Journal of Sustainable Development; Vol. 9, No. 4; 2016
ISSN 1913-9063 E-ISSN 1913-9071
Published by Canadian Center of Science and Education

Effect of Improving Housing Conditions on Early Childhood Health in Rural Sudan

Obbey A. Elamin[1]

[1] The British University in Egypt, Cairo, Egypt

Correspondence: Obbey A. Elamin, Department of Economics, The British University in Egypt, El Sherouk City, Suez Desert Road, Cairo 11837, P.O. Box 43, Egypt. E-mail: obbey.elamin@bue.edu.eg; obbey.elamin@outlook.com

Received: March 26, 2016 Accepted: June 7, 2016 Online Published: July 30, 2016

doi:10.5539/jsd.v9n4p96 URL: http://dx.doi.org/10.5539/jsd.v9n4p96

Abstract

Improving housing sector in rural areas is important to improve health status of under-five children. Propensity score matching using nonparametric kernel estimates is used to examine the effect of improving rural structure of houses in rural Sudan and provide them with services like access to clean piped water, sanitation on improving under-five children health. The prevalence of diarrhoea and cough in rural Sudan are used as measures of health outcome and data from the Sudan Household Health Survey in 2010 is used. Our results show that providing houses with piped water can reduce prevalence of diarrhoea and cough by 22 and 24 percentage points, respectively. Gas cooking fuel reduces the prevalence rates by 26 and 29 percentage points, respectively. Construction materials of walls have strong impact on reducing the prevalence of both illnesses. We recommend that the quality of piped water should be observed and maintained in good standard to ensure that clean water is supplies to the household sector. Developing the housing sector in the rural has many advantages in improving early childhood health in Sudan and it should be one of the priorities of the government.

Keywords: under-five, health, rural, development, nonparametric

1. Introduction

Due to the lack of economic growth, rural houses infrastructure has not improved in Sudan for decades. In addition to the poor building materials of the houses, housing sector in rural Sudan is still lacking access to the main services like piped water, clean sanitation and electricity, which make living condition difficult and unhealthy for individuals and particularly under-five children. The effect of in-house living environment on under-five children health is direct and sever, poor housing condition contribute negatively to under-five children health. Some of under-five children sicknesses are preventable or can be substantially reduced if sufficient fund is invested in improving housing conditions of the families living rural areas. A good living house that is well constructed, provided with clean water source, uses clean cooking fuel, supported with health facilities like flush toilet and constructed from strong solid materials like cement or wood, provide healthy living environment for children and reduces risks of child illnesses. However, it is rare in rural areas to find houses in qualities that provide such living standards. Accordingly, some diseases like diarrhoea and cough among under-five children are difficult to be prevented.

The major economic activity in rural areas in the country is agriculture. Houses structure and living conditions are highly influenced by the natural of the economic activity in the region. Sudan has not invested in improving the housing sector in rural areas for a number of years. Piped water was available to only 61.3% of the rural population in 1990, and dropped to 50.2% in 2010, according to the World Bank statistics. From the same source we find that only 18.3% of the rural population had access to improved sanitation facilities in 1990 this dropped to 13.4% in 2010. In addition to that only 15% of the rural population have access to electricity at home. The statistics clearly indicate decreasing trends in housing sector quality of living in the rural of the country. Children are affected by houses condition by being drinking unclean water, exposed to in-door pollution from firing wood of charcoal for cooking or lighting or from poor sanitation system. On the other hand, it is evident that Sudan is making progress in reducing under-five mortality rate, but that mostly by investing in immunization of the children and focusing on mother orientation in an attempt of developing healthy environment at home. Under-five mortality rate of has fallen from 127.5 per 1000 in 1999 to 80.2 per 1000 in year 2010. However, Influenza, Pneumonia, diarrhoeal diseases and lung disease are some of the leading causes of under-five death. Life expectancy at birth is 60.0 and 63.6 years for males and females, receptively, improves to 65.8 and 69.1, receptively, at age 5, which shows that a high risk of dying at infancy and up to age 5 years old is facing children in early childhood in the country. Educated mother can take actions such as treating drinking water before use to prevent child illness, and maintaining the house clean as much as possible to develop healthy in-house environment for the children. Generally, however, both female education and housing sector are poorly developed in rural areas in Sudan. Early childhood health is a matter of concern for the Government and the international organizations that operating in the country.

In this research we use nonparametric econometrics propensity score matching method to estimate the possible shifts in under-five children health if houses structure and housing sector services are improved. Econometric treatment evaluation technique is used to estimate the change in the prevalence of child illness in response to a set of treatments that include; access to piped water, type of cooking fuel, household treatment of drinking water, dwelling structure materials. The research attempts to estimate the effect of improving houses condition on under-five children heath by comparing prevalence of diarrhoea and cough among children living in poor structured houses with prevalence among those living in relatively better structured houses. Diarrhoea and cough among under-five year old children are two risky health conditions that are widely spread in rural Sudan. Diarrhoea is fatal for children if not treated appropriately. Cough is usually an indicator of health problem in the child respiratory system. The research uses cross section data from the Sudan Household Health Survey 2010 (SHHS2010). A sample from all the states in Sudan in rural and urban areas has been covered, in which the characteristics of the dwelling of each household in the sample are collected. SHHS2010 survey is ideal for our study objectives in terms that among other household surveys in Sudan it provides sufficient information about both child illness and household dwelling condition and facilities.

Our results show that piped water and solid dwelling walls building material have the strongest impact on reducing the prevalence of both diarrhoea and cough, however, maintaining the quality of piped water is crucial for sustaining this link. Walls that are built from palm or sod host many types of bacteria and dirt, children get in touch with them directly and expose to unhealthy risks, which is the reason that solid material wall show significant effect in reducing diarrhoea. The paper is organised as follows; Section 2 presents the related literature and the key findings on how household characteristics affect under-five children health. Description of the estimation techniques and demonstration of the econometric treatment evaluation using PSM method is provided in Section 3, Sections 4 and 5 present the data and discusses the results, respectively. The last section, Section 6, present the conclusions of the research. The software package "R" is used in the analysis with all codes of the kernel estimators, ATE and ATT developed by the author.

2. Related Literature

Aspects of improving childhood health are strongly related to healthy indoor environment at the household as well as to family socio-economic characteristics and living standard. Relationship between child health outcomes and economic and social factors is discussed by (Schultz 1984). Risks of indoor pollution on children health are illustrate by (Zhang & Smith 2003), (Emmelin 2007), (Choi et al. 2010) and (Oluwole et al. 2012) among many others. The link between unclean water and children diarrhoea is well established and thoroughly discussed in the literature, see (Bryce J Shibuya K, Black RE and the WHO Child Health Epidemiology Reference Group. 2005) (Wardlaw, T., Salama, P., Brocklehurst, C., Chopra, M., Mason 2010), (Kumar & Subita 2013), (Alemayehu et al. 2014), (Qasim et al. 2014), (Upadhyay et al. 2015), (Currie & Rossin-Slater 2013), (Acharya et al. 2015), (Bern et al. 1992) and (Black RE, Morris SS 2003). Effect of sanitation system and treatment of waste is discussed by (Esrey et al. 1991), (Edejer et al. 2005), (Waddington et al. 2009) and (Waddington & Snilstveit 2009) among many others, flush toilets have pronounceable effect on reducing diarrhoea risk for children. Effect of poor housing and environmental conditions in children health is considered early by many researcher including (R.M. & D’Souza 1997) and (Ferng & Lee 2002). The positive association of the quality of indoor environment on health is well documented in the literature in(Ferng & Lee 2002) and (Ferng & Lee 2002).

This paper is closely related to a vast literature in economics that discusses to role of indoor air pollution (IAP) on children health. For piped water, however, the paper highlights the debate in the literature about the effect of piped water in the prevalence of diarrhoea. In the literature, positive, negative or insignificant effect for the piped water on children’s diarrhoea are all well demonstrated and explained. The key researches that demonstrate the negative effect of piped water on prevalence of diarrhoea are led by (Gross et al. 1989) which shows, using a sample from Brazil, that prevalence of diarrhoea decreased by about 25 percentage points with improved water supply and sanitation. (Fewtrell et al. 2005) show that piped water and improving water quality at the point of use, i.e. treating drinking water, make pronounceable reduction in the prevalence of diarrhoea. An economic model and an estimation framework is developed by (Jalan & Ravallion 2003), using a sample from rural India, which show that piped water reduces both prevalence of diarrhoea and duration of illness for under-five children. Using propensity score matching method, (Jalan & Ravallion 2003) find that diarrhoea for under five children in rural India is significantly less on average in households with piped water supplied.(Yu 2011), on the other hand, finds strong relationship between indoor pollution from cooking and heating fuel in respiratory health for under-five children in rural (Zhang 2012) considers the same topic in China and reaches the same results using ATE method.

Negative effect for piped water on bad health outcomes, like diarrhoea incidences, is attained when the quality of the treatment method and the facilities of the network are carefully observed and maintained. In many countries, however, piped water is found insignificant in affecting diarrhoea or even positively increasing health risks. This point is raised in the literature by many researchers including (Semenza et al. 1998), who find that, in the case of Uzbekistan, lack of appropriate maintenance to the network and the distribution system of piped water makes it a source of disease transmission. (Lechtenfeld 2012) finds, using spatial econometrics, that broken pipes and interruptions of water supply cause most of the water pollution in Yemen, and estimates that risk of child diarrhoea increases by 4.6 percentage points. (Cotruvo & Trevant 2000) argue that inadequate access to training in developing countries in addition to management and insufficient finance are some of the main factors that contribute to the deterioration of piped water industry in rural areas.

In Sudan sort literature, however, including (Haroun et al. 2010), (Sinha & Srivastava 1993) and (Al Mubarak 2006) addresses the factors that affecting prevalence of under-five children illnesses like diarrhoea in small local districts in the country. The importance of this study is that it is the first in Sudan that uses econometric average treatment effect (ATE) and average treatment effect on the treated to examine the effect of household living house construction condition and characteristics on under-five children health. A few research papers have been found in this topic in Sudan, which are mostly following non-econometric approach, for example (Siziya et al. 2013).

3. The Econometric Technique

In economics researchers might be interested in studying the effect of exposure to some kind of treatment from non-randomised data, where conducting experiment is complicated. The outcome from being exposed to a (treatment) is examined against the outcome before or without the treatment (the control). Units, like individuals or households, are more common to be self-selected to the treatment. Selection, accordingly, might be dependent on factors related to the units themselves, that might be observed or unobserved to the researcher. Each individual is observed once either with or without the treatment. Using the means to approximate the outcome of each group is not advisable, since the groups are usually differ even in the absence of the treatment, which known as the selection bias. As shown by (LaLonde 1986) and (Fraker & Maynard 1987), replicating experimental results with conventional estimators in econometrics fails. (Rubin 1974) and (Rubin 1977) developed a causal model based on comparing the groups using counterfactual outcomes and matching individuals in the treatment and the control groups. Participant in the treatment are matched with non-participant(s) with the closest given pretreatment characteristics. This approach is discussed also by (Heckman 1998).

Methods of matching the observation are thoroughly discussed in micro-econometric literature and textbooks including, (Wooldridge 2001), (Imbens & Wooldridge 2009), (Caliendo et al. 2005) and (Cameron & Trivedi 2005). To illustrate the development of the model, let Y_i denotes the outcome variable for household i, $i = 1,2,\ldots,n$. Y_i could be continuous or binary. We are interested in studying the effect of a treatment dummy variable T_i on Y_i. The treatment T_i is assigned a value 1 if i is exposed to the treatment and 0 otherwise. The two potential outcomes are denoted by $\{Y_i(0), Y_i(1)\}$ for the outcome when not exposed to the treatment and the outcome when exposed to the treatment, respectively. It is not likely to observe both outcomes, so, for household i either $Y_i(0)$ or $Y_i(1)$ is observed. Thus, the observed outcome in the sample is

$$Y_i = T_i Y_i(1) + (1 - T_i) Y_i(0).$$

The outcome is assumed to be independent of the treatment, which is known as the exogenous condition, or the independence condition, denoted as $Y_i(0), Y_i(1) \perp T_i$. This assumption is automatically satisfied via randomised experiments if available. In non-experimental data using estimated value of the outcome of interest in relation to relevant pretreatment characteristics of all units that are expressed in the matrix X_i provides good approach for replicating experimental data. The key underlying assumption is that conditioning on X_i both T_i and Y_i are independent, which is known as the unconfoundedness (or conditional dependence) assumption, which is crucial for identifying the treatment effect. The unconfoundedness assumption is denoted as $Y_i(0), Y_i(1) \perp T_i|X_i$, (Rosenbaum & Rubin 1983), (Wooldridge 1999) and (Imbens 2013).

Researcher's interest focuses on two measures, the average treatment effect (ATE), defined as $\tau = E[Y_i(1) - Y_i(0)]$, which is the expected mean effect on a randomly drawn sample unit. More interested measure, however, is the average treatment effect of the treated (ATT), defined as $\tau = E[Y_i(1) - Y_i(0)|T_i = 1]$, which estimates the expected mean effect for the units that actually exposed to the treatment. The conditional estimators of ATE and ATT are available with suitable independent variables set, X_i. The average treatment effect conditional on X_i is defined as

$$\tau(x) = E[Y_i(1)|T_i = 1, X_i = x] - E[Y_i(0)|T_i = 0, X_i = x] \tag{1}$$

The ATE can be obtained by taking the expectation $\tau = E[\tau(x)]$. From the sample this is obtained by $\tau = E[\tau(X_i)]$. Defining

$$Y_i = g(X_i, T_i) + e_i,$$
$$= g(X_i, T_i = 0) + \tau(X_i)T_i + e_i.$$

Then $\tau(X_i) = \frac{cov(Y_i,T_i|X_i)}{var(T_i|X_i)}$. The estimator is then has the formula

$$\tau = E\left[\frac{(T_i - P(T_i|X_i)Y_i)}{var(T_i|X_i)}\right]. \tag{2}$$

For the identification of the estimator further assumptions need to be imposed, the ignorability in mean assumption and the overlap assumptions. The ignorability in mean assumption states that $E[Y_i(g)|X_i, T_i] = E[Y_i|X_1]$ for $g = 0,1$, which means

$$\tau(x) = E[Y_i(1) - Y_i(0)|X_i = x], \tag{3}$$

the overlap assumption on the other hand, states that for any set of covariates there is a chance to see both participants and non-participants units in the sample, that is, $0 < P(T_i|X_i) < 1$. Under those assumption it is possible to estimate the terms in the right hand side in Eq 1. To prevent the curse of dimensionality problem that is associated with large dimension X_i, one approach suggests using a function that summaries the information in X_i. The conditional probability of T_i given X_i is used, which is denoted as the propensity score. The outcome of participants is matched with the outcome of the closest non-participant(s) in the sample, which is found to be appropriately replicating experimental data results as illustrated by (Dehejia & Wahba 1999) and (Dehejia & Wahba 2002). Matching method based on the propensity score is defined as the propensity score matching (PSM). Either parametric or nonparametric econometric techniques can be used in the estimation. The advantage of nonparametric estimation on the propensity score, as illustrated in (Li, Maasoumi, et al. 2009), is that it does not enforce functional form to the conditional probability estimating the propensity score, but approximate it arbitrary from the data, which on the other hand, incorporates very low risk of model misspecification. Illustration about kernel nonparametric estimation approach is available in the literature in (Ullah 1988), (Härdle & Linton 1994), Ullah (ULLAH 2002), (Lordo 2005), (Li & Racine 2007) and (Li, Maasoumi, et al. 2009) among many other sources.

3.1 Kernel Estimator of the Treatment Effect

The set of independent variables, X_i, is assumed to be mixed of continuous and discrete variables types as specified by (Li, Racine, et al. 2009). Let q be the number of continuous variable and p be the number of discrete variables. (Li, Racine, et al. 2009) suggest an estimator of ATE and ATT that utilises a kernel estimated propensity score using smoothing method for mixed data type. The propensity score is estimated using local constant regression method and takes the form

$$\hat{p}(X_i) = \frac{\sum_{i=1}^{n} T_i W_{\mathbf{h},\gamma,ij}}{\sum_{i=1}^{n} {}_{,}W_{\mathbf{h},\gamma,ij}}, \tag{4}$$

where $W_{\mathbf{h},\gamma,ij}$ is the kernel product term. Each continuous independent variable, X_s^c, is smoothed with a continuous kernel functions, $k(\cdot)$ and a bandwidth, h_s, where $s = 1,2,\dots,q$ and the superscript c denotes that the variable is continuous. Similarly, each discrete independent variable, X_r^d, is smoothed using a discrete kernel functions $l(\cdot)$ and a bandwidth γ_r, where $r = 1,2,\dots p$. The kernel product for observation i measures the weighted distance between X_i and other observations in the sample, j, is given as

$$W_{\mathbf{h},\boldsymbol{\gamma},ij} = \prod_{s=1}^{q} h_s^{-1} k\left(\frac{X_{is}^c - X_{js}^c}{h_s}\right) \prod_{r=1}^{p} l\left(X_{ir}^d, X_{jr}^d, \gamma_r\right).$$

Using the estimator in Eq 4 in Eq 2, the kernel estimator of the average treatment effect is given as

$$\hat{\tau} = n^{-1} \sum_{i=1}^{n} \frac{(T_i - \hat{p}(X_i) Y_i)}{\hat{p}(X_i)(1-\hat{p}(X_i))} M_{ni}, \tag{5}$$

where M_{ni} is a trimming function that trims out the observations at the boundaries.

The bandwidths are estimated using the least-squares cross validation method by minimising the objective function

$$n^{-1} \sum_{i=1}^{n} [T_i - \hat{p}_{-i}(X_i)]^2 S(X_i), \tag{6}$$

with respect to $\mathbf{h}, \boldsymbol{\gamma}$. The bandwidth vectors are $\mathbf{h} = [h_1, h_2, \dots, h_q]$ and $\boldsymbol{\gamma} = [\gamma_1, \gamma_2, \dots, \gamma_p]$, where the bandwidths satisfy that $0 < h_s$ if the variable is continuous, and $0 < \gamma_r < 1$ if the variable is discrete. Very large bandwidth indicates that the corresponding variable is irrelevant to the probability estimating the propensity score. $S(X_i)$ is a trimming function that drops out the observations near the boundaries, to prevent estimating bandwidths that are affected by extreme values in X, in this research $S(X_i)$ is fixed at 1%, in contrast with M_{ni} which is fixed at the rule of thumb value, 2% to handle weak covariates overlap problem as will be interpreted in the next section. Li et al. (2009) suggest smoothing continuous variables by second order kernel function when $q \leq 3$ and by higher order kernel when $q \geq 4$. In our application $q = 3$, so, the Second-Order Gaussian kernel function, which takes the form

$$w(z) = (2\pi)^{-\frac{1}{2}} \exp(-z^2/2), \tag{7}$$

is used. Note that from the sample $z_{ij} = \frac{X_{ri}^c - X_{rj}^c}{h_r}$, is a measure of the distance between the i^{th} and the j^{th} observations for the variable X_r^c. For discrete unordered variables Aitchison and Aitken (1979) kernel function is used, which has the form

$$l\left(X_i^d, x^d, \gamma\right) = \begin{cases} 1-\gamma & \text{if } X_i^d = x^d \\ \gamma/(c-1) & \text{if } X_i^d \neq x^d \end{cases} \tag{8}$$

where c is the support of the discrete variable X^d. For the ordered variables the Geometrical Kernel function of (Wang and Van Ryzin 1981) is used, which takes the form

$$l\left(X_i^d, x^d, \gamma\right) = \begin{cases} 1-\gamma & \text{if } X_i^d = x^d \\ \frac{1}{2}(1-\gamma)\gamma^{\left|X_i^d - x^d\right|} & \text{if } X_i^d \neq x^d \end{cases} \quad \gamma \in [0,1]. \tag{9}$$

The ATE estimator based on kernel estimated propensity score as shown by (Li and Racine, Nonparametric Econometric Methods 2009), is semiparametrically and asymptotically efficient among the class of feasible ATE estimators.

3.2 Testing the Presence of Treatment Effect

The bootstrap method is used to test the hypothesis that there is no treatment effect, $H_0: \tau = 0$. The empirical distribution of $\hat{\tau}$ under the null is approximated as the follows

- From the outcome variable $\{Y_i\}_{i=1}^{n}$ select n sample units with replacement and denote it as $\{Y_i^*\}_{i=1}^{n}$.
- Use the bootstrap sample $\{Y_i^*, T_i, X_i\}$ to estimate the bootstrap value of $\hat{\tau}^*$. Both T_i and X_i are taken from the original sample. In this bootstrap sample the link between

the outcome and the treatment is broken since the bootstrap process is being performed under the null hypothesis.

- Repeat the above two steps large number of time, say B time, to approximate the distribution of τ^* under the null.
- Sort $\{\hat{\tau}_b^*\}_{b=1}^{B}$ in an ascending order and compute the α's percentile, $\hat{\tau}_\alpha^*$. Reject H_0 if at α level of significance if $\hat{\tau} > \hat{\tau}_\alpha^*$.

4. The Data

Sudan Household Health Survey in year 2010, SHHS2010, is conducted by the Government of National Unity (GONU) before the separation of the country to Sudan and South Sudan. The technical work of the survey is provided by Government of National Unity (GONU) and the Government of Southern Sudan (GOSS) with collaboration with UNICEF. The survey covers 14778 households in the northern states of Sudan before the separation of the south. Those are the states of what is now known as the Republic of the Sudan. The survey covers 9369 household in the southern states, which are excluded from the analysis in this research, since those are the states that currently form the independent Republic of the South Sudan. The number of the households in the survey that are living in rural areas is 10299, which counts about 67.6% of the respondents in the north.

The proportion of the rural household in the Khartoum State in the survey is only 17.3%. Khartoum State has low proportion of rural population compared with other states in the country, and has relatively large geographical size of its urban compared to the total size of the state. The urban of Khartoum is more developed than the urban of other states, it comprises the political capital and the centre of trade and industry in the country. Economically Khartoum State is advanced compared to other states. The rural of Khartoum is enjoying this advancement since it is small in size and easily connected to urban through active transportation network. Accordingly, rural Khartoum has different characteristics than the rural of other states in Sudan, which would bias our ATE and ATT estimator due to weak overlapping problem. Thus, we drop the rural of Khartoum state from the analysis to make the data more homogeneous.

SHHS2010 collected the data using a number of questionnaires, a questionnaire for under-five children, a questionnaire for women aged 15-45, a questionnaire for men and a general questionnaire to the household. If some of the inhabitants in a randomly chosen household are under five years old, the children questionnaire is passed to the mother, or to the responding adult in the household. Information about each under-five child in the household is collected. A question about whether the named under-five child had suffered from diarrhoea is asked, followed by a question whether the child had suffered from cough. In case of stating a positive answer, the respondent is asked further questions about the sickness and the actions that the parents did for treating the illness. From the household general questionnaire, that is used for all households in the survey, information about the house condition, building materials of the walls, the roof, the floor, sanitation are asked. This in addition to questions about the household economic condition, access to piped water, type of the cooking fuel, whether the household is treating drinking water. This research combines the information from those two questionnaire.

After choosing the households including under-five chil(ren) only, and removing observations with missing values the sample size reduces to 5850 households only. The estimated mortality rate for under-fives from the survey is 105 per 1000. About 69.38% of the households have more than one under five child. The proportion of household with at least one observed case of diarrhoea and at least one case of cough are 37.7% and 41.1%, respectively Among the households including more than one child 28.4% have two or more sick children with diarrhoea and 42.6% have two or more sick children with cough.

The list of variable that are used in the model are shown in Table 1. Total household income is not measured directly in the survey. The variable listed in Table 1 are used in the propensity score estimation. In the nonparametric kernel estimator only the variables lexpeduc, lexpfood and oldestM01 are specified as continuous variables, other variable are defined as discrete and are smoothed using either discrete ordered or discrete unordered kernel functions.

Table 1. Variables names

Name	label
lexpeduc	log expenditures on education
lexpfood	log expenditures on food
oldestM01	age of oldest woman in the household
nunder5	number of under 5 years old in the household
nhh	number of household members
HH12	number of women in age group 15-46
HC02	number of rooms in the household
HC8B	owning radio dummy
foodaid	Food aid dummy
mosqnet	mosquito net dummy
hhmale	male head of household dummy
headec2	education of the head of the household is primary
headec3	education of the head of the household is secondary of higher

Table 2. Summary Statistics of the overall sample and by access to piped water

		Mean	SE
Outcomes			
	diarrhoea	0.377	0.006
	cough	0.411	0.006
Covariates			
	lexpeduc	1.962	(029)
	lexpfood	5.697	(0.008)
	oldestM01	37.888	(0.219)
	nunder5	1.612	(0.009)
	nhh	6.253	(0.033)
	HH12	1.259	(0.009)
	HC02	3.344	(0.022)
	HC8B	0.518	(0.007)
	foodaid	0.064	(0.003)
	mosqnet	0.642	(0.006)
	hhmale	0.866	(0.004)
	headec2	0.278	(0.006)
	headec3	0.138	(0.005)
Treatments			
	Piped water	0.158	(0.005)
	Gas cooking fuel	0.199	(0.005)
	Treat drinking	0.163	(0.005)
	Treat piped	0.041	(0.003)
	Sand floor	0.946	(0.003)
	Flush toilet	0.015	(0.002)
	Solid walls material	0.570	(0.006)
	Solid roof material	0.080	(0.004)

1. Sample size is 5850.

Table 2 shows the summary statistics of the treatments, the variables that are used in the propensity score estimation and outcome variables and the covariates are shown in three different panels. The data shows that houses are poorly constructed in the rural areas and lack important facilities like water source supply and flush toilet. Piped water is available to only 15.8% of the rural households, gas cooking fuel is used by 20% of the household and only 16.3% of the household treat drinking water. Flush toilet is available in only 1.5% of the houses, in contrast to sand floor which exists in 95% of the houses. There is high dispersion between the percentages of the treatments. Solid martial wall type is the percentage for houses with walls that are built from mud or cement. Other types of walls include those formed from cane, palm, trunks or dirt. Similarly, solid type roofs in this research are the roofs that built from wood, cement, shingles or brick. With solid material walls and roof the house is better protected and provide healthier environment for the children. All treatments are expected to have negative effect on the illnesses prevalence except sand floor. Tests of means difference of the covariates are reported in Table 3, it is clear that the covariates are weakly balanced. Having perfectly balances covariates is extremely difficulty in our data. Limited household characteristics are covered in the survey, which is the reason that we use a full interaction specification in the propensity score estimation. This generates a weak overlap problem at certain ranges of X, accordingly, as has been stated earlier. The rule of thumb trimming method is used by matching observations with propensity score in the range $[0.1, 0.9]$ only.

Table 3. Means difference (standard errors) for the covariates

	Piped water	gas	treat drinking water	treat piped water	sand floor	flush toilet	solid material walls	solid material roof
lexpeduc	0.988 (0.078)	0.932 (0.071)	0.127 (0.078)	0.759 (0.145)	-0.603 (0.127)	0.845 (0.235)	0.458 (0.058)	0.137 (0.106)
lexpfood	0.173 (0.022)	0.242 (0.020)	0.013 (0.022)	0.084 (0.040)	-0.082 (0.035)	0.214 (0.065)	0.043 (0.016)	0.088 (0.029)
oldestM01	5.051 (0.598)	3.689 (0.547)	2.621 (0.591)	3.967 (1.108)	-1.303 (0.972)	2.756 (1.799)	2.195 (0.442)	1.260 (0.807)
nunder5	-0.092 (0.024)	-0.075 (0.022)	0.016 (0.024)	-0.139 (0.045)	0.075 (0.039)	0.001 (0.073)	-0.034 (0.018)	-0.060 (0.033)
nhh	0.505 (0.089)	0.350 (0.082)	0.207 (0.088)	0.257 (0.165)	-0.190 (0.145)	0.124 (0.268)	0.312 (0.066)	-0.038 (0.120)
HH12	0.258 (0.023)	0.217 (0.021)	0.099 (0.023)	0.220 (0.043)	-0.022 (0.038)	-0.113 (0.070)	0.125 (0.017)	0.029 (0.032)
HC02	1.225 (0.057)	1.018 (0.052)	0.430 (0.058)	1.792 (0.107)	-0.811 (0.095)	1.035 (0.177)	0.546 (0.043)	0.420 (0.079)
HC8B	0.097 (0.018)	0.138 (0.016)	0.074 (0.018)	0.112 (0.033)	-0.001 (0.029)	0.177 (0.054)	0.099 (0.013)	0.001 (0.024)
foodaid	-0.015 (0.009)	-0.032 (0.008)	-0.002 (0.009)	-0.032 (0.016)	0.021 (0.014)	0.004 (0.026)	0.007 (0.006)	-0.007 (0.012)
mosqnet	0.129 (0.017)	0.148 (0.016)	0.150 (0.017)	0.023 (0.032)	-0.001 (0.028)	0.063 (0.051)	0.147 (0.013)	0.001 (0.023)
hhmale	0.025 (0.012)	0.013 (0.011)	0.045 (0.012)	0.013 (0.023)	0.006 (0.020)	0.055 (0.037)	0.011 (0.009)	0.046 (0.016)
headec2	0.130 (0.016)	0.130 (0.015)	0.029 (0.016)	0.166 (0.030)	0.014 (0.026)	0.099 (0.048)	0.065 (0.012)	0.002 (0.022)
headec3	0.215 (0.012)	0.216 (0.011)	0.014 (0.012)	0.132 (0.023)	-0.137 (0.020)	0.160 (0.037)	0.079 (0.009)	0.111 (0.017)

1. Sample size is 5850.

2. Simple linear regression is used to compute the mean difference. [3]Numbers in parentheses are slope coefficient standard errors.

5. The Results

Table 4. Coefficients (standard errors) for the propensity score logit with linear specification of the covariates

	Piped water	gas	treat drinking water	treat piped water	sand floor	flush toilet	solid material walls	solid material roof
Intercept	-4.080	-5.412	-2.456	-3.386	3.114	-7.334	-0.339	-3.807
	(0.418)	(0.405)	(0.358)	(0.651)	(0.590)	(1.212)	(0.275)	(0.514)
lexpeduc	0.100	0.094	-0.014	0.072	-0.069	0.102	0.028	-0.009
	(0.022)	(0.020)	(0.020)	(0.037)	(0.032)	(0.060)	(0.016)	(0.027)
lexpfood	0.081	0.438	-0.091	-0.217	-0.008	0.361	-0.087	0.176
	(0.073)	(0.070)	(0.062)	(0.114)	(0.105)	(0.203)	(0.048)	(0.089)
oldestM01	0.026	0.023	0.004	0.015	-0.003	0.005	0.009	-0.002
	(0.003)	(0.003)	(0.003)	(0.006)	(0.005)	(0.010)	(0.003)	(0.005)
nunder5	-0.095	-0.050	0.040	-0.185	0.128	0.074	-0.073	-0.115
	(0.066)	(0.060)	(0.058)	(0.121)	(0.099)	(0.179)	(0.045)	(0.082)
nhh	-0.162	-0.185	-0.044	-0.179	0.041	-0.080	-0.037	-0.049
	(0.026)	(0.024)	(0.023)	(0.044)	(0.037)	(0.069)	(0.018)	(0.032)
HH12	0.377	0.384	0.159	0.189	0.180	-0.663	0.199	0.022
	(0.066)	(0.063)	(0.063)	(0.108)	(0.107)	(0.245)	(0.054)	(0.088)
HC02	0.302	0.229	0.126	0.453	-0.242	0.259	0.147	0.142
	(0.025)	(0.024)	(0.024)	(0.037)	(0.035)	(0.059)	(0.021)	(0.031)
HC8B	-0.112	0.130	0.092	0.014	0.267	0.395	0.130	-0.173
	(0.083)	(0.076)	(0.076)	(0.148)	(0.124)	(0.243)	(0.057)	(0.102)
foodaid	-0.198	-0.613	-0.018	-0.714	0.391	0.152	0.132	-0.098
	(0.171)	(0.174)	(0.148)	(0.379)	(0.282)	(0.431)	(0.112)	(0.208)
mosqnet	0.366	0.463	0.658	-0.185	0.126	-0.002	0.512	-0.076
	(0.090)	(0.083)	(0.084)	(0.150)	(0.126)	(0.242)	(0.058)	(0.104)
hhmale	-0.849	-0.935	0.279	-0.493	0.273	0.159	-0.319	0.457
	(0.162)	(0.149)	(0.155)	(0.271)	(0.235)	(0.493)	(0.113)	(0.217)
headec2	1.305	1.251	0.098	1.186	-0.145	0.721	0.465	0.152
	(0.096)	(0.087)	(0.084)	(0.167)	(0.146)	(0.265)	(0.065)	(0.119)
headec3	1.789	1.796	-0.034	1.078	-0.785	0.920	0.694	0.683
	(0.109)	(0.100)	(0.110)	(0.196)	(0.155)	(0.292)	(0.090)	(0.132)
Log lik	-2105.02	-2409	-2530.55	-844.7312	-1168.0734	-421.02467	-3793.0747	-1591.3442
AIC	4238	4846	5089.1	1717.5	2364.1	870.05	7614.1	3210.7
Pseudo R	0.1741	0.1747	0.0286	0.1512	0.0455	0.0780	0.0510	0.0242

1. Sample size is 5850.

Table 5. Bandwidths of kernel propensity score regression

	Piped water	gas	treat drinking water	treat piped water	sand floor	flush toilet	solid material walls	solid material roof
lexpeduc	3.907	0.769	1.466	3.800	3.106	0.651	2.014	2.085
lexpfood	0.379	0.464	0.603	0.505	0.662	0.322	0.787	0.282
oldestM01	9.135	19.023	11.262	7.596	6.304	5.802	7.597	11.178
nunder5	0.984	0.854	0.746	0.942	0.232	0.613	0.914	0.350
nhh	0.583	0.681	0.884	0.802	0.783	0.810	0.795	0.829
HH12	0.469	0.529	0.648	0.836	0.376	0.197	0.824	0.205
HC02	0.281	0.415	0.388	0.250	0.673	0.544	0.454	0.524
HC8B	0.673	0.366	0.561	0.589	0.695	0.169	0.280	0.268
foodaid	0.390	0.260	0.206	0.152	0.319	0.100	0.674	0.177
mosqnet	0.205	0.152	0.104	0.119	0.441	0.458	0.105	0.316
hhmale	0.695	0.385	0.284	0.358	0.374	0.114	0.828	0.190
headec	0.238	0.117	0.415	0.493	0.190	0.288	0.141	0.296
objective function	681.064	820.044	764.037	205.862	280.159	81.572	1347.255	414.144

1. Sample size is 5850.

2. Bandwidths are estimated using least squares cross validation method in Eq 6.

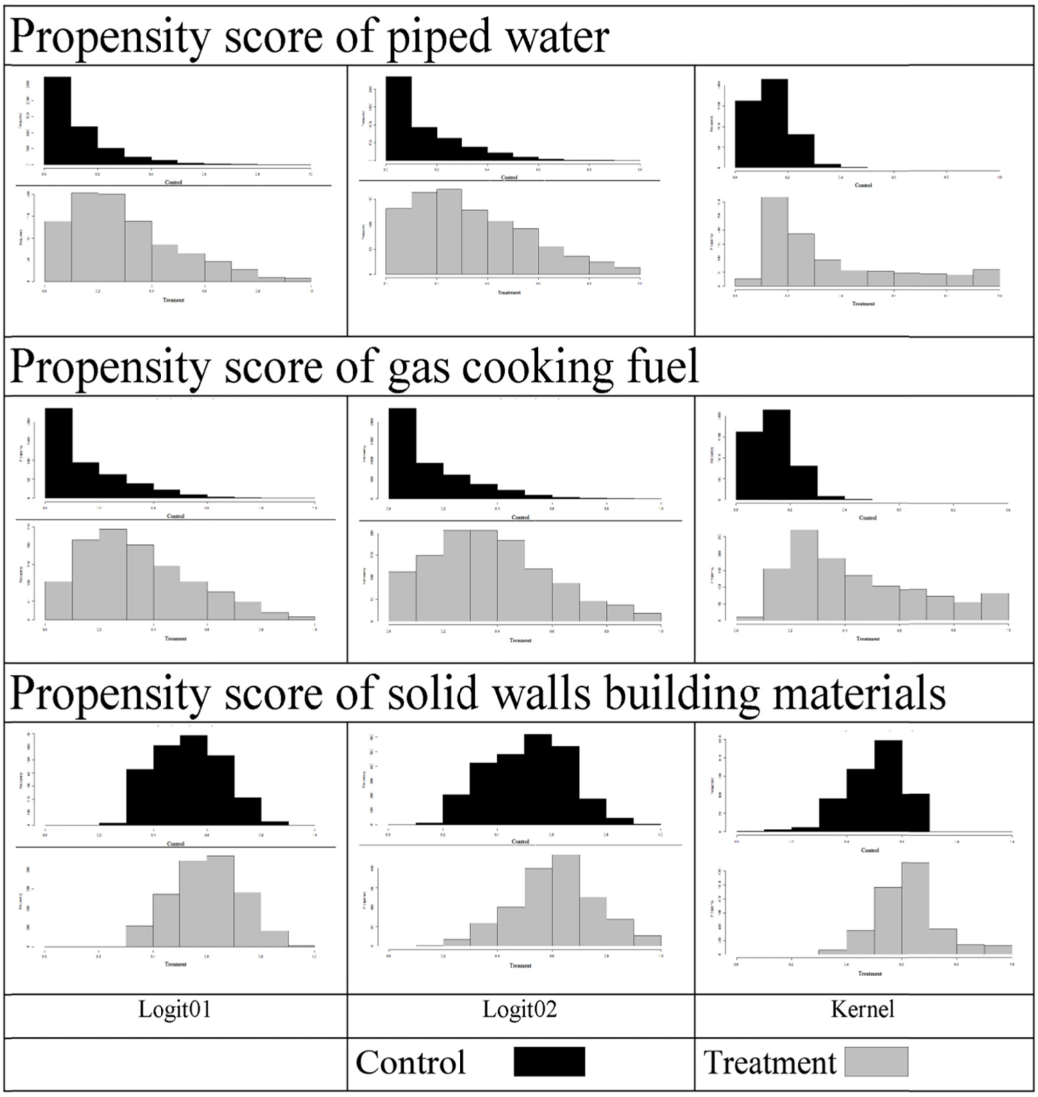

Figure 1. Histogram of the estimated propensity score for chosen treatments

The propensity score is estimated parametrically using two specifications, logit with linear specification of the covariates and logit model with full interaction between the covariates specification. Table 4 reports the coefficients and the standard errors of the first specification, where the coefficients of the latter are suppressed for brevity. The bandwidths of the kernel estimators of the propensity score, that are estimated using Eq 6 are presented in Table 5. The estimated bandwidths indicate nice cross validation objective function convergence, none of the bandwidths is estimated very small, close to zero, or too high. This indicates that all variables are relevant in corresponding regression.

Average treatment effect ATE estimates are reported in Table 6 followed by average treatment effects for the treated in Table 6. The bootstrap p-value of the significance test are shown in the table for each estimated average. The number of bootstrap sample is 1000. Propensity score estimated from logit two specification are very close in ATE, however, for ATT some differences are clearly noticeable. Negative significant effect in both diarrhoea and cough is captured for piped water, gas cooking fuel and wall building material. Treating piped water and roof building material are found effective in reducing prevalence of diarrhoea.

We focus on the results that is reported in Table 6, since they show the estimated average treatment effect on the household that actually treated. Generally, kernel based estimators are smaller in absolute value than logit based estimators and have smaller bootstrap p-value, which indicate that they are stronger. The effect of piped water is estimated negative by all estimators for both diarrhoea and cough. Kernel based PS shows that using gas cooking fuel reduces the prevalence of both diarrhoea and cough together. Piped water is estimated to reduce the prevalence of diarrhoea and cough by 27.8 and 26.5 percentage points, respectively. Gas cooking fuel is an efficient economically and practically which make it easier for the mothers to produce better stewed meals for their children. The gas cooking fuel is estimated to reduce prevalence of diarrhoea by 23.5 percentage points. This effect is explicitly captured by ATT that based on kernel PS only. Walls that are built from solid materials reduce prevalence of diarrhoea and cough slightly compared with piped water and gas cooking fuel, but significantly based on bootstrap p-value.

Table 6. ATE estimates and bootstrap significance test p-value [1]

		Diarrhoea	Cough
		value (boot. p-value)	value (boot. p-value)
Linear form (logit01)			
	Piped water	-0.267 (0.000)	-0.255 (0.000)
	Gas/electricity cooking fuel	-0.206 (0.000)	-0.190 (0.003)
	Treat drinking water	-0.064 (0.972)	-0.065 (0.976)
	Treat piped water	-0.367 (0.016)	-0.382 (0.158)
	Sand floor	0.278 (0.790)	0.279 (0.659)
	Flush toilet	-1.013 (0.957)	-1.014 (0.143)
	Solid wall	-0.032 (0.000)	-0.017 (0.010)
	Solid roof	-0.242 (0.000)	-0.194 (0.231)
Interactions form (logit02)			
	Piped water	-0.279 (0.000)	-0.267 (0.000)
	Gas/electricity cooking fuel	-0.221 (0.000)	-0.209 (0.003)
	Treat drinking water	-0.090 (0.985)	-0.098 (0.975)
	Treat piped water	-0.367 (0.016)	-0.382 (0.158)
	Sand floor	0.278 (0.790)	0.279 (0.659)
	Flush toilet	-1.010 (0.173)	-1.010 (0.290)
	Solid wall	-0.031 (0.000)	-0.009 (0.018)
	Solid roof	-0.261 (0.014)	-0.233 (0.414)
Kernel Estimator			
	Piped water	-0.278 (0.000)	-0.265 (0.000)
	Gas/electricity cooking fuel	-0.235 (0.000)	-0.222 (0.002)
	Treat drinking water	-0.132 (0.973)	-0.141 (0.987)
	Treat piped water	-0.370 (0.053)	-0.394 (0.201)
	Sand floor	0.324 (0.692)	0.351 (0.815)
	Flush toilet	-1.012 (0.663)	-1.012 (0.210)
	Solid wall	-0.030 (0.000)	-0.009 (0.006)
	Solid roof	-0.293 (0.005)	-0.283 (0.230)

1. Model Logit01 uses linear specification of the covariates are presented in Table 4. Model Logit02 uses quadratic for for continuous variables and interaction term between all the variables in the model.

2. ATE estimate exceeded 1.

Table 7. ATT estimates and bootstrap significance test p-value [1]

		Diarrhoea	Cough
		value (boot. p-value)	value (boot. p-value)
Linear form (logit01)			
	Piped water	-0.265 (0.039)	-0.315 (0.002)
	Gas/electricity cooking fuel	-0.322 (0.465)	-0.388 (0.007)
	Treat drinking water	-0.256 (0.842)	-0.285 (0.836)
	Treat piped water	-0.268 (0.994)	-0.323 (0.385)
	Sand floor	0.271 (0.770)	0.270 (0.633)
	Flush toilet	-0.848 (0.957)	-0.899 (0.143)
	Solid wall	-0.325 (0.006)	0.014 (0.664)
	Solid roof	-0.278 (0.150)	-0.295 (0.368)
Interactions form (logit02)			
	Piped water	-0.245 (0.023)	-0.280 (0.009)
	Gas/electricity cooking fuel	-0.307 (0.290)	-0.020 (0.275)
	Treat drinking water	-0.229 (0.996)	-0.274 (0.847)
	Treat piped water	-0.268 (0.994)	-0.323 (0.385)
	Sand floor	0.271 (0.770)	0.270 (0.633)
	Flush toilet	-0.691 (0.173)	-0.696 (0.290)
	Solid wall	-0.312 (0.003)	-0.323 (0.167)
	Solid roof	-0.245 (0.530)	-0.268 (0.610)
Kernel Estimator			
	Piped water	-0.217 (0.000)	-0.236 (0.000)
	Gas/electricity cooking fuel	-0.259 (0.068)	-0.285 (0.038)
	Treat drinking water	-0.224 (0.588)	-0.239 (0.868)
	Treat piped water	-0.237 (0.920)	-0.268 (0.698)
	Sand floor	0.324 (0.646)	0.350 (0.770)
	Flush toilet	-0.792 (0.663)	-0.811 (0.210)
	Solid wall	-0.282 (0.001)	-0.298 (0.117)
	Solid roof	-0.245 (0.257)	-0.245 (0.926)

1. Model Logit01 uses linear specification of the covariates are presented in Table 4. Model Logit02 uses quadratic for for continuous variables and interaction term between all the variables in the model.

2. Numbers in square brackets are the naive bootstrap 95% confidence intervals based on 1000 bootstrap samples.

The available information is not sufficient to capture the effect of flush toilet, very high negative effect is estimated but rejected by bootstrap p-value. We argue that, due to the very small number of houses that are supplied by flush toilet, the effect is likely being over-estimated. The ATE of flush toilet exceeded 100 percentage points is kept unreported. Sand floor has positive effect on both illnesses, but bootstrap p-value is insignificant. We argue that the data does not comprise enough variation to examine the significance of the effect of sand floor on the illnesses.

6. Conclusions

This paper uses the propensity score method to estimate the effect of improving houses construction and condition in rural Sudan on under-five children health. Prevalence of diarrhoea and cough are used as outcome variables. Access to piped water, type of cooking fuel, treatment of drinking water and living house characteristics are used as treatment variable. Kernel nonparametric method is used to estimate the propensity score for the treatments. Treatment effects that are based on nonparametrically estimated propensity score overcome those relay on parametric logit propensity score in terms of the captured effect. Our results show that piped water, using gas cooking fuel and house walls are that built from solid material have the strongest impact on reducing the prevalence of both diarrhoea and cough. This paper recommend that the government should maintain the quality of piped water in good standard to ensure that clean water is supplied to the household sector. The paper also recommences improving the housing sector in rural areas by improving houses building style and materials, and constructing sanitation system. Houses that built from palm, sod and other materials do not provide protection for the inhabitants, not merely under-five children. Our conclusion in this research is that, improving housing industry in rural areas is crucial for improving early childhood health.

References

Acharya, P., Mishra, S. R., & Berg-Beckhoff, G. (2015). Solid fuel in kitchen and acute respiratory tract infection among under five children: evidence from Nepal demographic and health survey 2011. *J Community Health, 40*(3), 515–521. Retrieved from http://link.springer.com/article/10.1007/s10900-014-9965-0

Alemayehu, M. et al. (2014). Household Fuel Use and Acute Respiratory Infections in Children Under Five Years of Age in Gondar city of Ethiopia. *Journal of Environment and Earth Science, 4*(7), 77–86.

Al Mubarak, R. F. (2006). *Pneumonia Case Management in Children Under-Five: A Study in First Referral Hospitals in Khartoum, Sudan*. University of Oslo. Retrieved from https://www.duo.uio.no/handle/10852/30020

Bern, C. et al. (1992). The magnitude of the global problem of diarrhoeal disease: a ten-year update. *Bulletin of the World Health Organization, 70*(6), 705–14. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2393403&tool=pmcentrez&rendertype=abstract

Black, R. E., & Morris, S. S. (2003). Where and Why are 10 million childrendying every year? *Lancet, 361*(9376), 2226–34. Retrieved from http://www.sciencedirect.com/science/article/pii/S0140673603137798

Bryce, J., Shibuya, K., Black, R. E., & the WHO Child Health Epidemiology Reference Group., B.C. (2005). WHO estimates of the causes of death in children. *Lancet, 365*(9465), 1147–1152. Retrieved from http://www.sciencedirect.com/science/article/pii/S0140673605718778

Caliendo, M., Caliendo, M., & Kopeinig, S. (2005). Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of economic surveys, 22*(1588), 31–72. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1467-6419.2007.00527.x/full

Cameron, A. & Trivedi, P., 2005. *Microeconometrics: methods and applications*, Cambridge University Press.

Choi, H. et al. (2010). Common household chemicals and the allergy risks in pre-school age children. *PLoS ONE, 5*(10), e13423. Retrieved from http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0013423

Cotruvo, J. A., & Trevant, C. (2000). Safe drinking water production in rural areas: a comparison between developed and less developed countries. *Schriftenreihe des Vereins fur Wasser-, Boden- und Lufthygiene, 108*, 93–123. Retrieved from http://europepmc.org/abstract/med/11225283

Currie, J., & Rossin-Slater, M. (2013). Weathering the storm: Hurricanes and birth outcomes. *Journal of Health Economics, 32*(3), 487–503. http://dx.doi.org/10.1016/j.jhealeco.2013.01.004

Dehejia, R. H., & Wahba, S. (1999). Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *Journal of the American statistical Association, 94*(448), 1053–1062. Retrieved from http://amstat.tandfonline.com/doi/abs/10.1080/01621459.1999.10473858

Dehejia, R. H., & Wahba, S. (2002). Propensity Score-Matching Methods for Nonexperimental Causal Studies. *Review of Economics and Statistics, 84*(1), 151–161. Retrieved from http://www.mitpressjournals.org/doi/abs/10.1162/003465302317331982

Edejer, T.T.-T. et al. (2005). Cost effectiveness analysis of strategies for child health in developing countries. *Bmj, 331*(7526), 1177. Retrieved from http://bmj.bmjjournals.com/cgi/content/abstract/331/7526/1177

Emmelin, A. (2007). Indoor Air Pollution A Poverty-Related Cause of Mortality Among the Children of the World. *CHEST Journal, 132*(5), 1615–1623. Retrieved from http://journal.publications.chestnet.org/article.aspx?doi=10.1378/chest.07-1398

Esrey, S. et al. (1991). Effects of improved water supply and sanitation on ascariasis, diarrhoea, dracunculiasis, hookworm infection, schistosomiasis, and trachoma. *Bulletin of the World Health Organization, 69*(5), 609–621. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2393264/

Ferng, S.-F., & Lee, L.-W. (2002). Indoor air quality assessment of daycare facilities with carbon dioxide, temperature, and humidity as indicators. *Journal of environmental health, 65*(4), 14–18, 22. Retrieved from http://search.proquest.com/openview/95f91a118b8f4635ff81ec08db411fc9/1?pq-origsite=gscholar

Fewtrell, L. et al. (2005). Water, sanitation, and hygiene interventions to reduce diarrhoea in less developed countries: a systematic review and meta-analysis. *Lancet Infect Dis, 5*(1), 42–52. Retrieved from http://www.sciencedirect.com/science/article/pii/S1473309904012538

Fraker, T., & Maynard, R. (1987). The adequacy of comparison group designs for evaluations of employment-related programs. *Journal of Human Resources, 22*(2), 194–227. Retrieved from http://www.jstor.org/stable/145902

Gross, R. et al. (1989). The impact of improvement of water supply and sanitation facilities on diarrhea and intestinal parasites: a Brazilian experience with children in two low-income urban communities. *Revista de Saúde Pública, 23*(3), 214–220. Retrieved from http://www.scielosp.org/scielo.php?script=sci_arttext&pid=S0034-89101989000300006&lng=en&nrm=iso&tlng=en

Härdle, W., & Linton, O. (1994). Applied Nonparametric Methods. *Handbook of Econometrics, 4*(1069), 2295–2339. Retrieved from http://www.sciencedirect.com/science/article/pii/S1573441205800078

Haroun, H. M. et al. (2010). Assessment of the effect of health education on mothers in Al Maki area, Gezira state, to improve homecare for children under five with diarrhea. *Journal of family & community medicine, 17*(3), 141–146. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3045106&tool=pmcentrez&rendertype=abstract

Heckman, J. (1998). Matching as an econometric evaluation estimator. *The Review of …, 65*(2), 261–294. Retrieved from http://restud.oxfordjournals.org/content/65/2/261.short

Imbens, G. W. (2013). Nonparametric Estimation of Average Treatment Effects Under Exogeneity : a Review. *Review of Economics and statistics, 86*(1), 4–29. Retrieved from http://www.mitpressjournals.org/doi/abs/10.1162/003465304323023651

Imbens, G. W., & Wooldridge, J. M. (2009). *Recent Developments in the Econometrics of Program Evaluation*. Retrieved from http://www.nber.org/papers/w14251

Jalan, J., & Ravallion, M. (2003). Does piped water reduce diarrhea for children in rural India? *Journal of Econometrics, 112*(1), 153–173. Retrieved from http://www.sciencedirect.com/science/article/pii/S0304407602001586

Kumar, S. G., & Subita, L. (2013). Diarrhoeal diseases in developing countries: A situational analysis. *Kathmandu University Medical Journal, 10*(2), 83–88.

LaLonde, R. (1986). Evaluating the econometric evaluations of training programs. *American Economic Review, 76*, 604–620. Retrieved from http://www.jstor.org/stable/1806062

Lechtenfeld, T. (2012). *Why does piped water not reduce diarrhea for children ? Evidence from urban Yemen*, Retrieved from https://www.econstor.eu/handle/10419/90462

Li, Q., Maasoumi, E., & Racine, J. S. (2009). A nonparametric test for equality of distributions with mixed categorical and continuous data. *Journal of Econometrics, 148*(2), 186–200. Retrieved from http://www.sciencedirect.com/science/article/pii/S0304407608002054

Li, Q., & Racine, J. S. (2007). *Nonparametric econometrics: Theory and practice*. Princeton Univ Pr.

Li, Q., Racine, J. S., & Wooldridge, J. M. (2009). Efficient Estimation of Average Treatment Effects with Mixed Categorical and Continuous Data. *Journal of Business and Economic Statistics, 27*(2), 206–223. Retrieved from http://pubs.amstat.org/doi/abs/10.1198/jbes.2009.0015

Lordo, R. (2005). Nonparametric and Semiparametric Models. *Technometrics, 47*(2), 234–234. Retrieved from http://amstat.tandfonline.com/doi/abs/10.1198/tech.2005.s259?journalCode=utch20

Oluwole, O. et al. (2012). Indoor air pollution from biomass fuels: A major health hazard in developing countries. *Journal of Public Health (Germany), 20*(6), 565–575. Retrieved from https://www.scopus.com/inward/record.url?eid=2-s2.0-84871297909&partnerID=40&md5=281cb07c01416cab2152614f652631a0

Qasim, M. et al. (2014). Health Risk Assessment of Indoor Air Quality in Developing Countries. *Asian Journal of Management Sciences & Education, 3*(2). Retrieved from http://www.ajmse.leena-luna.co.jp/AJMSEPDFs/Vol.3(2)/AJMSE2014(3.2-08).pdf

R.M., D., & D'Souza, R. M. (1997). Housing and environmental factors and their effects on the health of children in the slums of Karachi, Pakistan. *Journal of Biosocial Science, 29*(3), 271–281. Retrieved from http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L27271632\nhttp://dx.doi.org/10.1017/S002193209700271X\nhttp://vb3lk7eb4t.search.serialssolutions.com?sid=EMBASE&issn=00219320&id=doi:10.1017/S002193209700271X&atitle=Housing+and+env

Rosenbaum, P. R., & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika, 70*(1), 41–55. Retrieved from http://biomet.oxfordjournals.org/content/70/1/41.short

Rubin, D. B. (1977). Assignment to Treatment Group on the Basis of a Covariate. *Journal of Educational Statistics, 2*(1), 1–26. Retrieved from http://jeb.sagepub.com/content/2/1/1.short

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology, 66*(5), 688–701. Retrieved from http://psycnet.apa.org/journals/edu/66/5/688/

Semenza, J. et al. (1998). Water distribution system and diarrheal disease transmission: a case study in Uzbekistan. *The American Journal of Tropical Medicine and Hygiene, 59*(6), 941–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9886204

Sinha, A. K., & Srivastava, S. P. (1993). Awareness of diarrheal disease control in rural and urban areas of Bihar. *Indian Pediatrics, 30*(12), 1433–1439. Retrieved from http://europepmc.org/abstract/med/8077033

Siziya, S., Muula, A. S., & Rudatsikira, E. (2013). Correlates of diarrhoea among children below the age of 5 years in Sudan. *African health sciences, 13*(2), 376–83. Retrieved from http://www.mendeley.com/catalog/correlates-diarrhoea-among-children-below-age-5-years-sudan/

ULLAH, A. (2002). Improved Combined Parametric and Nonparametric Regressions: Estimation and Hypothesis Testing. *Handbook of applied econometrics and statistical inference, 165*, 159. Retrieved from Improved Combined Parametric and Nonparametric Regressions: Estimation and Hypothesis Testing.

Ullah, A. (1988). Non-Parametric Estimation of Econometric Functionals. *The Canadian Journal of Economics, 21*(3), 625. Retrieved from http://www.jstor.org/stable/135443?origin=crossref

Upadhyay, A. K. et al. (2015). Impact of indoor air pollution from the use of solid fuels on the incidence of life threatening respiratory illnesses in children in India. *BMC public health, 15*(1), 300. Retrieved from http://apps.webofknowledge.com.proxy-um.researchport.umd.edu/full_record.do?product=WOS&search_mode=GeneralSearch&qid=1&SID=1FRUQje5h6Eqr1Rvd6B&page=1&doc=3

Waddington, H. et al. (2009). *Water, Sanitation and Hygiene Interbentions to Combat Childhood Diarrhoea in Developing Countries*, International Initiative for Impact Evaluation New Delhi. Retrieved from http://www.popline.org/node/210650

Waddington, H., & Snilstveit, B. (2009). Effectiveness and sustainability of water, sanitation, and hygiene interventions in combating diarrhoea. *Journal of Development Effectiveness, 1*(3), 295–335. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/19439340903141175

Wardlaw, T., Salama, P., Brocklehurst, C., Chopra, M., & Mason, E. (2010). Diarrhoea: Why Children Are Still Dying and What Can Be Done. *The Lancet, 375*(9718). Retrieved from http://www.citeulike.org/group/13490/article/7498768

Wooldridge, J. M. (1999). Asymptotic Properties of Weighted M-Estimators for Variable Probability Samples. *Econometrica, 67*(6), 1385–1406. Retrieved from http://www.jstor.org/stable/2999565

Wooldridge, J. M. (2001). *Econometric Analysis of Cross Section and Panel Data*. MIT press.

Yu, F. (2011). Indoor Air Pollution and Children ' s Health : Net Benefits from Stove and Behavioral Interventions in Rural China. *Environmental & Resource Economics, 50*(4), 495–514. Retrieved from http://link.springer.com/article/10.1007/s10640-011-9479-5

Zhang, J. (2012). The impact of water quality on health: Evidence from the drinking water infrastructure program in rural China. *Journal of Health Economics, 31*(1), 122–134. http://dx.doi.org/10.1016/j.jhealeco.2011.08.008

Zhang, J., & Smith, K. R. (2003). Indoor air pollution: a global health concern. *British Medical Bulletin, 68*(1), 209–225. Retrieved from http://bmb.oupjournals.org/cgi/doi/10.1093/bmb/ldg029