,4	$C_2 = T_1\ S_1 + P_0 + T_2\ P_1$
3	1,2,3,4	$C_3 = T_1\ S_1 + P_0 + T_2\ P_1 + H_0 + T_3\ H_1$
4	2,4	$C_4 = P_0 + T_2\ P_1$
5	2,3,4	$C_5 = P_0 + T_2\ P_1 + H_0 + T_3\ H_1$
6	3,4	$C_6 = H_0 + T_3\ H_1$
7	4	$C_7 = O$

T_m, residence time in Markov state m (m=1, 2, or 3); S_1, out-of-pocket cost of in-home care; H_0, minimum per patient third-party care cost of in-home care; H_1, cost of long-term care facility; P_0, per-patient cost of an inpatient rehabilitation; P_1, per-patient cost of other rehabilitation treatments.

Equations (1) and (2) are executed using an aggregate method or a marginal method. The marginal method explicitly estimates care duration, morbidity, mortality, and cost for individual patients. Non-zero recycling transition probabilities for transitive states ($p_{11}>0;p_{22}>0;p_{33}>0$) represent a patient neither improving nor worsening over one Markov cycle. The number of cycles (state residence time T_m) estimates years of life spent in a given health state and medical care utilized when cycle length is 1 year. The per-patient path cost ($C_{n,k}$) is a function of the variable discrete costs of occupying Markov states on path k for a given quantity of cycles. Occupying a specific Markov state results in specific medical costs. These costs are either duration independent (physician visit) or duration-dependent costs (medication use, hospitalization, or caregiver's time). Resulting path cost formulas (C_k) are shown in Table II.

In the marginal method, expected cost depends on path sojourn probability (P_k) and state residence time (T_m) for each transitive state on path k. Predicted residing time $E[T_m]$ for transitive state m is

$$E[T_m] = \frac{1}{1-P_{mm}} \tag{3}$$

Expected path costs C_k are calculated using the appropriate expected residence time as shown in Table II. The aggregate method assumes that each patient experiences average progression along one of the k possible paths by compressing the entire time spent in a phase (T_1, T_2, or T_3 in the marginal method) into a single Markov cycle.

Recycling in the transitive states for the aggregate method is prohibited ($p_{11}=0$; $p_{22}=0$; $p_{33}=0$). Costs now represent average lumped cumulative expenses associated with a particular path along the Markov chain; the costs for paths k are explicitly defined and drawn from the literature.

Cost parameter estimates

Cost estimates and sources for the STARR program are in Table III; all costs were adjusted to 2015 US dollars (EPA 2016). Medical expenses in the marginal method (P_0, P_1, and S_1) were drawn from the average individual cost of medical treatment in stroke. Ambulance transportation is included as H_0.

Table III. Cost parameters used in simulating the STARR program

Stroke				
Aggregate method		Marginal method		
Parameter	Value	Parameter	Value	Source
C_1	$155,728	H_0	$5,320	Earnshaw (2009)
C_2	$242,858	H_1	$77,745	Earnshaw (2009)
C_3	$584,844	P_0	$21,688	Ramirez (2008)
C_4	$182,291	P_1	$10,941	Ramirez (2008)
C_5	$405,071	S_1	$300	Demaerschalk(2013)
C_6	$289,698			
C_7	$0			

Aggregate costs were generated from table 2 and using life expectancy from Nelson(2011).

Transition probability estimates

Average transition probabilities (q_{ij}) were estimated by a meta-analysis of previous stroke studies. PubMed, Medline, Econlit and Google Scholar were searched for articles addressing Stroke. Bibliographies were then hand searched. Non-peer reviewed publications were not included. Studies from 2005 to 2016 were used if they provided original numbers of total patients, hospitalizations, and deaths. The random-effects model outcomes were used because there was substantial between-study variability (Normand, 1999).

Thirty-eight studies were used; 30 describe stroke in the general population; 8 apply to patients in a long-term care facility. Ten were used to estimate q_{12}, fifteen to estimate q_{23}, and seventeen to estimate q_{34}. Because the sum of all state transition probabilities is one, the average probabilities of death are calculated as $1 - \sum_{j=1}^{j=4} q_{mj}$

These estimates all have dependent time-scales indicating the observation period. Cycle lengths in the aggregate method are average durations; transition probabilities do not need adjustment ($p_{ij}=q_{ij}$). The marginal method cycle length is 1 year, thus transition probabilities were time adjusted (Sonnenberg and Beck, 1993):

$$P_{ij} = 1 - exp\left[\frac{-q_{ij}}{l}\right] \tag{4}$$

Where:

p_{ij} = *time-adjusted transition probability.*

q_{ij} = *average transition probability,* $i \neq j$

l=overall observation period

Observation periods for individual studies were used when reported; otherwise, mean care duration was used as a surrogate. The overall observation period was the study-size weighted average of study specific observation periods. Only the transition probabilities for exiting transitive states could be extracted from the literature. Estimates of the probability of remaining intransitive state *m* were calculated as

$$1 - \sum_{j=1}^{j=4} p_{mj}, j \neq m.$$

Resulting transition probabilities are shown in Table IV.

Table IV. Transition probability estimates from the literature

Aggregate method		Marginal method		Both	
p11	0	p11	0.9749	p1	0.3105
p12	0.5140	p12	0.0129	p2	0.2438
p14	0.4860	p14	0.0122	p3	0.2484
p22	0	p22	0.8839	p4	0.1973
p23	0.6804	p23	0.079		
p24	0.3196	p24	0.0371		
p33	0	p33	.8823		
p34	1	p34	.1177		

Monte-Carlo simulations

Expected morbidity, mortality, and cost for N patients is determined by equations (1)–(3); Monte Carlo simulates results for individual patients; the resulting distributions show possible variability in morbidity, mortality, and cost. Monte Carlo simulations were conducted separately for each method. A number U was randomly chosen from a uniform distribution between 0 and 1, at the end of each cycle. If {m=1, 2, 3} and $U \leq p_{m4}$, the patient transitioned to State 4. If $p_{m4} < U \leq p_{m4} + p_{m(m+1)}$ and {m=1, 2}, the patient transitioned from state m to state m+1. Otherwise, the patient remained in the current state. Markov cycles were simulated until the patient transitioned to State 4. The path k taken and residence time in each transitive state was used to determine the morbidity, mortality, and resulting cost for each patient using the path costs in Table I. Ten thousand trials were conducted for each patient. These form distributions of possible medical care, care duration, and cost. Equations (1)–(3) correspond to the means of the distributions of patients along path k, state duration, and cost (Sonnenberg and Beck, 1993).

Statistical tests

Pearson chi-squared goodness-of-fit tests were used to test method similarity and model fit. One-tailed bioequivalence t-tests were used to test if the Monte Carlo results were equivalent by comparing distribution means (Berger and Hsu, 1996; McBride, 1999).

RESULTS

Equations (1)–(3) and Monte Carlo simulation results were compared with the reported morbidity of the 2015 STARR results for the 447 patients. Both methods were tested. The Markov model was then used to estimate morbidity, mortality, and cost.

STARR

The Stroke Telemedicine for Arizona Rural Residents (STARR) network was a randomized control trial consisting of a 1-hub, 4-spoke Telestroke system. The study was conducted by the Mayo clinic from 2008 to 2010 and included 447 patients. The studies primarily looked at the number of patients who received thrombolytic medication and the time to treatment in patients evaluated by telemedicine. The study also assessed the functional outcomes of acute stroke subjects by modified Rankin scale and assessed rate of intracranial hemorrhage post thrombolysis. Consequently, in the first 6 months of the program thrombolysis administration for qualified patients increased 10- to 20-fold from the participating spoke hospitals' past baseline (from roughly 0.5 to 1.0 per hospital per year to roughly 10 per hospital per year). (Demarschalk et al. 2012)

Simulations of the STARR cohort (N=447)

N=447 for Equations 1–3 and Monte Carlo simulations. Expected morbidity and mortality are shown in Table V. Monte Carlo distributions of patients traveling on paths 1-6 are shown in Figure 2. No significant difference was found between predicted and reported morbidity and mortality (p-values=0.7212-0.933) for either method. The marginal method had a lower Chi-square value and a higher p-value, making it a closer fit to the real data, compared to the aggregate method. While no significant difference was found between the aggregate and marginal methods (p-value=0.173) the low p-value suggests at least some differences between the two methods. Figure 2 shows that patients have a slightly higher chance of taking paths 1-3 in the marginal method than in the aggregate. This corresponds with the longer life and higher costs that we find in the predicted costs in Table VI.

Table V. Comparison of model results to reported STARR

	Path 1	Path 2	Path 3	Path 4	Path 5	Path6	Path 7	Chi-square	pvalue
STARR	76	22	47	36	74	104	88		
Aggregate	68	19.1	40.6	36.9	78.5	114.6	89.3	3.67	0.7212
Marginal	77.0	23.9	50.8	36.3	77.3	93.2	88.5	1.85	0.933

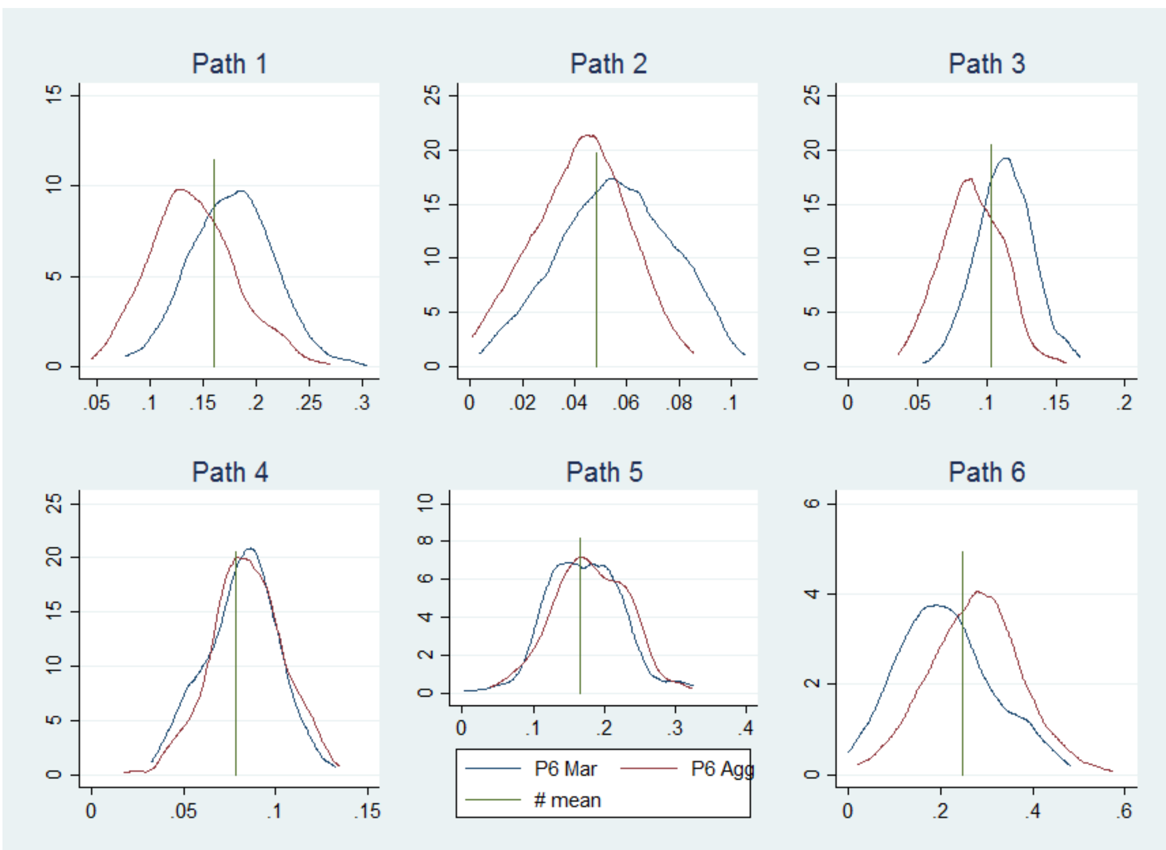

Figure 2. STARR morbidity/mortality distributions (447 patients, 10000 simulations)

Predicted outcome costs are shown in Table VI; probability distributions of stroke cost are in Figure 3. Costs are highly impacted by the length of life because as patients get older their medical costs increase exponentially. Method cost distributions were not equivalent (p>0.9999). There is an $89,347 difference in cost per patient from the two methods. Average predicted life expectancy post stroke was 5.3 years (3.87 QALYs; range=2.63–5.06 QALYs) for aggregate and 7.82 years (5.71 QALYs; range=3.93–8.30 QALYs) for marginal. The durability of health states and mortalities showed in Figure 4.

Table VI. Predicted costs for STARR cohort (N=447)

	Aggregate method	Marginal method
Expected cost (Equation 2)	$160,589	$249,936
Monte Carlo mean cost (10000 simulations)	$160,398	$249,878
Monte Carlo median cost (10000 simulations)	$160,061	$249,746

Cost of the entire STARR cohort (N=447)

Table V shows the patient distribution and how many took each path. The aggregate method predicted a total cost of $73.78m or $160,589 per patient. The marginal method predicted a total cost of $111.72m or $249,936 per patient.

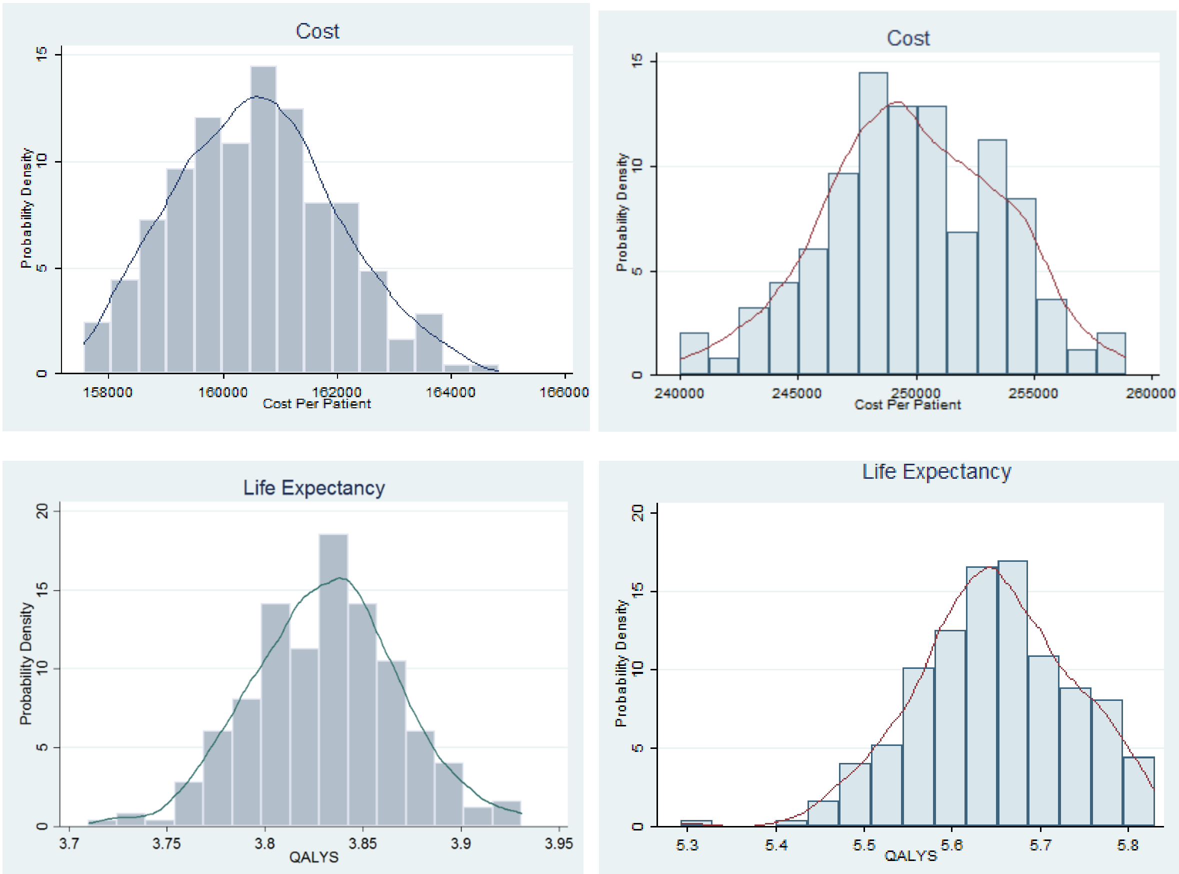

Figure 3.STARR Cost and Life Expectancy distributions (447 patients, 10000 simulations)

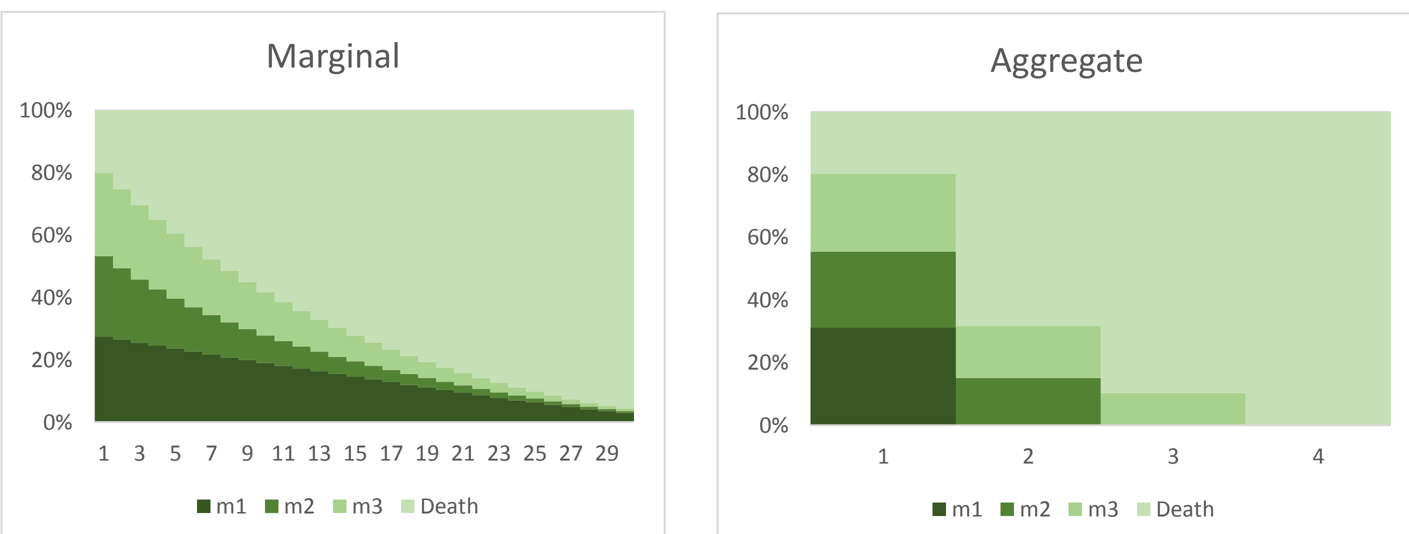

Figure 4. Durability of the Health States and Mortalities (447 patients, 10000 simulations)

These outcomes are consistent with the simulation results. Patients have a 17.8% probability in the marginal method and 13.9% probability in the aggregate method of taking path 1. Patients have a 20.4% probability to take path 6 by the marginal method and 26.7% probability by the aggregate method.

DISCUSSION

We developed a Markov model to predict morbidity, mortality, and resulting cost burden of stroke to an affected community. The affected community represents an area that is directly affected by stroke (individuals, families, or towns) and ranges to an entire nation (national stroke burden). The marginal method is a classic Markov model that estimates individual stroke behavior for every patient; the model explicitly estimates life expectancy and calculates separate time-dependent costs. The aggregate method is a Markov equivalent of existing linear combination models; it compresses the entire time spent in a health state into a single Markov cycle, estimates morbidity and mortality, and then assumes that each patient has mean illness performance for a particular morbidity or mortality. Models that do consider time (marginal method) and models that do not (aggregate method, existing linear combination models) can now be compared.

Both methods were tested against the STARR study and accurately predicted reported morbidity and mortality. Although not significantly different, the marginal method has a better fit for reported morbidity and mortality than the aggregate method (p-value: 0.933 versus 0.7212). The aggregate method may better estimate cost where specific medical care durations are known and so the variance is low. The differences in morbidity, mortality, and cost predicted by the two methods are the result of two distinct analytical approaches to the problem. The aggregate method assumes morbidity, mortality and costs depend only on the total number of cases. The marginal method assumes that morbidity, mortality, and cost depend on both the total number of cases and individual illness duration. This time-dependence assumption causes a higher sojourn probability for paths $k=1$, 2, and 3, and lower sojourn probabilities for paths $k=4$, 5, and 6, resulting in longer life expectancy and higher costs. Table VI shows a significant difference between marginal and aggregate cost estimates. This difference is caused by the cost of care resulting from the greater life expectancy. This difference in life expectancy is directly proportional to the difference in cost. Running the aggregate model with the mean life expectancy of the marginal model gives us nearly the same cost ($237,000 vs $249,000).

Long-term health effects of a stroke that may manifest months to years after the primary incident are not addressed directly. The model assumes the likelihood of any health outcome is identical for all patients. Several medical conditions are acknowledged as risk factors for physician visits (Hawton et al., 2013; Hallberg et al., 2016; MacDougall et al., 2008), hospitalization, and death (Jäkel et al., 2013; Kourlaba et al., 2012). However, these modifying conditions require pre-existing knowledge of their prevalence in the population; that information may be difficult to obtain. The implementation of this Markov model assumes that transition probabilities are constant for the duration of any health state. Physician visit and hospitalization patterns are not temporally fixed. Existing case reports and national burden of illness estimates do not provide enough information to determine how these change with time over the duration of the life of the patient. To estimate time-varying transition probabilities, future studies need to explicitly state when each patient visited a physician, was hospitalized, or died. Future versions of the Markov models can be improved by incorporating age-specific illness transition probabilities, greater symptom specificity, and greater flexibility in transitioning between stages.

Despite these limitations, our model has multiple applications. The aggregate method is an expedient simplification of the marginal method that sacrifices individual differences and temporal dependence to achieve computational efficiency. The aggregate method should only be used to provide a Markov equivalent to existing linear combination models, allowing statistical comparison of time-dependent models with prevailing linear total case models.

The marginal method has two major uses: (i) outcome modeling (using a historical or predicted number of patients); and (ii) national stroke burden estimates (using the estimated national stroke case total). For both, the marginal method provides more accurate estimates of morbidity and mortality than existing methods. The marginal method implementation of Equation (2) is ideal for the Centers for Disease Control and Prevention (CDC) Chronic Disease Cost calculator; users specify a given number of cases and receive morbidity, mortality, and economic cost estimates. The marginal method can also estimate time-dependent stroke behavior, such as when patients in historical cases sought medical care or when patients in future cases are expected to arrive at healthcare facilities. Officials looking to justify changes in stroke care, hospital procedure, and standard practice of care can use the model to calculate the costs of stroke under the current system and the reduction in morbidity, mortality, and cost under the suggested changes. Reduction in average costs, changes in estimated case severity, or reduced worst-case outcome probability can be presented as benefits of and justification for the change. This four-state, seven-path Markov framework is very flexible. It can be applied to any disease where morbidity is described by the medical care sought by using disease-specific transition probabilities and cost parameters.

Patient stroke outcomes are highly variable. Some result in minimal to no disability, while others require serious medical care to treat and still may result in severe long-term disability. This Markov model provides estimates of both expected and distributions of possible morbidity, mortality, and cost for stroke cases. The advantage of Monte Carlo simulation over the predictive equations is that the distributions from simulation offer a wider range of conceivable outcomes and, consequently, provide cost estimates for the unlikely, but more extreme, circumstances. For potential cases, system/regulation changes and national burden of illness estimates, mortality and mortality-associated costs (or the reductions in potential mortality) are more significant to decision makers than morbidity and morbidity-related costs. Thus, Equation (2) or the average cost from Monte Carlo simulation may be the best estimate of cost precisely because it is more sensitive to the influence of uncommon, but very expensive mortality.

In conclusion, this Markov model is a major improvement over the current methodology. Prevailing models for both estimating the total national and community-specific stroke burden do not consider care duration as a separate cost or morbidity predictor. Quality of health is a significant determinant of both pursuing further medical treatment (Hawton et al., 2013; Hallberg et al., 2016; MacDougall et al., 2008) and cost (Jäkel et al., 2013), thus ought to be included in any model that estimates health care costs. Although the difference between models was not significant in the small STARR cohort, current methods that do not explicitly model care duration and the time dependence of morbidity, mortality, and cost may significantly underestimate the number of hospitalizations, deaths, and economic burden of stroke.

Chapter 2

2005 Managua Water/Sanitation Access Expansions Effect on Income, Health, and Education: A Synthetic Control Method Analysis.

Background

Clean water is crucial for life. Yet, many impoverished people worldwide do not have access to proper sanitation and clean water. The use of unclean water sources is prevalent in low-income countries. In 2010 over 884 million peoples' primary source of drinking water was classified as unsafe (UNICEF 2010). The use of unsafe water and lack of access to proper sanitation has been linked to disease and sickness. In developing countries contaminated water results in thousands of deaths every day, mostly in children under five years of age (WHO 2015). The United Nation Development Project asserted that unsafe water and a shortage of basic sanitation caused 80% of diseases in the developing countries (UNDP 2014, Xuan-Long 2010). There are many studies on the connection between access to water/sanitation on health. Increased availability of clean water and sanitation reduces the incidence of water-related diseases in India (Dasgupta 2004). Piped water reduced diarrhea in children in rural India (Jalan and Ravallion 2003). The privatization of water services in Argentina resulted in greater access to clean water and sanitation, and as a result, reduced child mortality (Galiani *et al.* 2005). In Brazil, increased access to clean water and sanitation has led to a reduction in infant mortality (Macinko *et al.* 2005, Gamper-Rabindran *et al.* 2010). In Nicaragua, areas with limited access to clean water and sanitation have been linked to drops in tourism and increased illness (laVanchy 2017).

Nicaragua is a country rich in surface and underground water resources. However, water sources are being contaminated by a lack of treatment systems and wastewater runoff (WHO 2016). Deforestation and intensive land use also affect the recharge capacity of sources and aquifers. The territorial distribution of water resources in the country is uneven and large parts of the infrastructure are obsolete and need to be updated and improved. Nicaragua has one of the largest sources of renewable clean water for Latin American. However, Nicaragua also has the greatest economical water scarcity. This means that while Nicaragua has many clean sources of water, there is little infrastructure to get this water to the people. Neither, the rural or urban communities have the resources needed to invest in a large infrastructure project. As a result, Nicaragua does not have sufficient basic services to provide drinking water and sanitation to its population. Approximately 77% of households endure continuous water cutoffs and limited hours of service. Sanitary sewer coverage is only 35% and of the wastewater generated only 42% receive some type of treatment (DCP 2006). These precarious hygiene conditions represent the main cause of diarrheal diseases, especially among the most vulnerable groups, such as children under five years of age (WHOFS 2017). Over 1.4 million cases of diarrhea are reported in Nicaragua every year, with over 65% of children under five having had diarrhea at some point in the last year (UWHO 2017). Hence, the provision of an improved water supply is an important policy in Nicaragua. In 2005 the Managua government launched a project to improve access to clean water and sanitation.

2005 Investment program in drinking water and sanitation (PIAPS)

In 2005 Nicaragua started the Municipal Social Investment Program. For this program, they received $18.7 MM to further develop and update their water and sanitation systems. This program was expanded and renamed the investment program in drinking water and sanitation (PIAPS). With the expansion, Nicaragua received $37.9 MM more funding for water and sanitation. PIAPS was funded by a loan from the International Development Bank and local funding sources. These programs had three major components. First, was the updating, repair, and replacement of wells, pumping equipment, reserve tank, pipes, chlorination equipment, and macro/micro measurement programs complementary to those in progress. This was called the emergency plan component. The emergency plan component was for facilities and equipment that were indispensable to restoring the minimum service condition across the governmental region of Managua. The simple aim of this component was to reduce the number and duration of water service shortages. The second component, business strengthening, was to support and complement the activities of the business side of the water infrastructure. This paid for office buildings, equipment, communication systems, training of personnel, and management of human resources. The last and largest component was for the rehabilitation and optimization of potable water and sanitation systems. This component covers the expansion of drinking water distribution networks to low-income sectors, extensions in sewage networks, and improvement/extensions in wastewater treatment plants. The program also provided for low-income neighborhoods to connect to existing water and sewage networks. While investment in the Managua water and sanitation system continues, the PIAPS program was considered complete in 2009 (Fryer 2012).

The PIAPS program spent over $56.6 MM in Managua over four years. The GDP of Managua in 2005 was roughly $1.32 Billion (Factbook 2015), meaning that the program also could be seen as stimulates package. The package would have been 4% of the areas GDP. In comparison, the American Recovery and Reinvestment Act(ARA) of 2009 was about 1% of USA GDP(ACT 2011). It has been estimated that the ARA has lead to about 1.9 million jobs created over the last 9 years (CBO 2018). This is .059% of the population entering the workforce. Using the same rationale this would have lead to 8,326 jobs for Managua for a 1% stimulus in 2014, or 33,436 new jobs because of the larger size of the PIAPS in relation to Managua GDP. From the reports filed by the PIAPS project we know that the project led to 326,933 individuals getting access to piped water, 288,623 gaining sanitation (sewage) access and, 2,452 permanent jobs directly. The project could be seen to have affected but the supply and demand side of the labor market. The reduction in water fetching and increased health allowed for more time to be spent in wage labor. While the project created jobs both directly and indirectly.

This study aims to measure the impact of improved access to piped water and sanitation on several household welfare indicators, including off-farm income, education, and health, in Managua, using recent household surveys, and synthetic control estimators. The PIAPS program makes Managua a good case for studying the effects of improved water and sanitation access. First, Managua experienced a large increase of water/sanitation access from 1998 to 2014 (NHLSS 1998,2014). Improved water access increased from 75% to 92% and sanitation from 43% to 63% (Figure 1). This large increase will allow for good comparison analysts.

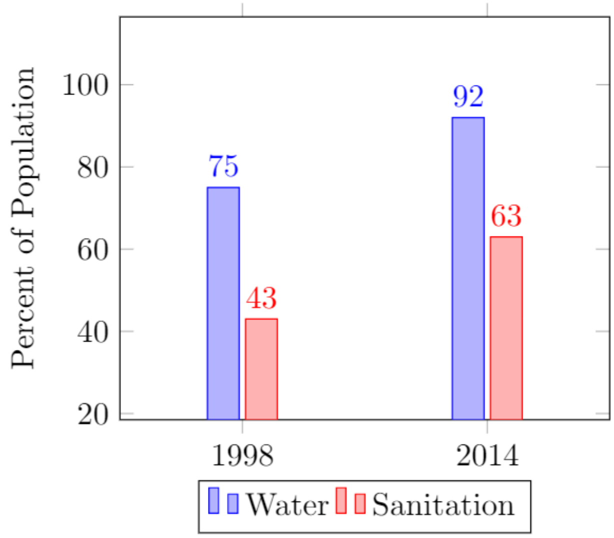

Figure 1: Clean Water and Sanitation Access. 1998 and 2014

During this same time, Nicaragua`s income per capita rose from $951 USD in 1998 to $1,975 USD in 2014 (World Bank 2017). We see from figure 2 Nicaragua seems to have experienced steady growth, and the 2009 world recession had little effect. This is believed to be due in part to foreign debt reduction, recovery of export demand and growth in its tourism industry (Factbook 2015). This paper aims to test what if any percent of this change in income is due to improved access to water and sanitation.

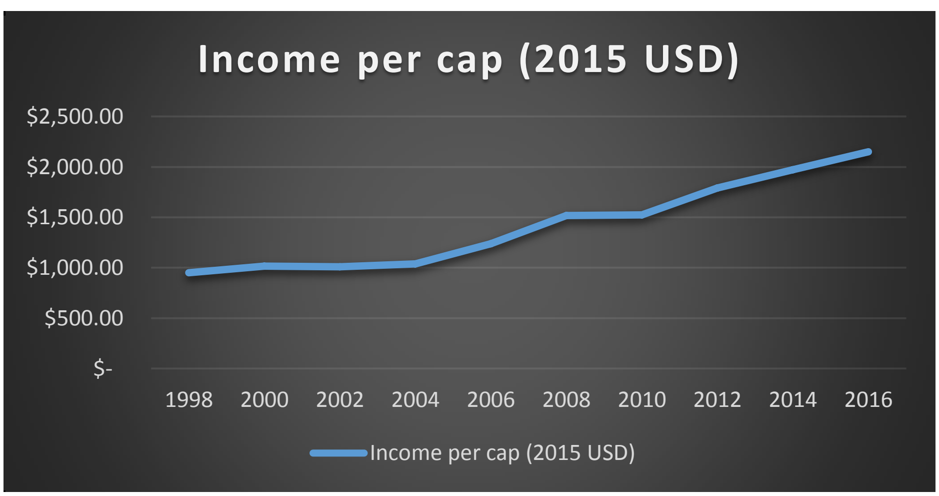

Figure 2: Nicaragua Income per capita. 1998 - 2014

Second, there are numerous studies on the effect of clean water and sanitation on health, yet only a few studies investigate the effect of clean water and sanitation on other household welfare indicators such as income and education. Waddington (2009) and Viet (2012) did show a positive effect from clean water and sanitation on labor supply and income. Their findings, however, were not statistically significant. In the long-term, clean water and sanitation can result in an increase in income through several channels. Unclean water can cause diseases, health problems and low labor force participation. The majority of households do not have piped water to their homes. Without this access to piped water, households have to use other water sources. The majority of these other sources are, usually, great distances from their home and require purification before use. Furthermore, although the adults may not themselves contract illness, they still must tend to any children that do become ill, causing them to lose wages. Thus, having piped water can save time and allow for activities that are more productive and increases to their income. In urban Morocco, Devoto et al. (2011) carried out an exceptional study, which examined the effect of water on labor and income. It was their conclusion that piped water could improve the households' leisure and social activities, but not necessarily their income or labor supply. McCauley (2015) however, did find that improved access to clean water led to greater female participation in parliament, attending school, and working for wages. McCauley does state, however, that the study cannot determine causality or fully address the endogeneity of water access. Several studies have focused on the quality of drinking water around the world (Hoang 1990, Le *et al.* 1993, Nguyen *et al.* 1994, Le and Munekage 2004, Agusa *et al.* 2006). Other studies have mentioned the adverse effects of unclean water on health, but have done so without quantitative evidence (Worid Bank 2000, 2004, Xuan-Long 2010, *Sue Khoe Newspaper* 2010). None have considered the socio-economic benefits of clean water and good sanitation on a quantitatively significant population. Third, there are no quantitative studies that measure the effect of clean water/sanitation on household welfare in Nicaragua. The Nicaragua Household Living Standard Surveys (NHLSS) 1998, 2005, 2009, and 2014 can be used to estimate the effect of improved access to water and sanitation beyond water-related diseases using the synthetic control method (SCM).

There are two problems with comparative case studies in economics. First, in comparative case studies, there is typically some degree of ambiguity about how comparison units are chosen. Researchers often select comparison groups based on subjective measures of affinity between affected and unaffected units. Second, there remains uncertainty about the ability of the control group to reproduce the counterfactual outcome trajectory that the affected units would have experienced in the absence of the intervention or event of interest. This type of uncertainty is not reflected by the standard errors constructed with traditional inferential techniques for comparative case studies. However, using SC removes the ambiguity of how comparison units are chosen and shows individual contribution to the synthetic calculation. Also, unlike other approaches like DID, SC can account for the effects of confounders changing over time, by weighting the control group to better match the treatment group before the intervention. The NHLSS contains data on water use of households and on household welfare indicators, including sickness, education, and income. The NHLSS also contains panel data for the difference-in-differences (DID) estimator. This study will then compare DID to SC method. There is a difficulty which arises in measuring the effects of improved water and sanitation access on household welfare because of the endogeneity of the water and sanitation systems. The DID with matching estimator can address the endogeneity bias, provided that this bias is caused by time-invariant unobserved variables. Matching has limitations, but with the use of SC endogeneity can be addressed with less dependence on propensity score matching.

This paper is structured into six sections. The second section introduces the conceptual framework. The third explains the cases and data sources used in this study. The fourth and fifth sections present the methodology and empirical findings of improved access to water/sanitation on income, health, and education. Finally, the sixth section gives the conclusion.

Conceptual framework

Everywhere in the developing world, impoverished households are mired in time-consuming domestic and childcare activities. These households also have to spend substantial amounts of time on activities such as collecting water and firewood. In 1998, on average Managua households spent over 2.66 hours a day getting water. Water fetching is normally a chore done by youth in the home. The long distances and time involved in water fetching have been shown to negatively affect school enrollment (Wadhwa 2016, Nauges 2015). When piped water is available in the home, that dropout rates decreases and educational attainment increases (Koolwal 2013, Nauges 2015, Ortiz 2015, Dreibelbis 2013). Higher overall levels of education lead to skilled labor jobs and higher wages. In 1998 35% of households reported having to fetch their water from a well or lake. This water is much more likely to be unclean due to crosscontamination or wastewater runoff (Cotton 1991, khan 2012). Unclean water can cause diseases, health problems and low labor force participation. Most Nicaraguans do not know how or have no system for checking for water pathogens (Scalr 2017). These pathogens can cause severe illness in both adults and children. Children are more likely to contract a pathogen, and their illness severity is likely to be greater. The child's sickness results in lost time at school for the child and lost income as the adult household member must tend to the sick child (Sclar 2017, Dreibelbis 2013). There is also the increased cost of medical care resulting from the illness.

Cases and data

This article uses balanced panel data from 17 governing regions in the country from 1998–2014 to analyze the influence of the Managua PIAPS on average household (HH) off-farm per-capita income in Managua. The 16 governing regions that are not the Managua region constitute the control group.[1] This data comes from the Living Standards Measurement Surveys performed by the Encuesta Nacional de Niveles de Vida. These nationally representative surveys follow the Living Standards Measurement Survey methodology developed by the World Bank (INEC, 2006). According to SCM, the weights of different components of the synthetic region are determined such that the economic conditions of the synthetic Managua are very close to the real Managua. Control variables included in the econometric specification are regional averages of people in HH, off-farm income per capita, age, age^2, and regional population percent of piped water to the home, rural, straw roof, fetches water, electricity, forages for cooking fuel, head of HH gender, and HH sanitation access. The primary outcome variable of interest is average HH per capita off-farm income, with other variables of interest being average HH education level and total household medical spending. Our level of analysis for this study is at the governmental regional. The data is taken from the LSMS and aggregated up to the governmental regional level.

In this study, off-farm income is the variable of interest. Agriculture-based societies tend to consume part of their production, receive in-kind payments for work performed, and engage in barter and trade with neighbors (Ravallion, 1992). Thus, accurate measurements of household consumption become difficult to calculate. However, off-farm income is a crucial measurement

[1] These 16 regions include Boaco, Carazo, Chinandega, Chontales, Estelí, Granada, Jinotega, Leon, Madriz, Masaya, Matagalpa, Nueva Segovia, Rivas, Rio San Juan, North Caribbean Coast Autonomous, South Caribbean Coast Autonomous Region

because it affects the household consumption bundle. When a household takes part in off-farm work duties this household is more likely to have access to cash, thereby increasing the liquidity of assets and gaining greater access to goods/services. Off-farm income is measured in Cordobas (base year 2014), which is the local unit of currency in Nicaragua.

Methods

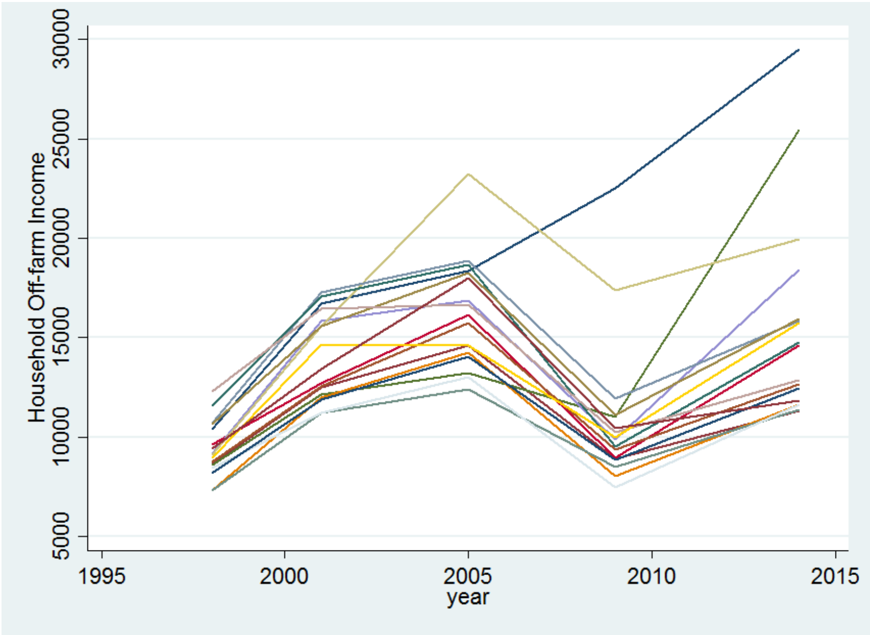

Figure 3: Average HH income per capita over time

DID

In this sub-section, the study provides simple difference-in-difference (DID) estimates and time-trend graphs in order to consider what can be learned from a DID framework. First, Figure 3 shows the time trends of off-farm income per capita for each governmental region. Managua is the solid blue line on top and all others control regionals being the different colors below. The comparison of the trends in the graph indicates that Managua`s income clearly shifted upward compared with the income trends of the other regions. However, there is a great deal of movement both before and after treatment, making it difficult to make the common trend argument needed for DID. Figure 3 also shows that Managua is already very different from the rest of Nicaragua. Managua has a higher population density, greater access to health care, education, and greater economic opportunity. All these things weaken the argument for DID, as

Managua and the rest of Nicaragua are not good comparison groups. While SC can adjust for these limitations by using more than one control unit and weights, DID is limited to using propensity score matching that can be highly selective and sensitive to control variables. This being said, doing a DID analyses as a bench march can give helpful insights. This study implements a simple OLS estimation to make clear the graphical implications mentioned above, with the following DID model:

$$Y_{it} = \alpha_0 + \alpha_1 \mathrm{W}_{\mathrm{it}} + \beta\grave{}X_{it} + \varepsilon_{it} \tag{1}$$

Where Y_{it} denotes off-farm income for household *i* at time *t*,W_{it} is access to water, X_{it} is a vector of control variables, and ε_{it} is an error term. Equation (1) could be estimated using Ordinary Least Squares (OLS) if there was not potential endogeneity between off-farm income and water. The presence of endogeneity is suspected on the basis of studies which reveal the significant impact of water on both income and consumption (Bridge et al., 3 2016a), and the significant impact of income on access to water (Louw et al., 2008; Pachauri and Spreng, 2004). We solve this endogeneity problem by estimating this relationship through a propensity score matching difference-in-differences approach. The difference-in-differences evaluates the effect of a treatment (access to water) on an outcome Y over a population of individuals (household off-farm income). The sample is broken down into two groups of households indexed by the treatment variable W, which is binary, i.e., e∈{0,1}, where 0 indicates households in the control group that do not gain access to water, and 1 indicates households in the treatment group that do gain access to water. The time variable is given as T, where two time periods are observed t∈{0,1}. Period 0 indicates a time period before the treatment group receives access to water, and 1 indicates the time period after the treatment group receives water.

Off-farm income for household *i* would then modeled by the following equation:

$$Y_i = \alpha + \beta_1 X_i + \beta_2 e_i + \beta_3 t_i + \beta 4(e_i * t_i) + \varepsilon_i \tag{2}$$

where β_2 is the treatment group specific effect, β_3 is the time trend common to both the control and treatment groups, and β_4 is the true treatment effect of gaining access to water.

Off-farm income can be indexed by the treatment and time-period variables as Y_t^e, indicating the offfarm income that would be realized given certain values of *e* and *t*. The difference-in-difference estimator is the difference in average outcome in the treatment group before and after treatment minus the difference in average outcome in the control group before and after treatment, so that

$$\widehat{\beta_{DD}} \ = \ (\bar{Y}_1^1 \ - \bar{Y}_0^1) - \ (\bar{Y}_1^0 - \bar{Y}_0^0) \tag{3}$$

Running the regression from Equation 1 would yield reasonable estimates only in the event that those households treated with water were treated at random. As there are many factors influencing whether or not a household becomes connected to water, it cannot be assumed that the treatment is random. Equation 3 allows for systematic differences but requires a common trend assumption. There are many areas however seen from Figure 3 that do have a common trend. We can attempt to address these two issues with the use of propensity score matching, where treated households are compared to non-treated households with similar observed characteristics. The propensity score is the probability of receiving treatment, conditional on X_i . There can be systematic differences but propensity score matching will address those differences by matching treated households with untreated households based on observed characteristics that predict the likelihood of being treated. In this way, matching attempts to make treated and untreated groups that are similar more comparable by trying to match like with like. But this is essentially a selection on observables methods so the method is only as good as the observed variables you are matching on.

The estimation of propensity scores can be done through a binary model as follows:

$$P(E_i \ = \ 1 \,|X_i) \ = \ G(\gamma_0 \ + \gamma_1 X_i) \tag{4}$$

where G(.) is the logistic function:

$$G(\gamma_0 \ + \gamma_1 X_i) = \frac{exp(\gamma_0 \ + \gamma_1 X_i)}{[1 + exp(\gamma_0 \ + \gamma_1 X_i)]} \tag{5}$$

The propensity score for household *i* is then given as:

$$\hat{P}(E_i \ = \ 1 \,|X_i) \ = \ G(\hat{\gamma}_0 \ + \hat{\gamma}_1 X_i) \ = \ \widehat{PS_i} \tag{6}$$

The last step prior to estimating the difference-in-differences estimator is to make certain to compare only households with similar propensity scores. In order to verify this, those households that are treated that have no similar propensity score match in the control group are dropped from the sample. This satisfies the common support assumption for proper identification of a difference-in-difference estimator (Lechner et al., 2011), which is given as:

$$P[TE = 1|X = x, (T,E) \in \{(t,e),(1,1)\}] < 1; \forall (t,e) \in \{(0,1),(0,0),(1,0)\}; \forall x \in X \tag{7}$$

SC

In order to estimate the influence of improved water access on off-farm income, this paper uses SCM to form a reasonable comparison for the treatment group. There are different time trends in treatment group and control group before policy shocks, and they may remain after using DID. There is also the problem of endogeneity of income and water/sanitation. Some of these time trends and endogeneity can be addressed with propensity score matching. However, the variables used to do the propensity score are highly selective and change the results drastically. As a result, it is difficult to accurately estimate the project`s effect. The SCM (Abadie et al. 2010) is used to address this difficulty. The idea of SC methods is as follows: we choose all the regions, which are not subject to PIAPS as the control group, and give a weight to each region. Each region's weight is chosen according to a great number of different weighted control groups to simulate the actual situation of the treatment group. We then choose a vector of weights that give the most unbiased simulation results. We use this vector of weights to form a counterfactual control group after the PIAPS project.

The synthetic control method is explained as follows:

Suppose we have observed the data of off-farm income per capita for J + 1 states. One of them is Managua which is affected by the 2005 water improvement project, the others are not. We observe *T* periods of income data in these areas. T_0 is the year that the project took effect. In this paper, T_0 refers to 2005. Y_{it}^N is the outcome in absence of the treatment effect for region *i* in period *t*. Y_{it}^I is the outcome in region *i* with the effect of the treatment for period t. Where Y_{it}= Y_{it}^N for the regions in the control group and Y_{1t} = Y_{1t}^N for the treatment group before T_0 , or 2005. Therefore, we set the model for:

$$Y_{it} = Y_{it}^N + \alpha_{it} D_{it}.$$

In the model, Managua is indexed as region 1, so for T > T_0, we have D_{it} = 1. We are mostly interested in α_{1t} for T>T_0, which stands for the project effect: α_{1t}=$Y_{1t}^I - Y_{1t}^N = Y_{1t} - Y_{1t}^N$, where

Y_{1t} is the actual off farm income per capita in Managua that has been observed, Y_{1t}^{N} is the hypothetical off farm income in Managua if it were not affected by PIAPS. Since these hypothetical values are unobservable, we need some other regions that are not affected by PIAPS to construct the counterfactual group, which is then used to evaluateY_{1t}^{N}.

Assume that the decision of the off-farm income equation is as follows:

$$Y_{it}^{N} = \delta_t + \theta_t Z_i + \lambda_t \mu_i + \varepsilon_{it}$$

In the equation above, δ_t is the time fixed effects which is the same for all areas; Z_i is the control variable in region *i* that can be observed; μ_i is the region fixed effect that cannot be observed for the treatment group. Unlike DID, these region fixed effects change over time. So for any t > T_0 μ_1 would be unobserved. Meaning that for any t > T_0 the above equation could not be analyzed because of lack of knowing μ_1. ε_{it} is white noise.

In order to estimate the influence of PIAPS on Managua's off-farm income, we need a weighted control group to construct a counterfactual group of Managua. We use W = (w_2,…, w_{J+1})` to represent the weight vector for each area in the control group, where $w_j \geq 0$, $\sum_{j=2}^{J+1} w_j = 1$. Given a vector W, we have:

$$\sum_{j=2}^{J+1} w_j \, Y_{jt} = \delta_t + \theta_t \sum_{j=2}^{J+1} w_j \, Z_j + \lambda_t \sum_{j=2}^{J+1} w_j \, \mu_j + \sum_{j=2}^{J+1} w_j \, \varepsilon_{jt}$$

If there exists another vector W*= (w_2^*,…, w_{J+1}^*)` such that $\sum_{j=2}^{J+1} w_j^* \, \mu_j = \mu_i$ and $\sum_{j=2}^{J+1} w_j^* \, Z_j = Z_1$, from Eq. (3), we know that the synthetic group characterized by w performs as a good estimate forY_{1t}^{N}. But recall that μ_1 is unobservable, we thus cannot find the ideal w* in the explicit way. However, suppose there is a set of weight vector W*=(w_2^*,…, w_{J+1}^*)` that could be applied to the observed control groups to estimate the counterfactual of the treatment group satisfying:

$\sum_{j=2}^{J+1} w_j^* \, Y_{j1} = Y_{11}, \ldots, \sum_{j=2}^{J+1} w_j^* \, Y_{jT_0} = Y_{1T_0}$, and $\sum_{j=2}^{J+1} w_j^* \, Z_j = Z_1$

If $\sum_{t=1}^{T_0} \lambda`\lambda$ is nonsingular, we further have:

$$Y_{1t}^{N}-\sum_{j=2}^{J+1} w_j^* Y_{jt}=\sum_{j=2}^{J+1} w_j^* \sum_{s=1}^{T_0} \lambda_s\left(\sum_{n=1}^{T_0} \lambda`\lambda\right)^{-1} \lambda_s\left(\varepsilon_{js}-\varepsilon_{1s}\right)-\sum_{j=2}^{T_0} w_j^*\left(\varepsilon_{js}-\varepsilon_{1s}\right)$$

Abadie et al. (2010) shows that the right-hand side of the above equation converges to zero as J increases under several quite parsimonious requirements. So $\sum_{j=2}^{J+1} w_j^* Y_{jt}$ is the unbiased estimation of Y_{1t}^{N}. When $t \geq T_0$, we cannot observe Y_{1t}^{N}. So we can use $\sum_{j=2}^{J+1} w_j^* Y_{jt}$ as an estimate of Y_{1t}^{N} to evaluate the policy effect. The weight vector W*= $(w_2^*, \ldots, w_{J+1}^*)`$ is found by minimizing the distance function $\|X_1 - X_0W\|_v = \sqrt{(X_1 - X_0W)`V(X_1 - X_0W)}$. In this function, X is a feature vector of regions, which corresponds to the control variable Z. This can be observed in every region and in Y before PIAPS impact. The symmetric positive semi-definite matrix V determines the importance of different feature vector X in structuring weight.[2] This give us $Y_{1t}^{N} = \sum_j^J W^* Y_{jt}$ that we can plug back into α_{1t}=$Y_{1t}^{I} - Y_{1t}^{N}$ to get the impact of the treatment α_{1t}.

SCM permits for impact estimation in settings wherever a single unit (a state, country, firm, etc.) is exposed to an event, intervention, or treatment. It provides a data-driven process to build SC units based on a convex combination of comparison units that approaches the characteristics of the unit that is exposed to the intervention. A mixture of comparison units regularly provides an improved comparison for the treated unit than any comparison single unit alone. Additionally, data-driven processes reduce discretion in the choice of the comparison control units, forcing researchers to validate the similarities between the affected and unaffected units using perceived quantifiable characteristics. SC expands the conventional linear panel data (difference in differences) model, permitting that the effects of unobserved variables on the outcome vary with time. Unlike the DID model, the synthetic comparison group is a weighted average of comparison individuals. Thus, it is made clear what each control group's contribution is in constructing the synthetic comparison group.

SCM has been widely used in recent years with comparative case studies. In political science, SC has been used to study the influence of Spanish Basque terrorist attacks on its

[2] In this paper, X1 is the people in HH, rural, straw roof, HH fetch h20, Toilet, electricity, Forages wood, male, age, and age2 before 2005. Accordingly, X0 is the people in HH, rural, straw roof, HH fetch h20, Toilet, electricity, Forages wood, male, age, and age2 of the control group areas before 2005. The details of Synthetic Control Method can be found in Abadie et al. (2010).

economy. The study constructs a counterfactual group of Basque or synthetic Basque. It constructs this by using the information on the areas that did not suffer terrorist attacks (Abadie and Gardeazabla 2003). They find that, after the outbreak of terrorism in the late 1960's, per capita GDP in the Basque Country declined about 10 % points relative to an SC region without terrorism. Using SCM Abadie et al. (2014) gage the influence of German reunification in 1990 for West Germany's economic development. They found that per capita GDP of West German fell by $1,600 a year between 1990 and 2003, because of the German reunification. In 2003, the per capita GDP of synthetic West German is 12 % higher than the real West German. In health economics, Kreif et al. (2015) compared the commonly used DID estimation with SC on hospital policy changes and mortality rates. The DID showed a reduction in mortality rates erroneously while SCM showed an increase in mortality rates. This paper uses SCM in a developmental economics setting and evaluates the impact of water policy on off-farm income per capita. The contribution of this paper is to improve the understanding of access to clean water and sanitation effects on income by means of health, and education. This paper also contributions to the recent but growing literature of the SC method and its uses.

Results

Trends and simple DID as a benchmark

The DID results are shown in Table 1. DID estimates for the 2005 water project suggest that part of the difference in income might be due to the 2005 project. The results for education and medical spending are inconclusive. The signs on the coefficients are what we would expect from the project (+education, -medical spending), yet they are not all significant. In summary, DID estimates suggest that PIAPS could have some positive ‘‘effect’’ on off-farm income.

Table 1: DID and SCM results

Panel A: No matching	HH off-farm income	Average HH education	Sickness
$\widehat{\beta_{DD}}$	10,729	0.0687	-0.0266
	(8,719)	(0.259)	(0.02)
Constant	2,516***	4,495***	0.753***
	(211.9)	(0.094)	(0.012)
Observations	10,776	10,776	10,776
R-squared	0.000	0.000	0.000

Panel B: Propensity score matching	HH off-farm income	Average HH education	Sickness
$\widehat{\beta_{DD}}$	8,496**	3.996***	-0.178***
	(3,855)	(1.313)	(0.0492)
Constant	4,495***	6.243***	0.706***
	(499.4)	(0.258)	(0.0284)
Observations	4,311	4,311	4,311
R-squared	0.014	0.037	0.005

Panel C: SCM	HH off-farm income	Average HH education	Sickness
$\widehat{\beta_{SCM}}$	13,654**	0.384***	-0.189***
	(7,623)	(0.112)	(0.0337)

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

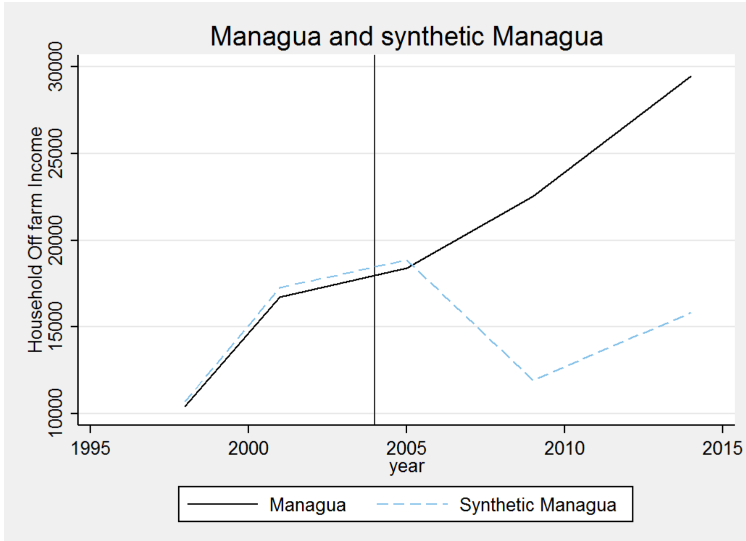

Figure 5: Average HH off-farm income

We examine SCM by presenting systematic graphical results for outcome variables of interest affected by PIAPS. Here, the results show that there is little difference in average off-farm income, between the synthetic Managua and the real Managua, prior to PIAPS (up to 2005; Figure 5). Average household characteristics were also similar between the real Managua and the synthetic Managua (Table 2). Table 3 displays the weights of each control region in the synthetic Managua and Figure 6 shows the time trends of off-farm income per capita for each governmental region used to produce synthetic Managua. The weights reported in Table 3 indicate that income trends in Managua prior to PIAPS is best reproduced by a combination of Granada, Masaya, Nueva Se, and Carazo. All other regions in the donor pool are assigned zero W-weights. The 'gap' between synthetic and real Managua before PIAPS started indicates the quality of the synthetic control region for comparison; the gap after PIAPS started can be attributed to the effect of improved water/sanitation access. For 2014 the estimated effect of PIAPS on HH off-farm income is $13,654 NIO (Table 4) ($453 USD), and we can reject the idea that these estimated effects are random ($p<0.00$). The SC method estimates a larger income affect then the DID method, but a smaller education effect and roughly the same effect on illness (Table 1).

Table 2: Means of HH characteristics measured before the 2005 PIAPS project

	Treated	Synthetic	Average
People in HH	5.175	5.107	5.310
Rural	0.104	0.183	0.470
Straw Roof	0.004	0.012	0.025
HH fetch H2O	0.083	0.211	0.454
Toilet	0.508	0.383	0.193
Electricity	0.976	0.837	0.692
Forages Wood	0.100	0.223	0.458
Male	0.634	0.683	0.699
Age	47.235	47.714	48.133
Age^2	2448.645	2524.851	2566.889

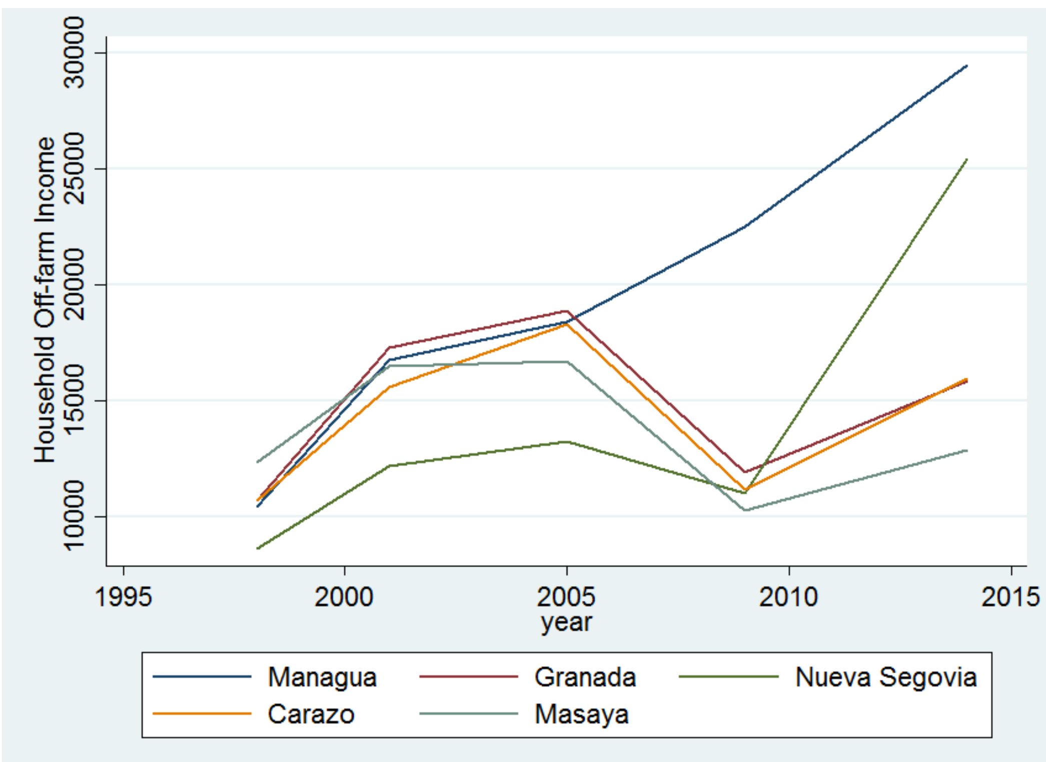

Figure 6: Average HH off-farm income per capita over time for contributing regions

Table 3: Region Synthetic Unit Weight

Region	Unit Weight
Nueva Se	0.145
Jinotega	0
Madriz	0
Estelí	0
Chinande	0
León	0
Matagalp	0
Boaco	0
Masaya	0.221
Chontale	0
Granada	0.565
Carazo	0.069
Rivas	0
Río San Juan	0
RAAN	0
RAAS	0

Table 4: HH Income effects by year post program

	estimates	pvals	pvals_std
c1	-480.4453	.9375	.75
c2	10594.05	0	0
c3	13654.66	0	0

Robustness

The same method used to get the results for Managua are allied to the other 16 regions one at a time excluding Managua to get placebo effects. Figure 7 shows the estimated gaps for the 16 placebo regions, demonstrating a good pre-intervention fit, and for the post-intervention period, randomly scattered around zero. There are two outliners but none as high as what we calculate for Managua. These outliners mean that our results would have a downward bias. The right panel shows the empirical distribution of the placebo ATTs. The distribution of average placebo effects (averaged over the years of each placebo trial) is roughly bell-shaped with a slight right skew and a peak at zero, indicating that ‘‘placebos’’ do not cause systematic impacts on control units.

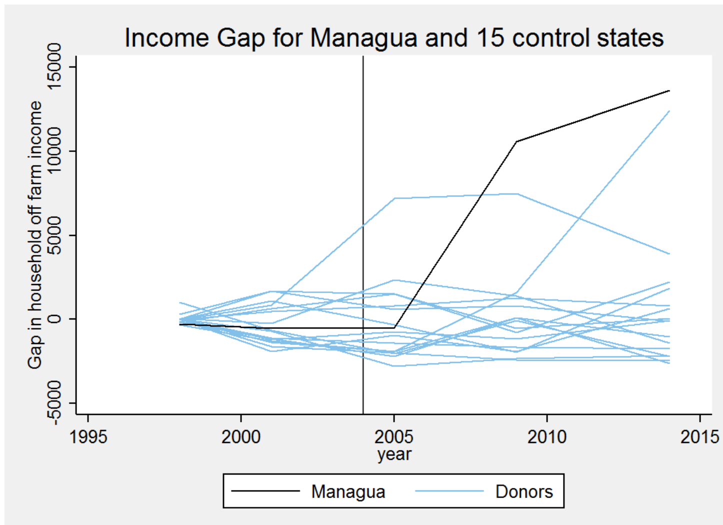

Figure 7: Managua vs. the synthetic Managua (black line) compared with the distribution of 16 placebo gaps (blue lines).

When it comes to the magnitude of average treatment effects, Figure 8 shows that the estimated average effect in Managua is larger than 99.5% of placebo effects (CDF > 0.995). This threshold is comparable to the 1% significance level in conventional two-sided tests. The information from both Figure 7 and 8 imply that it is very rare to randomly obtain the average treatment effects from the placebo distributions. Thus, it can be concluded that the 2005 PIAPS project has some positive noticeable effects on HH off-farm income.

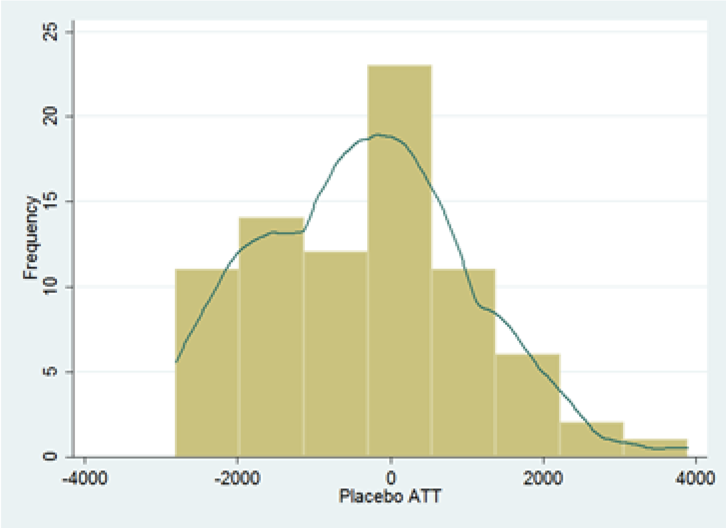

Figure 8: Distribution of estimated placebo average treatment effects

Indirect effects on HH income via education and health

These results strengthen the identifying assumptions of SCM. The results also suggest that the treated and synthetic units are comparable in the post-intervention period. This allows us to test some of the indirect effects that could be causing this continues upward trend in off-farm income compared to the synthetic region. Next, the paper examines the indirect effects via education and health. Figure 9 and 10 show the comparisons of average HH education and medical spending in Managua compared to synthetic Managua. The Figure 9 shows medical spending has changed relative to their SC units after the project. This provides the HH more liquidity and implies less time being sick or caring for the sick. This would allow more time to pursue education or other income sources (Koolwal 2013, Nauges 2015, Ortiz 2015, Dreibelbis 2013 Sclar 2017, Dreibelbis 2013). Figure 10 shows the change in educational attainment relative to their SC units before and after the project.

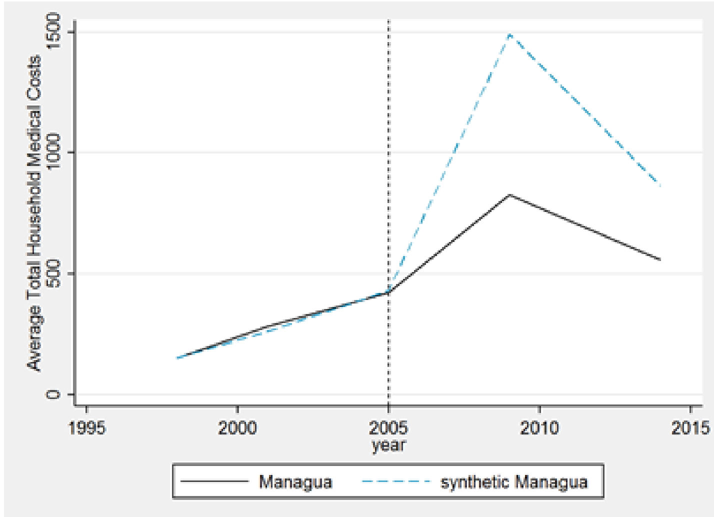

Figure 9: Average HH medical Costs

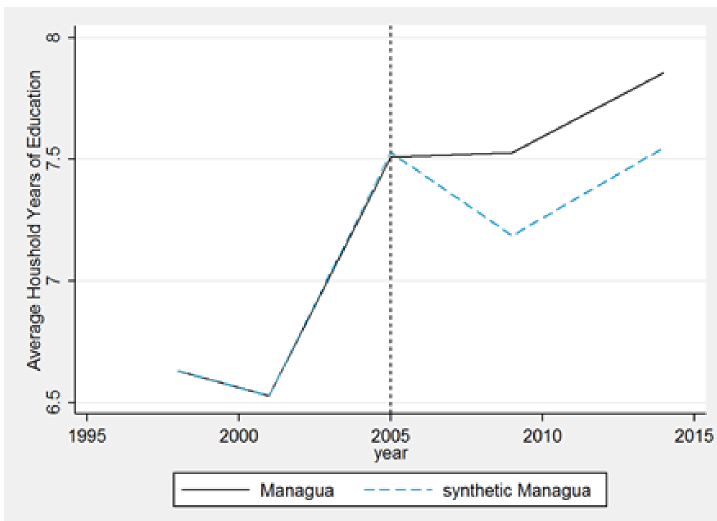

Figure 10: Average HH years of Education

Considering the magnitude of the SC estimates and the clear increase in HH education and decrease in medical spending compared with SC units in Managua, it is plausible to conclude that the sizable indirect effects on HH off-farm income caused by the 2005 PIAPS project in Managua have increased local employment and wages. As a whole, the project has caused higher local HH education and lower HH medical spending, which may be part of the cause for the higher per capita income levels.

Conclusion

Although clean water is vital for health and human development, many impoverished people still do not have access to clean water in Nicaragua. While approximately 85% of households in 2010 had access to improved water sources, these sources were only available 56% of the time (WHO 2010). When these improved water sources are not available, households use water from wells, and some households still use drinking water from rivers, ponds/lakes and simple wells without any purification. Without proper sanitation and wastewater treatment, many of these water sources are becoming unclean. Unclean water causes diseases, health problems and low labor force participation. The higher rates of illness and morality for children, along with time spent fetching water can lead to low educational attainment and medical fanatical burdens. This results in lower income and consumption to a comparable group. This study aimed to measure the effect of improved water access on household welfare indicators including income, education, and health. We found that the effect of the 2005 PIAPS project on HH off-farm income and education is positive and statistically significant. While income per capita in Managua increased by $923 USD, this paper suggests that $453 USD of this increase can be attributed to increased access to clean water and proper sanitation. The project also had a statistically significant negative effect on HH medical spending, likely caused by improved HH health.

This paper's findings contribute to the evaluation of water/sanitation infrastructure and development policies. To begin with, the 2005 PIAPS project has had a clear positive impact on local off-farm income levels. In this sense, improved water access may work as a place-based policy for increasing local off-farm income growth, where access to water is intermittent or limited.

By using SCM, this paper was able to take into account the endogeneity of income on water and sanitation by accounting for these effects in the control SC units. I was also able to investigate the possible causal mechanism of the effects of PIAPS on off-farm income levels through post-estimation comparisons between Managua and it`s SC units. The approach adopted in this study can be characterized as particularly useful when researchers and policymakers want to examine the impact of individual cases rather than a single average impact or when the number of cases is small and the estimation of an average effect could be difficult, potentially misleading, or have large endogeneity effects. However, from the point of view of a more detailed case study, that solely focuses on one case with better quantitative and qualitative data, the data-driven procedure of the SCM technique with comparatively restricted sets of pre-determined covariates could also be perceived as a crude research design. I would nonetheless argue that the data-driven procedure of SCM provides transparent results that are comparable across different cases and easily reexamined by other researchers. Case studies using more extensive qualitative and quantitative methods and focusing on individual water/sanitation improvement sites would complement, not replace, the findings of this study.

Chapter 3

Effects of Water Poverty on Household Welfare

Introduction

Currently, 44% of the world's population must leave their homes to fetch the water they need for drinking and other domestic needs. Women and children are known to be the main water carriers in low-income countries, often spending more than an hour per water collection trip and making multiple trips per day. (Montgomery, Elimelech 2007; Sorenson et al. 2011) Although the large global time burden of water fetching is well known, much of water-related impact research has focused on water quality and health (Hunter 2009; Schmidt, Cairncross 2009) leaving the relationship between water fetching and consumption understudied. (Wang, Hunter 2010) Improved water access is generally thought to be one way. As income increases households improve water access and sanitation. (Sorenson et al. 2011) However, few studies have looked at how improved water access can increase household consumption, let alone the mechanisms on how it does this.

The time cost and physical burden associated with water fetching translate into reduced volumes of water accessed by households using non-networked sources. Previous research has found an inverse association between volume of water used and walk time to the source; households whose water sources are located more than 30 minutes away often collect less water than is believed necessary for basic needs (Cairncross 1987). The proximity of water available to a household has also been demonstrated to correlate with the frequency of hygiene behavior. For example, mothers in Burkina Faso with piped water supplies in their yards were three times more likely to perform regular hand washing as compared to mothers using wells or public standpipes outside their yards (Curtis et al. 1995). Households in Latin America with access to piped water on their land have been found to use twice the volume of water for personal hygiene as compared to those without on-land access to water (Tumwine et al. 2002). Also, the quality of water from wells and public standpipes has been shown to be of poor quality. This is due to both the source being unclean and the contamination of the collection mechanism from improper and frequent use (Shaheed et al. 2014). Pesticide, agricultural runoff, along with human or animal waste can containment local lake, wells, and rivers. Also, improper sanitation of those using the common water source can be transfer back to the water source or directly to other users. Touching the water spout or well handle can pass bacteria from one to another. While these community water sources see frequent use they are not regularly cleaned or maintained. One community-specific study in Ethiopia found that installation of village taps reduced time spent fetching water and child mortality (Gibson, Mace 2006). A recent systematic review of studies investigating the relationship between the distance from the home to a water source and diarrheal disease identified only six studies, four of which did not adjust for possible confounding variables. The review authors were unable to calculate a quantitative relationship between distance and illness risk and concluded that more research is needed on this topic. (Subaiya et al. 2011; Wang, Hunter 2010) Lower levels of water use and poor water quality have been shown to increases sickness greatly in children. This increase in child morbidity and mortality has many spillover effects in the household. Household medical costs increase as they try and treat the child, along with the loss of time spent working as adult household members must take time off work to care for the child. These children are also more likely to drop out of school compared to their peers, because of the increase in absences and the time requirement of water fetching (Brown et al. 2013).

Previous research on water access and health has explored the impact of households gaining access to on-plot piped water connections, but little is known regarding the extent to which water fetching affects child health or education and resulting household consumption (Zwane, Kremer 2007). Also, Long water fetching distances have been shown to increase dropout rates for youth. (Brown et al. 2013) There have been no studies, to our knowledge, that have looked at how improved water access could improve consumption and the mechanisms that drive it.

To address this knowledge gap, this work investigates the association between household consumption and access to water. The mechanisms addressed in this study that could affect consumption are time spent fetching water and improved health. This would lead to lower school dropout rates and more time spent working.

The present study is novel in that it examines the effects of multiple manifestations of water poverty on both educational and health outcomes for individuals in a poor country. This paper uses data from Nicaragua's 2014 Living Standards Measurement Survey to look at piped water access effects on household consumption and individual-level effects on education, health, and time spent working. These relationships are estimated using an instrumental variables models to account for endogeneity. Specifically, the expected cost of extending water infrastructure to rural areas of a municipality is proxied by three plausibly exogenous factors: the mean slope gradient of the land in the municipality, the tree coverage, and the 2005 population density in the municipality. These instruments are strong predictors of the probability that a rural household reports having piped water, provisional on municipal fixed effects, but are not correlated with unobserved factors impacting consumption. This study finds that piped water has a significant positive impact on consumption. Piped water access also shows increased enrollment in primary and secondary school, improved health for children in the household, and a decrease in number of hours worked for children with an increase in the number of hours worked for adults in the home.

Background

Latin America has some of the largest renewable water sources in the world. However, parts of Latin America also suffer the most in the world from economic water sacristy (Malik, 2014). This study will focus on the situation of Nicaragua, which is the least developed country in Latin America. As seen in Figure 1 Nicaragua has the second highest child mortality rate due to diarrhea for all of Latin America (UNICEF 2018). As of 2014, it was ranked 132 out of 187 countries in the United Nations human development index. Compared to the other Central American countries Nicaragua has the lowest GDP per capita of any of the surrounding countries, it is among the least educated, lowest rates of water access, and the highest incidence of child mortality (Malik, 2014). Piped water services are the most reliable sources for safe drinking water (Irianti, Sri et. al 2016). As piped water may still be contaminated, the use of household water treatment and safe storage is also encouraged by local governments to eliminate disparities in the burden of water-related illnesses where community-based water supplies are infeasible (Irianti, Sri et. al 2016). Community wells or public water taps have been shown to be cleaner sources then lakes or rivers (WHO 2014), yet they are still below that of piped water to the home or private source. (WHO 2014) The high use of community water sources leads to a higher likelihood of water tap contamination. (WHO 2014)

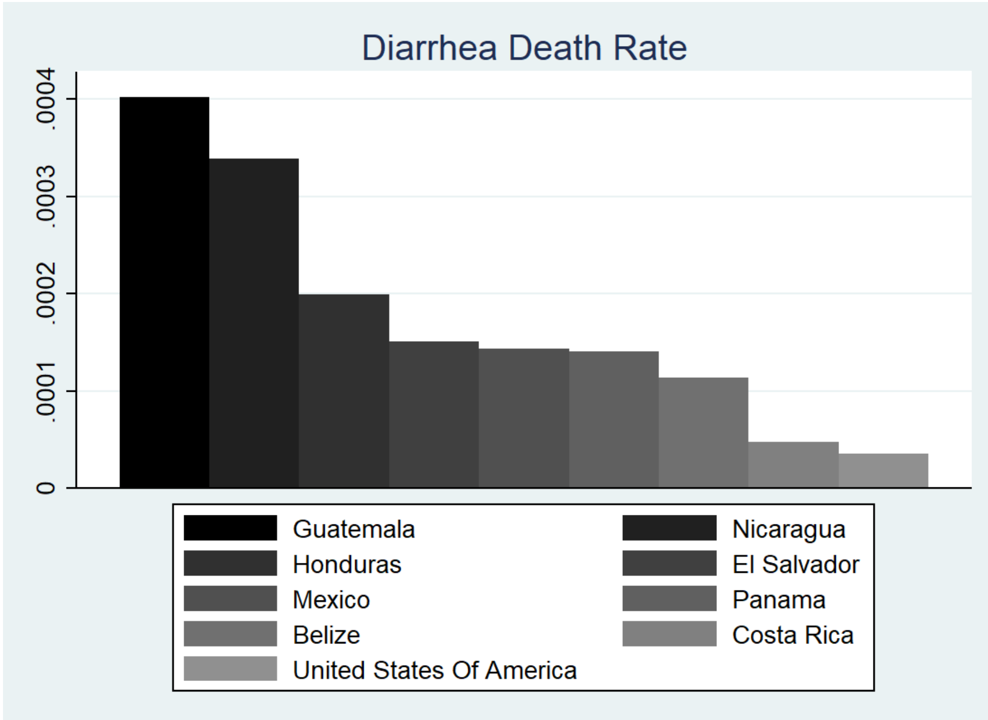

Figure 1: Diarrhea Death rate for Latin American Countries (2014)

Over 80% of Nicaragua`s rural population has to fetch their water. Meaning that they have to travel some distance to a community or public water source for all water access. Figure 2 shows the piped water access rates for urban and rural residents for 1998-2014, broken down by poverty group. Where piped water access is water obtained from a pipe on the property, either on their land or in the home. The largest factors for household piped water access are piped water education of household head, household wealth, population density, sanitation facilities, household sizes, tree coverage and slop of land. (Irianti, Sri et. al 2016) The largest factors for household piped water access are piped water education of household head, household wealth, population density, sanitation facilities, household sizes, tree coverage and slop of land (Irianti, Sri et. al 2016). Over this period the vast disparity in access to water between poverty groups in urban areas has been largely diminished, with around 84% of even extremely poor urban individuals sampled having at least some access by 2014. However, in urban areas, there are still people that have to fetch water. This is either because of improper infrastructure or because they built their home on land that just did not have access to piped water. These individuals while still in an urban setting because their living conditions are forced to obtain their water from some source outside their home.

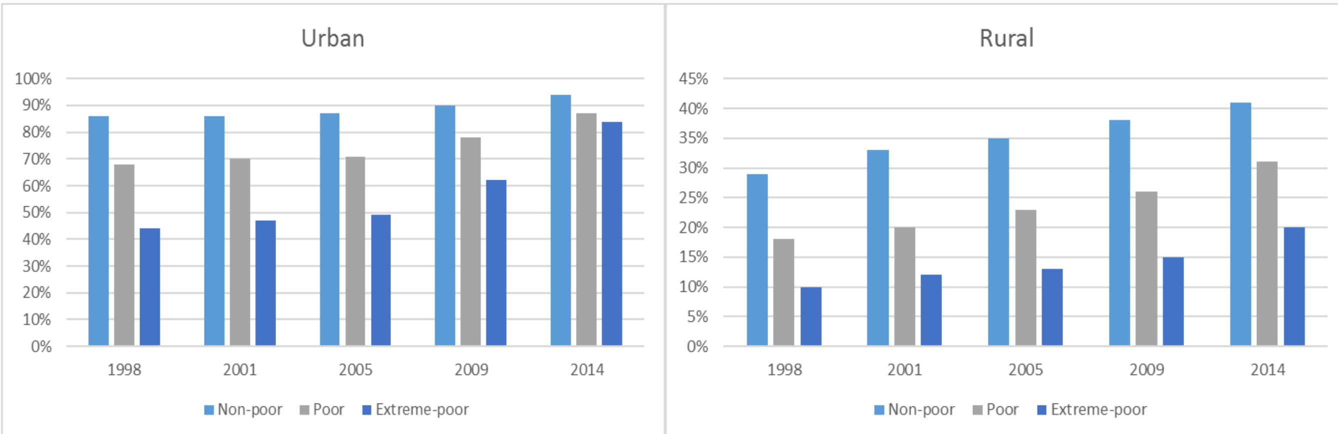

Figure 2: Piped Water Access Rates by Poverty Group

Rural populations, however, are still largely suffering from economic water scarcity. Table 1 shows that while water access is increasing through the years, a large proportion of poor and extremely-poor individuals are still lacking even a minimum amount of water access. Only 20% of extremely poor individuals sampled in rural Nicaragua have piped water access as of 2014. The largest sources of water for rural individuals are wells and lakes. These are commonly far from their place of residence. The average travel time to fetch water was 1.5 hours, and these trips are often made more than once a day.

Figure 3 displays the piped water access rates by the municipality. We can see that while there is high water access around the urban areas (deep blue), the majority of Nicaragua has very low water access rates.

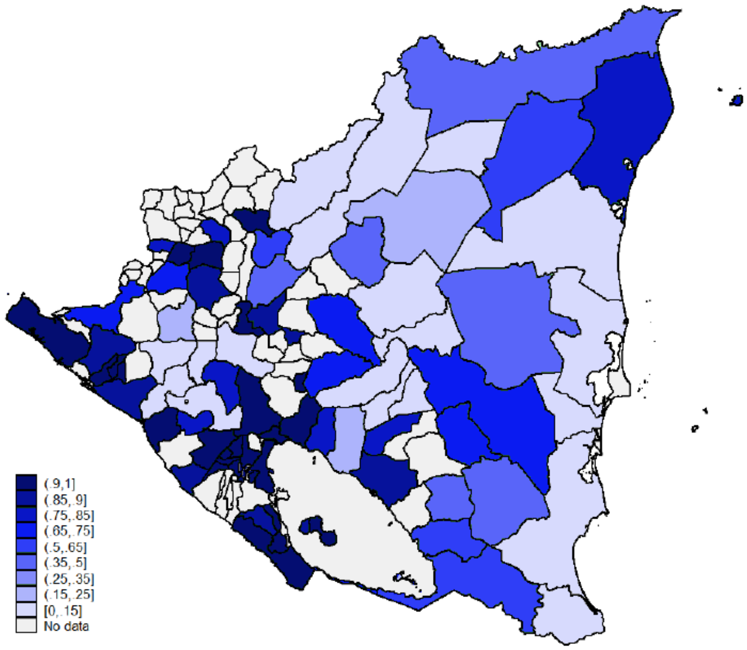

Figure 3: Spatial Distribution of Individual Rural Piped Water Access Rates (2014)

The individual's time spent fetching water is the second main aspect of water poverty that will enter into this analysis. Figure 4 shows rural averages of time spent fetching water, broken down into poverty groups. The average time spent fetching water also does not vary much based on poverty status. Rural individuals that fetch water spend about 1.48 hours a day doing so. The Non-poor households that do fetch water have a higher average for time spent fetching water. This is because fewer people overall fetch their water and there are less available water sources (wells, lakes, rivers, etc.) in richer settings.

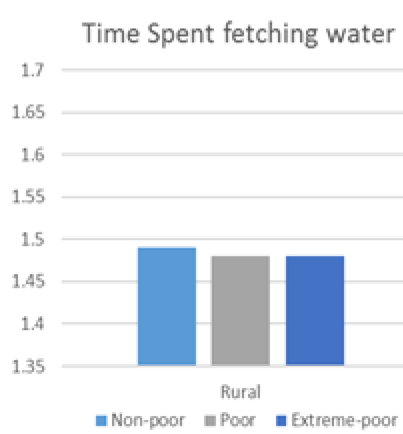

Figure 4: Time spent fetching water by Poverty Group

Data

The data for this study comes from the living standards measurement surveys (LSMS) conducted in Nicaragua in 2014 (INIDE, 2014). This is a nationally-representative survey which follows the methodology developed by the World Bank, which contains living-standards information for households and individuals. To econometrically estimate the effects of water poverty on human development, it is necessary to have an exogenous variation in the data with regards to water use. Exogenous variation has a greater likelihood for piped water access in rural areas than in urban. As seen in Figure 2, water access is becoming ubiquitous in the urban areas of Nicaragua. For this reason, we only use observations of rural individuals when estimating the impacts of water. This household survey data was combined with municipal population density data from the 2005 National Census (INIDE, 2006), as well as geographic data on the mean slope of the land at the municipal level. This is calculated using ArcGIS version 10.5 geographical software and municipal maps from ESRI Data and Maps (2014). The slope gradient as an instrument was also used in a study of infrastructure projects in South Africa (Dinkleman 2011). This geographic data was compiled by Grogan and Sadanand (2013). Finally, we add tree cover data at the department-level (Global Forest Global Forest Watch, 2000) to complete the data set used for this analysis.

Our instruments might be associated with other factors that have direct impacts on our outcomes of interest. This is in part because the vast majority of potential employment in rural areas is agricultural. For example, the slope of the land in a municipality may reflect agricultural productivity, which would impact the potential wages, and thus consumption. As well, population density in a municipality could also reflect the size of potential agricultural plots available, or impact the price of land. Higher population densities or lower slopes could lead to large access to schools. It might also be correlated with the extent of non-farm employment. Rural areas with more water access could also have higher wages. Thus, while our instruments are arguably exogenous to a given household, we are particularly concerned about municipal unobservables which could be correlated both with our instruments and with our outcomes of interest. To ensure that our instrumental variables strategy is valid, we include controls for municipal and household level factors that reflect the local labor market situation. In all models, we control for the fraction of the municipal population which currently resides in higher population areas. This means that the historic municipal population density instrument, also partially captures the sparsity of the population. As well, we control for the distance between an individual's household and the nearest school. These two variables are assumed to proxy the level of dynamism of the local labor market. Together, these controls help ensure that unobservable municipal level factors are not biasing our estimated impacts. Including municipality fixed effects in all models allows us to control for local, but unobserved, fixed factors. We are also able to test the joint validity of our instruments because we have multiple instruments for one endogenous variable.

The consumption variable in the data is an aggregated continuous variable that measures per capita yearly expenditures of food, beverages, and nonfood products and services (e.g. housing, health, education, furnishings, transportation, personal expenses, and home maintenance). This consumption variable is also used to classify households into three poverty categories; extremely poor, generally poor, and non-poor. Extremely poor households were classified as such if their food consumption levels fell below the minimum daily calorie requirements. Because minimum daily requirements vary greatly according to gender and age. We calculated household-specific daily kcal per capita requirement based on household individual equivalencies (INIDE, 2011). The level of extreme poverty for rural households sampled is 15%.

The level of annual per capita consumption required to meet minimum daily caloric requirements plus a sufficient amount for housing, transportation, education, health, and clothing is assumed to be twice that of extreme poverty. If a household's consumption level falls between this line and that for extreme poverty then it is classified as generally poor. Households with consumption levels higher than the general poverty line are classified as non-poor.

In order to observe high variation in water access and educational outcomes, only those households residing in rural areas are used in estimating water access impacts on quality of life. This leaves 4,598 observations. Out of these individuals, 1,437 have piped water access. The time spent fetching water for the 3,160 without access is 1.47 hours per day. Piped water access is a dichotomous variable measuring whether or not a household has access to water in their home or on their land. When estimating the impact of piped water on education, a dichotomous variable is used indicating whether an individual of correct age was currently enrolled in school (primary 6-12, secondary 12 -14). When estimating the impact of water on health a dichotomous variable is used indicating whether an individual had any non-chronic illness in the last month. To estimate the piped water effect on labor the variable hours worked is used corresponding to the number of hours work on average per week over the last 6 months. Other characteristics that are used in this analysis include, whether the individual lives in a dwelling with a dirt floor and/or straw roof, poverty status, as well as the gender, age, and education level of the head of household.

Table 1. Descriptive statistics by water access

	Piped water		No Piped water		
	Mean	Sd	Mean	Sd	Diff
Household per capita consumption	20,097.33	12,163.21	24,193.73	14,939.61	-4,096.39***
Enrolled in primary school	0.12	0.33	0.13	0.33	-0.01
Enrolled in secondary school	0.09	0.29	0.11	0.31	-0.02*
Individual years of education	3.91	3.65	5.33	4.13	-1.42***
(max) sick	0.89	0.32	0.86	0.35	0.02**
Age	25.05	18.97	26.20	19.54	-1.15*
Years of Education(head)	3.06	3.41	4.62	3.98	-1.57***
Have straw roof or similar	0.05	0.23	0.00	0.04	0.05***
Dirt Floor	0.71	0.45	0.49	0.50	0.23***
Male	0.52	0.50	0.49	0.50	0.03
Population Density(log)	3.54	1.27	4.63	1.30	-1.08***
Percentage of Department Tree Cover	-0.60	0.15	-0.55	0.13	-0.05***
Mean slope Gradient of Municipality	-11.67	6.36	-12.49	7.20	0.81***

Table 1 gives descriptive statistics of the primary variables used in our estimations. We can see that those households with piped water access have higher levels of education are more likely to be enrolled in school, have higher levels of consumption and work more. They are less likely to be poor, have dirt floors, or straw roofs.

Consumption

The scope of this study is to investigate the ways and mechanisms in which access to piped water sources impacts quality-of-life for households in the developing world. Figures 2,3 and Table 1 give an initial impression that piped water is negatively correlated with poverty in Nicaragua. In order to arrive at a more in-depth understanding of these impacts, we turn to more rigorous methods.

Human development as measured by education, health, and consumption is co-determined with water use. Care is required, however, in estimating these endogenous relationships. It should be easy to measure, for example, how higher income levels lead to great access to water. It should also be fairly obvious that a decrease in fetching time or illness may result in an increase in income through enhanced labor productivity. This endogenous relationship can reasonably be expected to reveal itself in the estimation procedure.

The codetermination of water and health, or water and education may be a bit more complicated. While water source may have a direct effect on health and education measurements, the inverse effect will likely come indirectly through the consumption component. Indirect effects often are subject to time horizons that fall outside of the scope of cross-sectional data. This must be kept in mind throughout the proceeding estimation efforts.

This endogeneity will be addressed through the use of instrumental variables. When estimating two equations simultaneously, the requirements of a valid instrument require that it is correlated with the dependent variable in the first equation while being uncorrelated with the error term in the second equation. In the current application, this requires that one or more variables are used that is correlated with having access to piped water while being uncorrelated with consumption. These variables are (1) the mean slope gradient of the land in the municipality, (2) the population density in the municipality as measured in 2005 by the Nicaragua census (INIDE, 2006), and (3) the amount of tree cover in the municipality. An increase in population density will result in an increase in the availability of piped water access. Whereas higher levels of tree cover and land slope would make extensions of the water grid more challenging.

The estimation technique that will be followed is an instrumental variable approach. Equations 1 and 2 show this estimation strategy:

$$Y_i = \alpha_0 + \alpha_1 E_i + \vec{\beta}' \vec{X_{ci}} + \epsilon_{ci} \tag{1}$$

Where Y_i is the per capita consumption for household i, α_0 is an intercept, E_i is a dichotomous variable equal to one if household i has access to piped water and equal to zero otherwise, $\vec{X}_{ci}$ is a vector of regressors for household i. These controls are educational attainment(head), extrema poor, poor, male, age, age^2, straw roof, dirt floor, and one room home. The same controls are used in all models. The error term is represented by ϵ_{ci}. The piped water equation is given as:

$$E_i = \gamma_0 + \vec{\gamma}' \vec{z}_i + \vec{\delta}' \vec{X_{ei}} + \epsilon_{ei} \tag{2}$$

Where γ_0 is an intercept, $\vec{z}_i$ is a vector of instrumental variables, $\vec{X}_{ei}$ is a vector of regressors relating to the piped water access of household i, while ϵ_{ci} is an error term. This estimation will take place in two stages. First, equation 2 will be estimated using OLS. Once this is estimated, the predicted value of piped water ($\hat{E}_i$) will be used to replace the regressor for piped water in Equation 1, with Equation 1 becoming:

$$Y_i = \alpha_0 + \alpha_1 \hat{E}_i + \vec{\beta}' \vec{X}_{ci} + \epsilon_{ci} \tag{3}$$

All models in this paper use municipality level fixed effects and clustering errors by household.[3] The municipality level fixed effects are particularly important because these refer to small geographic areas where local economic conditions, access to schools, and water infrastructure are likely similar.

Consumption Results

The results of Equations 2 and 3 are found in Table 2. A Hansen's J-test for overidentifying restrictions returns a p-value of 0.9834, failing to reject the null hypothesis of valid overidentifying restrictions. It is also shown in Table 2 that the three instruments chosen are highly statistically significant in the piped water equations. Our F-statistic for the instruments, shown at the bottom of Table 2, confirms that these instruments are not weak. The p-value of the Chi^2 test also confirms that the instruments jointly satisfy the overidentification restrictions. Thus, the null hypothesis of weak instruments is firmly rejected.

Using instrumental variables, it is observed that a household with piped water access consumes \$2,194(US\$ 90.29) per capita more than a household without. The magnitude of this effect is quite large, especially in light of the magnitudes of the other regressor coefficients.

[3] As a robustness check, equations 2 and 3 are estimated using a number of various techniques. We find and that our results remain qualitatively the same. Results available in appendix A.

Table 2. Consumption

	OLS	Consumption	Piped water
Years of Education(head)	1973.1***	498.4***	0.0162***
	(35.34)	(47.48)	(0.00189)
Extreme Poverty	-20981.8***	-18158.9***	-0.0632***
	(1200.4)	(452.7)	(0.0199)
General Poverty	-16972.7***	-13905.5***	-0.0203
	(506.9)	(330.0)	(0.0144)
Male	352.8	387.5	-0.0134
	(319.8)	(287.5)	(0.0123)
Age	175.7***	53.16***	0.000187
	(8.548)	(8.123)	(0.000348)
Age^2	1.697***	0.195*	0.00000708
	(0.119)	(0.106)	(0.00000452)
Have straw roof or similar	8171.9	1741.9**	-0.0808**
	(5954.9)	(792.6)	(0.0334)
Dirt Floor	-4875.3***	-3065.3***	-0.125***
	(447.6)	(349.9)	(0.0138)
One Room home	-3740.7***	-974.8***	-0.0513***
	(432.6)	(341.6)	(0.0150)
Piped Water	3702.9***	2405.3***	
	(514.6)	(909.9)	
Population Density(log)			0.120***
			(0.00573)
Distance to Nearest Public School			-0.0231***
			(0.00322)
Percentage of Department Tree Cover			-0.204***
			(0.0539)
Mean slope Gradient of Municipality			-0.00366***
			(0.000980)
Observations	4532	4532	4532
F-stat		93.64	
P-value > Chi^2		0.00	

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2 shows the positive and significant effect of piped water access on household per capita consumption across the sample. However, we still have little detail about the mechanisms that piped water affects consumption.

Mechanism Econometric Methodology

The second analysis is to understand the mechanisms through which access to piped water impacts the quality of life for individuals and the resulting consumption in the developing world. Educational outcomes and water poverty are likely simultaneously determined in the long run, but this impact will be indirect. In other words, an individual that obtains access to piped water will not immediately experience an increase in the quantity or quality of their education. Rather this effect will happen over time. As less time is spent sick from unclean water and less time is spent fetching water the individual is less likely to fall behind in school or drop out (Wolf et al. 2014). Primary/Secondary school enrollment is a standard measurement of education in the developing world (Smith, 2010). This paper will look at whether an individual is enrolled in primary (ages 6-12) or lower secondary (ages 12-14) school.

Regarding health, Nicaragua shows a high incidence of child death due to diarrhea (WHO 2014). The primary cause of diarrhea is unclean water. Wells, lakes, and rivers are the primary sources of water when piped water is not used. These sources have been shown to be 10 times more likely to be unclean water sources. The hypothesis is that water poverty will impact health primarily through the means of clean water via pipes. The outcome of interest here is how the probability of being enrolled in primary/secondary school, being sick, or the number of hours worked is affected by water access, as seen in Equation 4.

Education

The results of Equation 3 for school enrollment are shown in Table 3.

Table 3. Education

	Primary(age 6-12)	Secondary(age 12-14)	Edu. Att.(age6-14)
Piped Water	0.155**	0.197**	1.438***
	(0.0636)	(0.0850)	(0.237)
Years of Education(head)	0.00606*	0.00905**	0.0431***
	(0.00340)	(0.00460)	(0.0127)
Extreme Poverty	-0.233***	-0.154***	-0.791***
	(0.0338)	(0.0436)	(0.132)
General Poverty	0.0138	-0.0617*	-0.326***
	(0.0258)	(0.0337)	(0.0987)
Male	-0.0519**	0.0281	-0.324***
	(0.0225)	(0.0288)	(0.0837)
Age	0.0256***	-0.0984***	0.698***
	(0.00655)	(0.00733)	(0.0162)
Age^2	0.00000167	0.0000152	0.000139***
	(0.00000878)	(0.0000116)	(0.0000344)
Have straw roof or similar	-0.140**	0.0745	-0.444*
	(0.0574)	(0.0883)	(0.234)
Dirt Floor	-0.0110	-0.0501	0.280***
	(0.0263)	(0.0357)	(0.102)
One Room home	0.0404	0.0151	-0.123
	(0.0251)	(0.0338)	(0.0939)
Observations	774	864	1148
F-stat	13.5	19.81	28.45
P-value > Chi^2	0.00	0.00	0.00

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Here it is shown that there is a positive and significant relationship between an individual (age 6-12, or 12-14) having piped water and enrollment in primary/secondary school, with piped water access predicting an approximate 11.5% and 18.9% increase in the probability enrollment. As expected, these results estimate an extremely poor individual is significantly less likely to be enrolled in school. An interesting result is that rural males are less likely to be enrolled in primary school than females. This could be due to males being more likely to be engaged in agricultural labor at younger ages. The household head's level of education is a highly positive and significant predictor of school enrollment, as might be assumed.

Health

The health measurement that will be used is whether the individual suffered any type of non-chronic illness during the prior month. Results are included in Table 4, again with standard errors in parentheses. It is first observed that in this model, piped water has a statistically significant effect on health for vulnerable populations.

Table 4. Illness

	Primary Age	Secondary Age	Over 50
Piped Water	-0.154**	-0.131**	-0.123**
	(0.0649)	(0.0650)	(0.0582)
Years of Education(head)	0.00397	0.0142***	0.00590*
	(0.00348)	(0.00358)	(0.00358)
Extreme Poverty	-0.0558	-0.0232	0.0187
	(0.0365)	(0.0342)	(0.0386)
General Poverty	-0.0418	-0.0162	0.0131
	(0.0267)	(0.0263)	(0.0268)
Male	0.00614	-0.0206	-0.0367
	(0.0227)	(0.0225)	(0.0225)
Age	-0.00237	0.00859	0.00310**
	(0.00674)	(0.00568)	(0.00142)
Age^2	0.0000140	0.0000311***	0.0000151
	(0.00000936)	(0.00000904)	(0.00000989)
Have straw roof or similar	-0.0528	0.0260	0.0537
	(0.0640)	(0.0695)	(0.0935)
Dirt Floor	0.00777	0.0753***	-0.0240
	(0.0272)	(0.0274)	(0.0255)
One Room home	-0.0542**	-0.00114	-0.0509*
	(0.0251)	(0.0264)	(0.0292)
Observations	774	864	716
F-stat	18.90	20.74	15.36
P-value > Chi^2	0.00	0.00	0.00

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Hours worked

The decrees in time spent fetching water and increases in child's health results in more time that could be used for other things. To see if some of this time is spent on income-generating activities we look at number of hours worked per week on average. Using Equation 3 again gives the following results in Table 5.

Table 5. Hours worked

	Primary Age	Secondary Age	Over 20
Piped Water	-2.601***	-5.827**	11.68***
	(0.791)	(2.812)	(2.057)
Years of Education(head)	-0.00351	-0.549***	-0.168
	(0.0424)	(0.155)	(0.112)
Extreme Poverty	-0.210	-1.640	-5.124***
	(0.445)	(1.480)	(1.299)
General Poverty	0.00410	0.426	-3.825***
	(0.325)	(1.139)	(0.858)
Male	1.068***	14.83***	25.77***
	(0.277)	(0.972)	(0.721)
Age	0.347***	2.397***	-0.101***
	(0.0822)	(0.246)	(0.0272)
Age^2	0.000103	-0.00104***	-0.00146***
	(0.000114)	(0.000391)	(0.000275)
Have straw roof or similar	-0.451	-0.936	-1.143
	(0.782)	(3.010)	(2.400)
Dirt Floor	-0.108	1.460	0.676
	(0.332)	(1.187)	(0.865)
One Room home	-0.563*	-2.171*	0.567
	(0.306)	(1.145)	(0.854)
Observations	774	864	3063
F-stat	18.90	20.74	71.58
P-value > Chi^2	0.00	0.00	0.00

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Having piped water is a significant determinate of working hours. Interesting that those of school age decrease and adults in the household increase. As the children are less likely to be sick or having to fetch water it seems that most of this time then is not used on working. Based on the results from Table 3 we can assume that they are spending more time in school than their counterparts in lieu of working. Also, the impact is larger for secondary school age (12-14) as school is assumed to have a larger opportunity cost for this group. We see that adults in the household spend more time working when there is piped water access. Also, being male had a significant effect on determining number of hours worked.

Table 6. Placebo

	Edu. (over 50)	Illness (20-50)	Chronic Illness (6-12)
Piped Water	0.237	0.00286	0.0646
	(0.408)	(0.0395)	(0.154)
Years of Education(head)	0.728***	0.000717	0.00846
	(0.0250)	(0.00212)	(0.00945)
Extreme Poverty	-0.453*	-0.0202	0.129
	(0.270)	(0.0253)	(0.102)
General Poverty	-0.249	-0.0219	0.0812
	(0.187)	(0.0164)	(0.0706)
Male	0.309**	-0.00454	-0.0739
	(0.157)	(0.0138)	(0.0593)
Age	-0.0913***	0.000345	0.0102***
	(0.00992)	(0.000871)	(0.00374)
Age^2	0.000537***	0.0000214***	0.0000750***
	(0.0000692)	(0.00000525)	(0.0000261)
Have straw roof or similar	-0.220	-0.00634	0.130
	(0.655)	(0.0432)	(0.247)
Dirt Floor	-0.0709	0.0187	-0.184***
	(0.178)	(0.0168)	(0.0673)
One Room home	-0.202	-0.00811	-0.0684
	(0.204)	(0.0160)	(0.0770)
Observations	716	2301	774

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Robustness

We show in Table 6 that our identification strategy does not result in us predicting direct impacts of piped water access for outcomes where this is not plausible. We next examine the sensitivity of our consumption results in the inclusion of municipal controls. To our original specifications, we add variables to control for the male primary school enrollment ratio and paved road access to the community. The results are presented in Table 7, along with those of the first stage regressions Table 8. To summarize, coefficients on the piped water access dummy are essentially unaltered. This gives us confidence that municipal level unobservable were not driving our findings.

Table 7. Various controls Omitted

	Consum.	Consum.	Consum.	Consum.
Piped Water	3234.38***	2321.97***	2405.27***	2399.67**
	(1121.73)	(879.28)	(909.91)	(1149.82)
Years of Education(head)		574.04***	498.40***	251.22***
		(48.82)	(47.48)	(63.31)
Extreme Poverty		-18980.70***	-18158.93***	-16838.26***
		(455.07)	(452.67)	(587.27)
General Poverty		-14508.61***	-13905.52***	-11740.33***
		(329.88)	(330.00)	(468.11)
Male		385.51	387.54	244.57
		(290.85)	(287.54)	(401.06)
Age		52.79***	53.16***	15.04
		(8.21)	(8.12)	(12.00)
Age^2		0.35***	0.20*	-0.13
		(0.11)	(0.11)	(0.16)
Have straw roof or similar			1741.88**	639.03
			(792.60)	(1111.88)
Dirt Floor			-3065.30***	-996.66**
			(349.87)	(457.03)
One Room home			-974.85***	-660.83
			(341.56)	(446.48)
Male Primary Enroll. Ratio				-339.92
				(635.60)
Paved				6393.45***
				(770.58)
Observations	4532	4532	4532	4532

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 8. First

	All	6-12	12-14	20+	50+
Years of Education(head)	0.016***	0.013***	0.018***	0.016***	0.010
	(0.002)	(0.005)	(0.005)	(0.003)	(0.007)
Extreme Poverty	-0.063***	-0.057	-0.050	-0.081***	-0.088
	(0.020)	(0.052)	(0.047)	(0.029)	(0.061)
General Poverty	-0.020	0.013	-0.017	-0.031	-0.018
	(0.014)	(0.040)	(0.037)	(0.020)	(0.044)
Male	-0.013	-0.027	0.000	-0.020	-0.031
	(0.012)	(0.033)	(0.031)	(0.017)	(0.037)
Age	0.000	-0.015	-0.006	0.001	0.001
	(0.000)	(0.010)	(0.008)	(0.001)	(0.002)
Age^2	0.000	-0.000	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Have straw roof or similar	-0.081**	-0.059	-0.123	-0.070	-0.058
	(0.033)	(0.085)	(0.085)	(0.051)	(0.137)
Dirt Floor	-0.125***	-0.095**	-0.170***	-0.129***	-0.126***
	(0.014)	(0.038)	(0.035)	(0.019)	(0.041)
One Room home	-0.051***	-0.072*	-0.059	-0.054**	-0.087
	(0.015)	(0.039)	(0.038)	(0.022)	(0.054)
Population Density(log)	0.120***	0.123***	0.120***	0.124***	0.129***
	(0.006)	(0.016)	(0.014)	(0.008)	(0.019)
Distance to Nearest Public School	-0.023***	-0.026***	-0.013*	-0.028***	-0.037***
	(0.003)	(0.008)	(0.007)	(0.005)	(0.012)
Percentage of Department Tree Cover	-0.204***	-0.137	-0.335**	-0.216***	-0.210
	(0.054)	(0.152)	(0.135)	(0.075)	(0.157)
Mean slope Gradient of Municipality	-0.004***	-0.005*	-0.005**	-0.004***	-0.010***
	(0.001)	(0.003)	(0.003)	(0.001)	(0.003)
Observations	4532	774	864	3063	716

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

To summarize the impact of piped water access on consumption is statistically the same in the model with no control variables and only instrumentation for piped water access. This model also passes the overidentification test. Neither historic population density, tree coverage, nor land slope seems to have direct effects on consumption in rural Nicaragua.

Conclusion

Water poverty in the developing world is a factor in nearly all of the human development indicators. Nicaragua is one area of the world with high levels of economic water scarcity and relatively low levels of human development. This paper investigates how water poverty impacts consumption and three key human development indicators that might be causing pathways in Nicaragua: education, health, and hours worked.

Using instrumental variables, it was found that piped water had a significant positive effect on consumption. It is believed that this was due to the positive effects resulting from primary/secondary school enrollment, improved health, and results in a change in household labor distribution. These results are significant as they show the important role that water plays in achieving the primary goals of policymakers in developing countries: increasing education, improving health outcomes, and increasing income levels. The increase in education and reduction in illnesses seems to have resulted in a higher likelihood of working more. This is a significant finding that shows that income does not just lead to better access to water, but that improved water access can also lead to better income.

References

Brown, Joe, et al. "Relative Benefits of on-Plot Water Supply over Other 'Improved'sources in Rural Vietnam." Tropical medicine & international health 18.1 (2013): 65-74. Print.

Cairncross, Sandy. Domestic Water Supply in Rural Africa. Institute of Commonwealth Studies, 1987. Print.

Curtis, Valerie, et al. "Potties, Pits and Pipes: Explaining Hygiene Behaviour in Burkina Faso." Social science & medicine 41.3 (1995): 383-93. Print.

Dinkelman, Taryn. "The Effects of Rural Electrification on Employment: New Evidence from South Africa." American Economic Review 101.7 (2011): 3078-108. Print.

ESRI Data and Maps (2014). ARCGIS10. Redwood, California: Environmental Systems Research Institute.

Gibson, Mhairi A, and Ruth Mace. "An Energy-Saving Development Initiative Increases Birth Rate and Childhood Malnutrition in Rural Ethiopia." PLoS Medicine 3.4 (2006): e87. Print.

Grogan, Louise, and Asha Sadanand. "Rural Electrification and Employment in Poor Countries: Evidence from Nicaragua." World Development 43 (2013): 252-65. Print.

Hunter, Paul R. "Household Water Treatment in Developing Countries: Comparing Different Intervention Types Using Meta-Regression." Environmental science & technology 43.23 (2009): 8991-97. Print.

INIDE (2014). Encuesta de Medición de Nivel de Vida - 2014. Technical report, Instituto Nacional de Información de Desarrollo, Managua, Nicaragua.

INIDE (2006). VIII Censo de Población y IV de Vivienda. Technical report, Instituto Nacional de Informaci{ó}n de Desarrollo, Managua, Nicaragua.

INIDE (2011). Aspectos Metodológicos. Technical report, Instituto Nacional de Informaci{ó}n Desarrollo, Managua, Nicaragua.

Malik, Khalid. Human Development Report 2014: Sustaining Human Progress: Reducing Vulnerabilities and Building Resilience. United Nations Development Programme, New York, 2014. Print.

Montgomery, Maggie A, and Menachem Elimelech. "Water and Sanitation in Developing Countries: Including Health in the Equation." ACS Publications, 2007. Print.

Ravallion, Martin. "Does Undernutrition Respond to Incomes and Prices? Dominance Tests for Indonesia." The World Bank Economic Review 6.1 (1992): 109-24. Print.

Schmidt, Wolf-Peter, and Sandy Cairncross. "Household Water Treatment in Poor Populations: Is There Enough Evidence for Scaling up Now?" Environmental science & technology 43.4 (2009): 986-92. Print.

Shaheed, Ameer, et al. "Why? Improved? Water Sources Are Not Always Safe." Bulletin of the World Health Organization 92.4 (2014): 283-89. Print.

Smith, Alan. "The Influence of Education on Conflict and Peace Building, Background Paper Prepared for the Education for All Global Monitoring Report 2011 the Hidden Crisis: Armed Conflict and Education, Paris: Unesco." (2010). Print.

Sorenson, Susan B, Christiaan Morssink, and Paola Abril Campos. "Safe Access to Safe Water in Low Income Countries: Water Fetching in Current Times." Social science & medicine 72.9 (2011): 1522-26. Print.

Subaiya, Saleena, and Sandy Cairncross. "Response to Wang and Hunter: A Systematic Review and Meta-Analysis of the Association between Self-Reported Diarrheal Disease and Distance from Home to Water Source." The American journal of tropical medicine and hygiene 84.3 (2011): 504-04. Print.

Tumwine, James K, and KK Munguti. 30 Years of Change in Domestic Water Use & Environmental Health in East Africa: Uganda Country Study. Vol. 2: IIED, 2002. Print.

UNICEF (2018). Estimates of child cause of death, Diarrhoea 2018 UNICEF WHO-CHERG. 'Estimates for child causes of death 2000–2013. 2014. http://www.who.int/healthinfo/global_burden_disease/estimates_child_cod_2013/ en/.

Wang, Xia, and Paul R Hunter. "A Systematic Review and Meta-Analysis of the Association between Self-Reported Diarrheal Disease and Distance from Home to Water Source." The American journal of tropical medicine and hygiene 83.3 (2010): 582-84. Print.

Wolf, Jennyfer, et al. "Systematic Review: Assessing the Impact of Drinking Water and Sanitation on Diarrhoeal Disease in Low- and Middle-Income Settings: Systematic Review and Meta-Regression." Tropical Medicine & International Health 19.8 (2014): 928-42. Print.

Zwane, Alix Peterson, and Michael Kremer. "What Works in Fighting Diarrheal Diseases in Developing Countries? A Critical Review." The World Bank Research Observer 22.1 (2007): 1-24. Print.

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American statistical Association, 105*(490), 493-505.

Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science, 59*(2), 495-510.

Abadie, A., & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque Country. *American economic review, 93*(1), 113-132.

Act, Recovery. "The American Recovery and Reinvestment Act of 2009." (2011).

Agusa, T., Kunito, T., Fujihara, J., Kubota, R., Minh, T. B., Trang, P. T. K., . . . Tanabe, S. (2006). Contamination by arsenic and other trace elements in tube-well water and its risk assessment to humans in Hanoi, Vietnam. *Environmental Pollution, 139*(1), 95-106.

Aquastat, F. (2016). Water withdrawal by sector, around 2006. In: Food and Agriculture Organization of the United Nations. Last Modified November.

Congressional Budget Office(CBO), Estimated Impact of the American Recovery and Reinvestment Act on Employment and Economic Output in 2017 (February 2018)

Cotton, A., Franceys, R., Hutton, L., Pickford, J., & UNICEF. (1991). *WATSAN 2000; proceedings of the UNICEF Orientation/Training Workshop for Water and Sanitation Staff.* Paper presented at the UNICEF Orientation/Training Workshop for Water and Sanitation Staff.

Dasgupta, P. (2004). Valuing health damages from water pollution in urban Delhi, India: a health production function approach. *Environment and development economics, 9*(1), 83-106.

Devoto, F., Duflo, E., Dupas, P., Parienté, W., & Pons, V. (2012). Happiness on tap: piped water adoption in urban Morocco. *American Economic Journal: Economic Policy, 4*(4), 68-99.

Dreibelbis, R., Greene, L. E., Freeman, M. C., Saboori, S., Chase, R. P., & Rheingans, R. (2013). Water, sanitation, and primary school attendance: A multi-level assessment of determinants of household-reported absence in Kenya. *International Journal of Educational Development, 33*(5), 457-465.

Factbook, C. I. A. "The World Factbook; 2010." *See also: http://www cia gov/library/publications/the-world-factbook, accessed January* 30 (2015).

Fryer, M., Sadeghi, P., Soldano, M., Elías, C., Rogers, D., García, S., & Vargas, S. (2012). *Country Program Evaluation: Nicaragua (2008-2012)*. Retrieved from

Galiani, S., Gertler, P., & Schargrodsky, E. (2005). Water for life: The impact of the privatization of water services on child mortality. *Journal of political economy, 113*(1), 83-120.

Gamper-Rabindran, S., Khan, S., & Timmins, C. (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.

INEC, C. (2006). *Encuesta Nacional de Hogares sobre Medición de Niveles de Vida 2005: Manual del Usuario de la Base de Datos*. Retrieved from

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

Khan, S., Shahnaz, M., Jehan, N., Rehman, S., Shah, M. T., & Din, I. (2013). Drinking water quality and human health risk in Charsadda district, Pakistan. *Journal of cleaner production, 60*, 93-101.

Komives, K., Whittington, D., & Wu, X. (2001). *Infrastructure coverage and the poor: A global perspective* (Vol. 2551): World Bank Publications.

Koolwal, G., & Van de Walle, D. (2013). Access to water, women's work, and child outcomes. *Economic Development and Cultural Change, 61*(2), 369-405.

Kreif, N., Grieve, R., Hangartner, D., Turner, A. J., Nikolova, S., & Sutton, M. (2016). Examination of the synthetic control method for evaluating health policies with multiple treated units. *Health economics, 25*(12), 1514-1528.

LaVanchy, G. T. (2017). When wells run dry: Water and tourism in Nicaragua. *Annals of Tourism Research, 64*, 37-50.

Macinko, J., Guanais, F. C., & De Souza, M. D. F. M. (2006). Evaluation of the impact of the Family Health Program on infant mortality in Brazil, 1990–2002. *Journal of Epidemiology & Community Health, 60*(1), 13-19.

McCauley, N. (2015). Working, Not Walking, for Water.

Nauges, C., & Strand, J. (2017). Water hauling and girls' school attendance: Some new evidence from Ghana. *Environmental and Resource Economics, 66*(1), 65-88.

Nguyen Viet, C., & Vu, T. (2013). The impact of piped water on household welfare: evidence from Vietnam. *Journal of Environmental Planning and Management, 56*(9), 1332-1358.

Organization, W. H. (2015a). *Progress on sanitation and drinking water: 2015 update and MDG assessment*: World Health Organization.

Organization, W. H. (2015b). *Progress on sanitation and drinking water: 2015 update and MDG assessment*: World Health Organization.

Organization, W. H. (2016). *World health statistics 2016: monitoring health for the SDGs sustainable development goals*: World Health Organization.

Organization, W. H. (2017). Protecting children from the harmful impact of food marketing [fact sheet].

Ortiz-Correa, J. S., Resende Filho, M., & Dinar, A. (2016). Impact of access to water and sanitation services on educational attainment. *Water Resources and Economics, 14*, 31-43.

Programme, U. N. D., & Malik, K. (2014). *Human Development Report 2014: Sustaining Human Progress-Reducing Vulnerabilities and Building Resilience*: UN.

Ravallion, M. (1992a). Does undernutrition respond to incomes and prices? Dominance tests for Indonesia. *The World Bank Economic Review, 6*(1), 109-124.

Ravallion, M. (1992b). Poverty Comparisons: A Guide to Concepts ands Methods. In *Living Standards Measurement Working Paper*: World Bank.

Robert N. Kaplan, R. D. A., Camilo E. Garzón, Ana María Linares , Miguel Coronado , Gustavo Martínez , y Cinthya Cordero. (2006). DOCUMENTO CONCEPTUAL DE PROYECTO (DCP) "Programa de Inversiones en Agua Potable y Saneamiento".

Sclar, G., Garn, J., Penakalapati, G., Alexander, K., Krauss, J., Freeman, M., . . . Clasen, T. (2017). Effects of sanitation on cognitive development and school absence: A systematic review. *International journal of hygiene and environmental health, 220*(6), 917-927.

Supply, W., Council, S. C., & UNICEF. (2000). Global water supply and sanitation assessment 2000 report: summary of the report. In *Global water supply and sanitation assessment 2000 report: summary of the report*.

Waddington, H., Snilstveit, B., White, H., & Fewtrell, L. (2009). Water, sanitation and hygiene interventions to combat childhood diarrhoea in developing countries. *New Delhi: International Initiative for Impact Evaluation*.

Wadhwa, D. (2016). *Does access to sanitation and drinking water facilities relate to students' learning outcomes? Evidence from Andhra Pradesh, India*: Georgetown University.

WHO, U. (2010). Joint monitoring programme for water supply and sanitation. *Estimates for the use of improved drinking-water sources. Updated March 2010*.

Who, U. (2010). *Progress on sanitation and drinking-water, 2010 Update*: World Health Organization.

WHO, U., & Mathers, C. (2017). Global strategy for women's, children's and adolescents' health (2016-2030). *Organization, 2016*(9).

World Bank. "World Bank national accounts data, and OECD National Accounts data files." (2017).

Berger RL, Hsu JC. 1996. Bioequivalence Trials, Intersection-Union Tests and Equivalence Confidence Sets. Statistical Science 11(4): 283–302.

Briggs A, Sculpher M. 1998. An introduction to Markov modeling for economic evaluation. PharmacoEconomics 13(4): 397–409.

Carod-Artal, F. J., et al. "Direct cost of cerebrovascular disease during the first year of follow-up." Revista de neurologia 28.12 (1998): 1123-1130.

Demaerschalk, Bart M, et al. "Efficacy of Telemedicine for Stroke: Pooled Analysis of the Stroke Team Remote Evaluation Using a Digital Observation Camera (Stroke Doc) and Stroke Doc Arizona Telestroke Trials."; Telemedicine and e-Health 18.3 (2012): 230-37. Print.

---. "Cost Utility of Hub-and-Spoke Telestroke Networks from Societal Perspective."; The American journal of managed care 19.12 (2013): 976-85. Print.

Earnshaw, Stephanie R., et al. "Cost-effectiveness of patient selection using penumbral-based MRI for intravenous thrombolysis." Stroke 40.5 (2009): 1710-1720.

Fattore, Giovanni, et al. "The social and economic burden of stroke survivors in Italy: a prospective, incidence-based, multi-center cost of illness study." BMC neurology 12.1 (2012): 137.

Gallet, Craig Arthur. "The Impact of Public Health Spending on California STD Rates." International Advances in Economic Research 23.2 (2017): 149-159.

Gandjour A, Stock S. 2007. A national hypertension treatment program in Germany and its estimated impact on costs, life expectancy, and cost-effectiveness. Health Policy 83(2–3): 257–267.

Go, Alan S., et al. "Executive summary: heart disease and stroke statistics—2013 update: a report from the American Heart Association." circulation 127.1 (2013): 143-152.

Hawton, Annie, et al. "Squinting through layers of fog: assessing the cost-effectiveness of treatments for multiple sclerosis." Applied health economics and health policy 11.4 (2013): 331-341.

Hallberg, S., et al. "Healthcare costs associated with cardiovascular events in patients with hyperlipidemia or prior cardiovascular events: estimates from Swedish population-based register data." The European Journal of Health Economics 17.5 (2016): 591-601.

Hervas-Angulo, A., J. M. Cabases-Hita, and T. Forcen-Alonso. "Costs deriving from strokes from a social perspective. A retrospective incidence approach with a follow-up at three years." Revista de neurologia 43.9 (2006): 518-525.

Jäkel, Anne, et al. "A systematic review of economic evaluations in second and later lines of therapy for the treatment of non-small cell lung cancer." Applied health economics and health policy 11.1 (2013): 27-43.

Kirsch, Florian. "A systematic review of Markov models evaluating multi-component disease management programs." Value in Health 18.7 (2015): A502.

Kourlaba, Georgia, Vassilis Fragoulakis, and Nikos Maniadakis. "Economic Evaluation of Clopidogrel in Acute Coronary Syndrome Patients without ST-Segment Elevation in Greece." Applied health economics and health policy10.4 (2012): 261-271.

Lackland, Daniel T., et al. "Factors influencing the decline in stroke mortality." Stroke 45.1 (2014): 315-353.

MacDougall L, Majowicz S, Doré K, Flint J, Thomas K, Kovacs S, Sockett P. 2008. Under-reporting of infectious gastrointestinal illness in British Columbia, Canada: who is counted in provincial communicable disease statistics? Epidemiology and Infection 136: 248–256.

Mar J, Sainz-Ezkerra M, Miranda-Serrano E. 2008. Calculation of prevalence with Markov models: Budget impact analysis of thrombolysis for stroke. Medical Decision Making 28(4): 481–490.

McBride GB. 1999. Equivalence Tests Can Enhance Environmental Science and Management. Australian & New Zealand Journal of Statistics 41(1): 19–29.

Morris, Stephen, et al. "Cost-utility analysis of stenting versus endarterectomy in the International Carotid Stenting Study." International journal of stroke 11.4 (2016): 446-453.

Müllhaupt, Beat, et al. "Modeling the health and economic burden of hepatitis C virus in Switzerland." PloS one 10.6 (2015): e0125214.

Navarrete-Navarro, P., et al. "The societal costs of intracerebral hemorrhage in Spain." European journal of neurology 14.5 (2007): 556-562.

Nelson, Richard E, et al. "The Cost-Effectiveness of Telestroke in the Treatment of Acute Ischemic Stroke." Neurology 77.17 (2011): 1590-98. Print.

Nichols-Larsen, Deborah S., et al. "Factors influencing stroke survivors' quality of life during subacute recovery." Stroke 36.7 (2005): 1480-1484.

Normand ST. 1999. Meta-Analysis: Formulating, Evaluating Combining and Reporting. Statistics in Medicine 18(3): 321–359.

Persson, Josefine, José Ferraz-Nunes, and Ingvar Karlberg. "Economic burden of stroke in a large county in Sweden." BMC health services research 12.1 (2012): 341.

Ramirez, Lucas, et al. "Trends in acute ischemic stroke hospitalizations in the United States." Journal of the American Heart Association 5.5 (2016): e003233.

Sonnenberg FA, Beck JR. 1993. Markov Models in Medical Decision Making: A Practical Guide. Medical Decision Making 13(4): 322–338.

Spieler, Jean-François, Jean-Louis Lanoë, and Pierre Amarenco. "Costs of stroke care according to handicap levels and stroke subtypes." Cerebrovascular diseases 17.2-3 (2004): 134-142.

The United States Environmental Protection Agency. 2016. Cost of Illness Handbook, Washington, DC.

Wong JB, McQuillan GM, McHutchison JG, Poynard T. 2000. Estimating future hepatitis C morbidity, mortality, and costs in the United States. American Journal of Public Health 90(10): 1562–1569.

WHO. (2014). UN-water global analysis and assessment of sanitation and drinking-water (GLAAS) 2014 report: Investing in water and sanitation: Increasing access, reducing inequalities. Geneva: Author. Retrieved from http://www.who.int/water_sanitation_health/glaas/2014/glaas_report_2014/en

Irianti, Sri, Puguh Prasetyoputra, and Tri Prasetyo Sasimartoyo. "Determinants of household drinking-water source in Indonesia: An analysis of the 2007 Indonesian family life survey." Cogent Medicine 3.1 (2016): 1151143.