Evaluating Endogenous Program Interventions with Heterogeneous Treatment Intensity using Bayesian Potential Outcomes Approach

Andinet Woldemichael[*], Degnet Abebaw[†], and Assefa Admassie[‡]

May, 2014

Abstract

In this paper, we implement Bayesian potential outcomes model to evaluate the impact of program interventions using non-randomized data. The approach jointly addresses selection bias in program placement, heterogeneous treatment intensity among the treated, and heterogeneity in treatment effects. Using data from a non-randomized household survey, we evaluate the impact of Ethiopia's Health Extension Program on fertility and child mortality outcomes. We find that there is significant selection bias in both program placement and intensity of exposure to the program among the treated. On average, the program has significant impact in reducing fertility and child mortality. However, there is notable heterogeneity in the treatment effects ranging from negative impacts for some individuals to positive impacts for the majority in the sample. We recover individual-level treatment effects and present the distributions graphically.

Keyword: Endogenous Program Placement, Heterogeneous Treatment Intensity, Bayesian Potential Outcomes, Heterogeneous Treatment Effects, Health Extension Program, Fertility, Child mortality.

JEL Classification: C11, C31, D04, I15, I18, J13.

[*]Correspondence to: Pharmaceutical Health Services Research Department, School of Pharmacy, University of Maryland, Baltimore, 220 Arch Street, 12th Floor, Baltimore, MD 21201, U.S.A. Correspondence e-mail: awoldemichael1[at]gsu.edu.

[†]Ethiopian Economic Association/Ethiopian Economic Policy Research Institute, Addis Ababa, Ethiopia

[‡]Ethiopian Economic Association/Ethiopian Economic Policy Research Institute, Addis Ababa, Ethiopia

1 Introduction

Advances in causal inference and program evaluation methods in the past few decades have made significant contribution in quantifying the impact of social policy interventions in various frontiers. However, several practical issues undermine the soundness of inference drawn from non-randomized program evaluations. Treatments could be assigned on the basis of factors related to outcome/s of interest resulting in selection bias. Deviations from treatment protocols are common in terms of compliance and adherence which gives rise to heterogeneity in treatment intensity (HTI) as individuals are autonomous and they pick and choose parts of an intervention. Programs have heterogeneous treatment effects (HTEs); some individuals gain, some individuals lose, and others do not gain or lose from exposure to the program. This results in difficulty to have a good summary measure of treatment effects which accounts for heterogeneity. These are opportune problems which could arise together and make sound program evaluation difficult.

There are several approaches in the causal inference literature, mainly in the classical framework, which are vastly discussed elsewhere in the econometrics and statistics literature. See for instance studies on matching methods in Rosenbaum and Rubin (1983a,b); Rubin (1986); Dehejia and Wahba (1999), instrumental variable methods in Heckman and Robb Jr (1985); Heckman et al. (1997); Angrist et al. (1996), and recent papers such as Heckman and Navarro-Lozano (2004); Wooldridge (2007); Imbens and Wooldridge (2008); Heckman et al. (2014) for a comprehensive treatment of several methods. Rubin stated that "*all problems of causal inference should be viewed as problems of missing data: the potential outcomes under the not-received treatment are the missing data*"(Rubin, 2005b). In this spirit, he proposed Bayesian approach of potential outcomes in causal inference in his 1978 paper (Rubin et al., 1978). For decades, Rubin's Bayesian potential outcomes (BPOs) approach received little attention mainly due to the computational difficulties involved in estimating such models. Since early 1990s, however, advances in Bayesian estimation methods such as the use of Markov Chain Monte Carlo (MCMC) simulation methods, *data augmentation* techniques, and availability of cheap computing power have made estimation of BPOs models feasible and efficient (Tanner and Wong, 1987; Chib and Hamilton, 2000; Rubin, 2005a; Li and Tobias, 2005; Heckman et al., 2014). The framework has a nice property which allows us to explicitly model the problem of missing data in the Roy-Neyman-Rubin potential outcomes framework.

In this paper, we apply BPOs approach to evaluate the interim impact of the Ethiopian Rural Health Extension Program (rHEP) on fertility and child mortality outcomes. In particular, we use ordered probit Potential Outcomes (POs) approach proposed in Li and Tobias

(2008) to model Potential Fertility Outcomes (PFOs) and Potential Child Mortality Outcomes (PCMOs), both of which are count variables (see Albert and Chib (2001); Munkin (2011) for similar approach), and jointly address issues of selection bias, HTI, and HTEs. We use quasi-randomized household survey data collected by the Ethiopian Economic Association/Ethiopian Economic Policy Research Institute (EEA/EEPRI) for the purpose of evaluating the interim impact of the program on selected health outcomes.

Reports show that Ethiopia is one of the few developing countries who have strongly gained in terms of fertility and child survival outcomes. According to the Ethiopian Demographic and Health Survey (DHS) 2011, fertility has declined from 5.5 to 4.8 children per woman between 2000 and 2011. Similarly, under-five mortality rate has declined from 166 children per 1,000 live births in 2000 to 88 children in 2011(Central Statistical Agency [Ethiopia] and ICF International, 2012). These gains could be attributed to the country's rHEP through better access to family planning and basic health services (UNICEF, 2013; Workie and Ramana, 2013), which has been a flagship program of the overall health sector development plan and strategy of the government reflecting a marked shift in paradigm from a traditional approach of providing access to basic health care services to an innovative community-based approach. However, there is no clear scientific evidence in the literature which quantifies the causal impact of the rHEP on fertility and child mortality outcomes.

The first challenge in quantifying the impact of the rHEP using the EEA/EEPRI survey data is selection bias in the timing of implementation of the program and deployment of HEWs in villages. Villages which implemented the program earlier could have been selected on the basis of factors related to their prior or current fertility and child mortality outcomes. The second challenge is HTI among the treated in that some treated villages received only one HEW and others received two or more HEWs. As discussed in Ravallion and Wodon (1998), serious bias arises in evaluating targeted programs in developing countries which arises from purposive targeting of geographic areas and targeting of individual recipients. HTE and HTI are more likely to attenuate the estimated ATEs towards zeros. Endogeneity in program placement overstates or underestimates treatment effects, depending on the direction of the correlation.

In BPOs framework, we address selection bias on observable and unobservable dimensions through joint estimation of the POs and program assignment equations. We address HTI among the treated by directly modeling the ordered program assignment mechanism; ordered as non-exposure, exposure to one HEW, and exposure to two or more HEWs. We deal with HTEs by exploiting the advantage of *data augmentation* and MCMC simulations to recover individual-specific treatment effects (iTEs) and present the whole distribution of TEs (Chib and Hamilton, 2000). Our approach is Intent-To-Treat (ITT) in that all indi-

viduals living in the treated villages were subject to treatment regardless of compliance or adherence to treatment protocols. This is particularly intuitive as the rHEP is a community-based intervention primarily aimed at prevention and promotive services and transfer and dissemination of health knowledge and information.

We control for selection bias on observables by including village level characteristics in the program assignment equation and individual-level and village-level characteristics in the potential outcomes equations.The estimated correlations between the error terms in program placement and potential outcomes equations are up to 57% implying that there is significant selection bias in the assignment of HEWs on the unobservable dimension as well and ignoring it would bias the estimated ATEs. After controlling for selection bias on observable and unobservable dimensions, the results show that the rHEP has significant impact in reducing fertility and child mortality among mothers and reproductive-age women in rural villages in Ethiopia. Had a typical mother been exposed to only one HEW as opposed to non-exposure, she would have given birth to 0.25 less number of children in the past five years. The reduction in fertility would have been 0.27 had she been exposed to two or more HEWs as opposed to non-exposure. In terms of child mortality, had a typical mother been exposed to only one HEW as compared to non-exposure she would have lost 0.17 less number of children due to death. Had she been exposed to two or more HEWs instead, she would have lose 0.18 less number of children to death. However, there is significant HTEs ranging from negative impacts for few individuals to positive impacts for the majority in the sample. We recover iTEs from the MCMC simulations and present the distributions graphically.

The rest of the paper is organized as follows. Section 2 describes the program, Section 3 discusses the data and Section 4 presents the Bayesian potential outcomes model. While Section 5 discusses the results Section 6 concludes the paper.

2 Program Description

The rHEP is part of the overall health sector development strategy of the Ethiopian government launched in 2003. The program focuses on preventive and promotive health services including basic health care services, immunization, maternal health, attended delivery, family planning (FP), and non-referral services such as transfer of health information and knowledge. The program aims at reducing travel and time costs that rural residents would otherwise have to spend to get access to basic health care services in urban centers. It is designed to cover all 15,000 villages throughout the country by deploying a team of two trained female HEWs in each village. They are expected to spend about 75% of their time conducting door-to-door and community-level outreach services (Admassie et al., 2009a).

The HEWs have high school diplomas (10th grade and above) and one year of training in basic health services from Vocational and Technical Colleges (VTCs). In addition to deploying a pair of HEWs in each village (Kebele), a Health Post (HP) is constructed, which is the smallest unit of basic health care facility, serving as a docking station for HEWs. At an estimated HP construction cost of $80,000, HEW training cost of $5,000, and total monthly salary of $167 per village, the rHEP is one of the least-cost programs which provides access to basic health care services to about 84% of the country's population who resides in 15,000 rural villages (Admassie et al., 2009b).

The federal government mandates regional states to recruit, train, and assign HEWs, as well as construct and equip a HP in each village. Big and populous regions implement the program in a stepwise manner due to the large number of villages they have to cover, shortage of qualified candidates in target villages, and limited capacity of regional VTCs dedicated to train HEWs. At the district level, health officials consider various factors such as availability of health centers in the nearby villages, history of child and maternal health outcomes, and other health issues in the village before recruiting, training, and deploying HEWs (Admassie et al., 2009b). As a result, there is variation in the timing and intensity of implementing the rHEP. Out of the 30,000 HEWs needed for complete coverage, as of June 2007, only 17,653 HEWs were trained and deployed to their respective villages. Some villages received HEWs as early as in 2003 and others were yet to receive HEWs during the time of field survey in 2007. In addition, among villages who received HEWs some received two or more HEWs in compliance with the program design and others received only one HEW giving rise to HTI.

3 The Data

3.1 Sampling strategy

We use quasi-randomized household survey data collected by the Ethiopian Economic Association/Ethiopian Economic Research Institute (EEA/EERI) from 130 control and treatment villages. The survey was conducted with the primary objective of conducting interim impact evaluation of the rHEP on selected health outcomes. The data was collected using a combination of purposive and multi-stage random sampling exploiting the spatial and temporal variations in the deployment of the program to identify villages which have received and not yet received HEWs. The survey identifies districts in each region which started implementing the program in some but not all villages covering three regions, Amhara, Oromia and SNNP, 10 districts, and 130 control and treatment villages. From each district, the study

randomly selected control group from a pool of villages which did not receive HEWs and treatment group from a pool of villages which received HEWs. It excludes treatment villages which started implementing the rHEP after 2005 focusing on villages which were exposed to the rHEP for a reasonable amount of time. In the third stage, the study uses village level population census records to randomly select households which have at least one under-five child. Sample sizes of villages and households were determined using power calculation on the basis of selected child and maternal health outcomes. Sampling and data collection was conducted between May and September 2007 (see Admassie et al. (2009b,a)).

The dataset has information on individual-level demographic characteristics, child immunization and health, maternal health, contraceptive use, fertility outcomes, and village-level characteristics. After excluding observations with missing values the outcome variables, we have 2,490 mothers and 3,739 reproductive-age women who reside in 66 treatment villages and 60 control villages.

3.2 Outcome Variables

The outcome variables of interest are the number of under-five children born to a mother in the past five years and the number of children who has died. We have very few mothers and women who gave birth to two or more children in the past five year and mothers who lost two or more children to death. For ease of modeling and estimation, we top-code the outcome variables to two if the respondent reported to have two or more (under-five children or dead) children. Table (1) presents descriptive statistics for the number of under-five children, the number of children ever borne, the number of live children, and the number of dead children by treatment status for a subsample consisting of only recent mothers in panel (a) and a subsample consisting of all reproductive-age women in panel (b). In panel (c) we present summary statistics of the same outcome variables but for women older than 50.

The total fertility rates (TFR) are 3.4 and 5.1 children for the whole sample of reproductive-age women and the subsample of mothers only, respectively. The TFR for women who have passed childbearing ages is 5.6. This figure is slightly lower than the Demographic and Health Survey (DHS) of Ethiopia report of 6 TFR in rural areas in 2005. Not only were fewer under-five children were born to mothers and reproductive-age women in treatment villages but also fewer number of children were dead compared to control villages. This underscores the positive impact that the rHEP has in helping women and mothers achieve their fertility goals as well as reduce child mortality.

3.3 Treatment Variable

We categorize villages into control and treatment groups. Further, we categorize treated villages into two; exposed to only one HEW and exposed to two or more HEWs. The treatment assignment variable is, therefore, count variable taking values 0, 1, and 2 if no HEWs (control), only one HEW or two HEWs, respectively, were deployed.

HTI could arise from the extent to which HPs are equipped and supplied with the necessary drugs and medical supplies, topology and access to modern transportation means, sparsity and density of the settlement, and motivation and dedication of HEWs. We focus on HTI arising from differences in the number of HEWs assigned to villages. Everything else remains the same, a team of two HEWs in a village better transfer health information and knowledge and provide basic health care services than single HEW.

3.4 Control Variables

In Table (2), we summarize control variables included in the potential fertility and child mortality outcomes equations which are demographic characteristics of mothers such as age, marital status, relation to the head, occupation, level of education, religion, and the number of female household members between the age of 14 and 45 in the household. We control for wealth effect by constructing a single wealth index using principal component analysis on household durables, furnitures and tools, farming tools and type of housing[1]. Then, we create dummy variables indicating the wealth quintile that a household belongs to. In addition, we control for village-level characteristics such as literacy rate, distance to tarmac road as well as district and regional dummies to control for spatial variations.

In the HEWs assignment equation, we include village-level variables to account for factors that could potentially influence the decision of a typical district administrator to assign certain number of HEWs to each village. Table (3) presents summary statistics of control variables by treatment status. These include dummy variable indicating whether there is basic adult education centers in the village, village-level literacy rate, and dummy variable indicating whether there is primary school in the village to control for educational characteristics and pre-existing health knowledge. We include dummy variable indicating whether the village is located in lowland area and hence susceptible to vector-borne diseases.

[1] Access to electricity, ownership of watch, radio, Television set, tape recorder, gold jewelry, mobile phone, fixed telephone, refrigerator. Ownership of furniture and household tools such as table, chair, bed, electric mitad (griddle), kerosene lump, gas medeja (stove), bicycle, cart, sewing machine. Ownership of farm tools such as plough, ax, sickle, machete, spade and hoe. Housing construction materials and condition such as corrugated iron roof, grass (sar) or bamboo roof, mud wall, wood wall, bamboo wall, mud floor and size of cultivated land.

To control for variation in access to health services, we include distance from the center of the village to the nearest tarmac road and dummy variable indicating availability of health center in the nearby villages. We also include dummy variables indicating access to clean water and whether or not waterborne disease is prevalent in the village. Furthermore, we include district and regional dummies to account for spatial variation in health practice, geography, budgetary, and cultural norms.

4 Bayesian Potential Outcomes Framework

4.1 The Model

Our empirical model is in the realm of the Neyman-Roy-Rubin potential outcomes framework. We follow Li and Tobias (2008) approach of modeling the *counterfactuals* in Bayesian estimation framework with *data augmentation* technique (Tanner and Wong, 1987; Albert and Chib, 1993) and reparametrization. In this framework, the count potential outcomes and treatment levels are modeled using ordered probit model (see similar studies such as Chib and Hamilton (2000)).

To fix ideas, let $y_{iv} \in \{0, 1, 2\}$ be the observed number of outcomes for the i^{th} individual in the v^{th} village, where $i = 1, \cdots, N$, $v = 1, \cdots, V$, N and V are the number of observations and villages, respectively. The observed treatment variable, denoted by T_v, is the number of HEWs taking values 0, 1, and 2 for no HEW (control), only one HEW, and two or more HEWs, respectively, are deployed. This can be represented by three dummy variables $\{T_{0v}, T_{1v}, T_{2v}\}$ indicating assignment of $h = \{0, 1, 2\}$ HEWs to village v. Then, the single observed outcome, y_{iv}, for any given individual i in village v can be written as

$$y_{iv} = T_{2v}y_{2,iv} + T_{1v}y_{1,iv} + (1 - T_{2v} - T_{1v})y_{0,iv}. \tag{1}$$

The observed treatment variable, T_v, can be related to the latent index as

$$T_v = h \quad \text{iff} \quad (\gamma_h \leq z_{T_v} \leq \gamma_{h+1}), \quad h = 0, 1, 2, \tag{2}$$

where z_{T_v} is the latent utility of typical district administrator from assigning h number of HEW/s in village v, $\{\gamma_h\}$ are the cut-points that map the continuous latent utility values into the observed discrete treatment values (i.e, number of HEWs). For instance, according to assignment rule (2), the district administrator assigns two HEWs in village v if and only if the latent utility from assigning two HEWs is greater than or equal to the latent utility

of assigning only one HEW. For assignment T_v the potential outcomes, $y_{h,iv}$, takes discrete values according to the rule

$$y_{h,iv} = k \quad \text{iff} \quad (\widetilde{\gamma}_{h,k} \leq z_{h,iv} \leq \widetilde{\gamma}_{h,k+1}), \quad k = 0, 1, 2, \tag{3}$$

where $z_{h,iv}$ is the continuous latent potential utility of individual i in village v with h number of HEWs, $\widetilde{\gamma}_{h,k}$ are the cut-points for each potential outcome equation mapping the latent utility defined in (3) into the observed level of outcomes.

Suppose a typical district administrator's latent utility, z_{T_v}, can be written as a function of village level vector of observables, w_v, and unobservable factors, ϵ_{T_v}. Also assume that individual i's latent potential utility, $z_{h,iv}$, in treated state h can be expressed as a function of a vector of observable factors $\{x_{0,iv}, x_{1,iv}, x_{2,iv}\}$ and unobservable factors $\{\epsilon_{0,iv}, \epsilon_{1,iv}, \epsilon_{2,iv}\}$. Then, the system of rHEP assignment and POs equations can be written as

$$\begin{aligned} z_{T_v} &= w_v\beta_T + \epsilon_{T_v}, \\ z_{2,iv} &= x_{iv}\beta_2 + \epsilon_{2,iv}, \\ z_{1,iv} &= x_{iv}\beta_1 + \epsilon_{1,iv}, \\ z_{0,iv} &= x_{iv}\beta_0 + \epsilon_{0,iv}, \end{aligned} \tag{4}$$

where $x_{iv(1\times k)}$ is a vector of covariates, which we let to be the same under the three treatment status, i.e. $x_{0,iv} = x_{1,iv} = x_{2,iv} = x_{iv(1\times k)}$, $w_{v(1\times m)}$ is a vector of village level covariates influencing district administrator's utility, $(\beta_T, \beta_2, \beta_1, \beta_0)$ are vector of "structural parameters" to be estimated, and ϵ_{T_v}, $\epsilon_{2,iv}$, $\epsilon_{1,iv}$, and $\epsilon_{0,iv}$ are unobserved factors influencing the decision to assign h number of HEWs and the POs, respectively. For the sake of identification, we restrict the cut-points as follows: $\widetilde{\gamma}_{0,k} = \gamma_0 = -\infty$, $\widetilde{\gamma}_{1,k} = \gamma_1 = 0$, and $\widetilde{\gamma}_{3,k} = \gamma_3 = \infty$. Such restriction reduces the number of unknown cut-points in each equation from that of four to only one. The remaining unknown cut-points are $\widetilde{\gamma}_{2,2}, \widetilde{\gamma}_{2,1}, \widetilde{\gamma}_{2,0}$, and γ_2 for the POs and HEWs assignment equations, respectively.

Assuming joint normality, endogeneity of program placement is handled by explicitly allowing the program assignment and the POs equations to be correlated. The covariance matrix can be written as

$$\begin{pmatrix} \epsilon_{T_v} \\ \epsilon_{2,iv} \\ \epsilon_{1,iv} \\ \epsilon_{0,iv} \end{pmatrix} | x_{iv}, w_v \overset{iid}{\sim} N_4 \left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \sigma_{2T} & \sigma_{1T} & \sigma_{0T} \\ \sigma_{2T} & 1 & \sigma_{21} & \sigma_{20} \\ \sigma_{1T} & \sigma_{21} & 1 & \sigma_{10} \\ \sigma_{0T} & \sigma_{20} & \sigma_{10} & 1 \end{pmatrix} \right) \equiv N_4(0, \Sigma), \tag{5}$$

where the diagonal elements of the covariance matrix are normalized for identification purpose and the off-diagonal elements are covariances among the system of equations representing HEWs assignment decision and the associated POs. The last three elements σ_{21}, σ_{20} and σ_{10} are not directly identified by the data since they do not enter the likelihood function, but can be updated through learning owing to the positive definiteness condition imposed on Σ (see Poirier (1998); Li and Tobias (2005); Poirier and Tobias (2003)).

The likelihood contribution of individual i in village v is given by

$$p(y,T|\Gamma) \equiv \prod_{i:T_v=2} Pr(y_{2,iv}=y_{iv},T_v=2|\Gamma) \times \prod_{i:T_v=1} Pr(y_{1,iv}=y_{iv},T_v=1|\Gamma) \times \prod_{i:T_v=0} Pr(y_{0,iv}=y_{iv},T_v=0|\Gamma), \tag{6}$$

where β=(β_T, β_2, β_1, β_0), $\gamma = (\gamma_2, \widetilde{\gamma}_{22}, \widetilde{\gamma}_{12}, \widetilde{\gamma}_{02})$, σ=(σ_{2T}, σ_{1T}, σ_{0T}, σ_{21}, σ_{20}, σ_{10}), $\Gamma = [\beta \quad \gamma \quad \sigma]$, and the corresponding joint probabilities are estimated from a bivariate normal cumulative density function (cdf). Although estimating the posterior distribution of the regression parameters from the augmented joint posterior distribution is fairly straightforward, the difficulty lies in drawing a non-standard covariance matrix Σ due to the restriction we impose on the diagonal elements. One can still draw the restricted Σ using Metropolis-within-Gibbs algorithm by restricting the variances and imposing positive definiteness. However, this method suffers from slow mixing due to high correlation between simulated cut-points and the latent data (Cowles and Carlin (1996); Nandram and Chen (1996); Li and Tobias (2005)). Instead, we use a reparameterization approach proposed by (Li and Tobias, 2005) which makes drawing from a standard Wishart distribution feasible. Their approach avoids the difficulty of directly drawing posterior distribution of the cut-points, and performs efficiently[2].

We put the following non-informative conjugate priors[3][4]: $\beta_0^* \sim N(b_0, V_{\beta^*})$, and $\Sigma_0^* \sim IW(v,R)$. After reparametrization, all the cut-points are known after and hence we do not

[2]We perform the following reparameterization using the largest unknown cutpoint: $\sigma_T = \frac{1}{\widetilde{\gamma}_2^2}$, and $\sigma_h = \frac{1}{\widetilde{\gamma}_{h(2)}^2}$. Then multiplying the latent variable in equations (4) above and their parameters including the cut-points by $\sqrt{\sigma_T}$, $\sqrt{\sigma_2}$, $\sqrt{\sigma_1}$, and $\sqrt{\sigma_0}$, respectively, we obtain $z_{T_v}^* = \sqrt{\sigma_T} z_{T_v}$, $\beta_T^* = \sqrt{\sigma_T}\beta_r$, $\epsilon_{T_v}^* = \sqrt{\sigma_T}\epsilon_{T_v}$, and $z_{h,iv}^* = \sqrt{\sigma_h} z_{h,iv}$, $\beta_h^* = \sqrt{\sigma_h}\beta_h$, $\epsilon_{h,iv}^* = \sqrt{\sigma_h}\epsilon_{h,iv}$

[3]The following values are chosen for the priors in all specifications: $\beta_0 = 0_{(k\times 1)}$, $V_{\beta^*} = 100I$, $v = 6$, $R = I_{4\times 4}$ with no prior values on the cutpoints. We draw 10,000 MCMC iterations discarding the first 5,000 as burn-in periods. We assess convergence through examining trace plots of posterior model parameters as well as formal Convergence Diagnostic (CD) tests proposed by Geweke (1992). We tested the algorithm on simulated data before implementing it to real data. The algorithm recovers all the true parameters in an efficient manner.

[4]The prior for the cut points is proportional to a constant, $\pi(\gamma_2^*, \widetilde{\gamma}_{2(2)}^*, \gamma_{1(2)}^*, \widetilde{\gamma}_{0(2)}^*) \propto c$. However, in our case they are all known.

place a prior on them. The posterior simulator which fits the reparameterized system of ordered treatment assignment and POs equations draws the latent data, the regression parameters, and the covariance matrix from their complete posterior conditional distributions. Details of the algorithm is presented in Appendix A.

In terms of model parameter identification, we use exclusion restriction approach. The strategy is to exclude variables that significantly influence district administrator's decision to assign HEWs but do not affect the POs associated with a particular level of treatment. Similarly, we exclude variables which influence individuals' PFOs and PCMOs but do not affect district administrator's decision to assign HEWs from the assignment equation.

Based on a battery of Sargan-Hansen overidentification tests[5], which under the joint null hypothesis are uncorrelated with the error terms, we jointly exclude variables indicating whether the village has access to pipe water and prevalence of malaria. Similarly, we exclude individual-level demographic and wealth variables from the HEWs assignment equation.

4.2 Treatment Effects

We estimate the treatment effects for each individual (iTE) as follows

$$\begin{aligned} iTE(x;\Gamma) &\equiv E(y_h - y_l|x,\Gamma) \\ &= \sum_{j=1}^{J} j\{P(y_h = j|x,\Gamma) - P(y_l = j|x,\Gamma)\}, \quad h > l. \end{aligned} \tag{7}$$

The iTE can be estimated using posterior parameter values obtained from the MCMC simulation as

$$\widehat{iTE}(x) = E_{\Gamma|y,T}(iTE(x;\Gamma)) \approx \frac{1}{R}\sum_{r=1}^{R} iTE(x;\Gamma_R), \tag{8}$$

where $\Gamma_R \sim p(\Gamma|y,T)$ is a vector of post-convergence parameters obtained using the algorithm described above. Then, the ATE can be obtained by averaging $\widehat{iTE}$ over sample observations, i.e. $\widehat{ATE} = \frac{1}{N}\sum_{i=1}^{N}\widehat{iTE}$ for the population and $\widehat{ATE}_0 = \frac{1}{N_0}\sum_{i=1}^{N_0}\widehat{iTE}_0$, $\widehat{ATE}_1 = \frac{1}{N_1}\sum_{i=1}^{N_1}\widehat{iTE}_1$, and $\widehat{ATE}_2 = \frac{1}{N_2}\sum_{i=1}^{N_2}\widehat{iTE}_2$ for the subpopulations which are not exposed to the rHEP, exposed to one HEW and exposed to two HEWs, respectively.

[5]The variables that we run the Sargan-Hansen Overidentification test on include indicators for access to pipe water, prevalence of malaria, waterborne diseases, lowland area, drought prevalent, and prior NGOs services.

5 Results and Discussions

5.1 The rHEP Placement and Endogeneity

Table (4) presents the posterior means of coefficients of the ordered probit model of assigning h number of HEWs in village v. We let PFO0, PFO1, and PFO2 to denote potential fertility outcomes had a typical women been not exposed, exposed to one HEW, and exposed to two or more HEWs, respectively. Similarly, PCMO0, PCMO1, and PCMO2 to denote the potential child mortality outcomes had a typical mother been not exposed, exposed to one HEW, and exposed to two or more HEWs, respectively. All coefficients in the fertility model are obtained from the joint estimation of the system of HEWs assignment and PFOs equations. Similarly, the coefficients in child mortality model are obtained from the joint estimation of the system of equations of HEWs assignment and the associated PCMOs.

In terms of direction, the coefficients are generally consistent across the two models and the subsample consisting of mothers only and the subsample consisting of all reproductive-age women. This implies that HEWs assignment decisions, i.e., selection bias observables, are mainly derived by a common vector of variables regardless of the outcome of interest. Villages which have basic adult education centers, primary schools, higher literacy rates, access to pipe water, and that waterborne disease is their primary health problem have better chance of receiving higher number of HEWs. These coefficients are statistically significant, which are within the bounds of two standard deviations. On the other hand, villages in remote places which are far from tarmac roads and villages located in lowland (lower altitude) areas are less likely to receive HEWs. Factors such as the presence of health centers in neighboring villages and regions do not influence the decision to assign HEWs.

We capture the level of endogeneity in the assignment of different number of HEWs directly through the joint estimation of correlation parameters. Tables (6) and (7) present the posterior means and standard deviations of the correlations. Panel (a) shows the correlation between the unobserved factors which influence the decision to assign HEWs and the unobserved factors which influence the PFOs. Panel (b) shows the correlations amongst the PFOs had individuals been exposed to different number of HEWs. PFO2 and the decision to assign HEWs have a correlation of 52% for the subsample consisting of mothers only. Similarly, PFO1 and the decision to assign HEWs have a positive correlation of 17%. However, the PFO0 and the decision to assign HEWs have a negative correlation of 57%. The same pattern is observed when we expand our sample to include reproductive-age women but with lower degree of correlation. In the child mortality model, the level of endogeneity is still significant. While the correlation between HEWs assignment equation and the PCMO2 is 39%, the correlation with PCMO0 is negative 46%, which are statistically significant.

5.2 The POs and Causal Effect of rHEP

We present the estimated treatment effects averaged over the population under consideration and sub-populations (i.e., ATEs) in tables and the densities and cumulative densities of iTEs are plotted in figures. In Table (5) panel (a), we present the posterior means and standard deviations of the PFOs associated with non-exposure and exposure to a different number of HEWs. In panel (b) we present the ATEs of exposure to the rHEP on fertility outcomes had a typical mother and a typical reproductive-age woman been exposed to different number of HEWs.

The PFOs due to exposure to higher number of HEWs are higher than the PFOs due to non-exposure. For instance, had a typical mother been exposed to the rHEP with two or more HEWs (PFO2) she would have given birth to 0.26 (std = 0.06) number of children in the past five years. Similarly, the potential number of under-five children that a typical mother would have given birth to had she been exposed to one HEW (PFO1) is 0.28 (std = 0.06). Whereas the number of under-five children born to a mother had she not been exposed to the rHEP (PFO0) is 0.53 (std = 0.08). The same pattern is obtained when we expand our sample to include all reproductive-age women, except now the PFOs are lower for each level of exposure as compared to the sub-sample that contains only mothers. Columns (ii), (iii), and (iv) show the average PFOs for individuals not exposed to the rHEP, exposed but with one HEW, and exposed to two or more HEWs, respectively.

In panel (b), we present the ATEs of exposure to the rHEP with different levels of intensity. Exposure to the rHEP significantly reduces the number of under-five children that would have been born to a typical mother. However, the impact of the rHEP in reducing fertility is higher for individuals who were exposed to two or more HEWs. The program reduces the number of under-five children by 0.27 (std = 0.109) for the average mother had she been exposed to two or more HEWs as opposed to not exposed to the program. Exposure to only one HEW as opposed to non-exposure reduces the number of under-five children that would have been born to a mother by 0.25 (std = 0.107). Although, exposure to two or more HEWs than to only one HEW reduces fertility by a small number, the impact is trivial and statistically insignificant implying that adding one or more HEWs after the first one does not have significant impact in reducing fertility.

The ATEs for sub-populations categorized by the level of exposure are presented in panel (b), columns (ii), (iii), and (iv). For individuals who were not exposed to the rHEP, the ATEs on fertility are much higher compared to the ATEs for the population. For instance, had a typical mother in a control village been exposed to two or more HEWs, she would have given birth to 0.4 less number of children in the past five years. The impact is even higher had she been exposed to only one HEW. Similarly, a typical reproductive-age women who

was not exposed to the rHEP would have 0.31 (std = 0.07) and 0.28 (std = 0.08) number of under-five children had she been exposed to the program with only one HEW and two or more HEWs, respectively. Looking at the sub-population who were exposed to the rHEP but with different number of HEWs, the ATEs due to exposure to higher number of HEWs is always to reduce fertility (negative ATEs). Whereas had a typical woman not been exposed or been exposed to few number of HEWs than she was actually exposed, she would have given birth to higher number of children (positive ATEs). For instance, in column (iv), had a mother who was exposed to two or more HEWs been exposed to only one HEW she would have given birth to 0.21 (std = 0.09) higher number of children.

Table (6) presents the posterior means and standard deviations of PCMOs and the ATEs for the population and the sub-populations under consideration. As shown in panel (a), the average PCMOs under the state of exposure to two or more HEWs is lower than the PCMOs when exposed to only one HEWs or not exposed at all. The potential number of children who dies had a typical mother been exposed to two HEWs (PCMO2) is 0.17 (std = 0.05). Similarly, the average potential number of children that a typical mother would have lost to death had she been exposed to only one HEW is 0.18 (std = 0.05), which is a little higher than PCMO2. The potential number of children who dies had a typical mother not been exposed to the rHEP (PCMOs) is 0.35 (std = 0.07).

The posterior means and standard deviations of the ATEs of the rHEP on child mortality are reported in panel (b). For the population, the ATEs in terms of reduction in child mortality is higher had a typical mother been exposed to higher number of HEWs. Had a typical mother been exposed to two or more HEWs as opposed to non-exposure, the number of children who would have died decreases by 0.18 (std = 0.09). Similarly, as compared to non-exposure, exposure to one HEWs reduces average number of child mortality by 0.17 (std = 0.1). The magnitude of the ATEs is much higher for the sub-population of mothers who were actually not exposed to the program (see column (ii) in panel (b)). For instance, had a typical mother residing in a control village been exposed to two or more HEWs she would have lost 0.32 (std = 0.10) lower number of children to death.

5.3 Heterogeneous Treatment Effects

In this section, we show how the estimated treatment effects vary across individuals by plotting densities and cumulative densities of iTEs. HTEs arise because individuals respond to exposure to a certain level of treatment differently; some gain, some lose and others do not gain or lose. Figures (1)-(3) show the distributions of iTEs of exposure to different number of HEWs on fertility and child mortality for sub-samples consisting of mothers and

reproductive-age women. Although, on average exposure to the rHEP reduces the number of under-five children borne to a mother, the effect varies widely from that of reduction in fertility for some mothers to increased level of fertility for others. As shown in Figure (1), panels (a) and (b), the iTEs due to exposure to two or more HEWs are in the negative domain for 80% of mothers in the sample. For the rest 20%, the iTEs due to exposure to two or more HEWs as opposed to non-exposure are in the positive domain implying that the rHEP increases the number of children borne in the past five years.

Among mothers who were not exposed to the rHEP, more than 90% of them would have given to birth to fewer number of under-five children had they been exposed to two HEWs. The distributions of iTEs on fertility and child mortality can be read off from the plots in a similar fashion. The take home message is that there is considerable level of heterogeneity in magnitude and direction of the treatment effects of the rHEP on fertility and child mortality outcomes among mothers and reproductive-age women.

6 Conclusion

Endogeneity in program placement, heterogeneity in treatment intensity and heterogeneity in treatment effects are common in program evaluations when non-randomized datasets are used. A novel approach to address these issues jointly is Bayesian potential outcomes framework. We apply this method to evaluate the impact of Ethiopia's Health Extension program on fertility and child mortality outcomes using non-randomized household survey data. We control for selection bias on observables by including village level characteristics in the program assignment equation and individual-level and village-level characteristics in the potential outcomes equations. We find significant selection bias on the unobservable dimension as well and ignoring it would significantly bias the estimated ATEs. After controlling for selection bias on observable and unobservable dimensions, the results show that the rHEP has significant impact in reducing fertility and child mortality among mothers and reproductive-age women in rural villages in Ethiopia. However, there are significant HTEs ranging from negative impacts for some individuals to positive impacts for the majority of them in the sample. We recover individual-level treatment effects and present the whole distribution graphically.

References

Admassie, A., D. Abebaw, and A. D. Woldemichael (2009a). Impact evaluation of the ethiopian health services extension programme. *Journal of Development Effectiveness 1*(4), 430–449.

Admassie, A., D. Abebaw, and A. D. Woldemichael (2009b). Impact evaluation of the ethiopian health services extension programme. *Global Development Network Working Paper Series* (22).

Albert, J. H. and S. Chib (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American statistical Association 88*(422), 669–679.

Albert, J. H. and S. Chib (2001). Sequential ordinal modeling with applications to survival data. *Biometrics 57*(3), 829–836.

Angrist, J. D., G. W. Imbens, and D. B. Rubin (1996). Identification of causal effects using instrumental variables. *Journal of the American statistical Association 91*(434), 444–455.

Central Statistical Agency [Ethiopia] and ICF International (2012). *Ethiopia Demographic and Health Survey 2011*. Addis Ababa, Ethiopia and Calverton, Maryland, USA: Central Statistical Agency [Ethiopia] and ICF International.

Chib, S. and B. H. Hamilton (2000). Bayesian analysis of cross-section and clustered data treatment models. *Journal of Econometrics 97*(1), 25–50.

Cowles, M. K. and B. P. Carlin (1996). Markov chain monte carlo convergence diagnostics: a comparative review. *Journal of the American Statistical Association 91*(434), 883–904.

Dehejia, R. H. and S. Wahba (1999). Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *Journal of the American statistical Association 94*(448), 1053–1062.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. *Evaluating the accuracy of sampling-based approaches to calculating posterior moments*. cited By (since 1996)48.

Heckman, J. and S. Navarro-Lozano (2004). Using matching, instrumental variables, and control functions to estimate economic choice models. *Review of Economics and statistics 86*(1), 30–57.

Heckman, J. J., H. Ichimura, and P. E. Todd (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The review of economic studies 64*(4), 605–654.

Heckman, J. J., H. F. Lopes, and R. Piatek (2014). Treatment effects: A bayesian perspective. *Econometric reviews 33*(1-4), 36–67.

Heckman, J. J. and R. Robb Jr (1985). Alternative methods for evaluating the impact of interventions: An overview. *Journal of Econometrics 30*(1), 239–267.

Imbens, G. M. and J. M. Wooldridge (2008). Recent developments in the econometrics of program evaluation. Technical report, National Bureau of Economic Research.

Li, M. and J. Tobias (2005). Bayesian modeling of school effects using hierarchical models with smoothing priors. *Studies in Nonlinear Dynamics & Econometrics 9*(3).

Li, M. and J. L. Tobias (2008). Bayesian analysis of treatment effects in an ordered potential outcomes model. *Advances in econometrics 21*, 57–91.

Munkin, M. K. (2011). The endogenous sequential probit model: An application to the demand for hospital utilization. *Economics Letters 112*(2), 182–185.

Nandram, B. and M.-H. Chen (1996). Reparameterizing the generalized linear model to accelerate gibbs sampler convergence. *Journal of Statistical Computation and Simulation 54*(1-3), 129–144.

Poirier, D. J. (1998). Revising beliefs in nonidentified models. *Econometric Theory 14*(04), 483–509.

Poirier, D. J. and J. L. Tobias (2003). On the predictive distributions of outcome gains in the presence of an unidentified parameter. *Journal of Business & Economic Statistics 21*(2), 258–268.

Ravallion, M. and Q. T. Wodon (1998). *Evaluating a targeted social program when placement is decentralized*. World Bank Publications.

Rosenbaum, P. R. and D. B. Rubin (1983a). Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *Journal of the Royal Statistical Society. Series B (Methodological)*, 212–218.

Rosenbaum, P. R. and D. B. Rubin (1983b). The central role of the propensity score in observational studies for causal effects. *Biometrika 70*(1), 41–55.

Rubin, D. B. (1986). Statistics and causal inference: Comment: Which ifs have causal answers. *Journal of the American Statistical Association 81*(396), 961–962.

Rubin, D. B. (2005a). Bayesian inference for causal effects. In D. Dey and C. Rao (Eds.), *Bayesian Thinking Modeling and Computation*, Volume 25 of *Handbook of Statistics*, pp. 1 – 16. Elsevier.

Rubin, D. B. (2005b). Causal inference using potential outcomes. *Journal of the American Statistical Association 100*(469).

Rubin, D. B. et al. (1978). Bayesian inference for causal effects: The role of randomization. *The Annals of Statistics 6*(1), 34–58.

Tanner, M. A. and W. H. Wong (1987). The calculation of posterior distributions by data augmentation. *Journal of the American statistical Association 82*(398), 528–540.

UNICEF (2013). Committing to child survival: a promise renewed (progress report 2013). new york: Unicef.

Wooldridge, J. M. (2007). Inverse probability weighted estimation for general missing data problems. *Journal of Econometrics 141*(2), 1281–1301.

Workie, N. W. and G. N. Ramana (2013). The health extension program in ethiopia.

Appendices

A Algorithm

Steps:

0. Initialize parameters
1. Draw β^* from

$$\beta^*|Z^*, \Gamma_{-\beta^*}, y, T \sim N\left(D_{\beta^*} d_{\beta^*}, D_{\beta^*}\right), \tag{9}$$

where $D_{\beta^*} \equiv \left[X'(\Sigma^{*-1} \otimes I)X + V_{\beta_0^*}\right]^{-1}$ and $d_{\beta^*} \equiv \left[X'(\Sigma^{*-1} \otimes I)Z^* + V_{\beta_0^*}\mu_{\beta_0^*}\right]$

2. Draw Σ^* from

$$\Sigma^*|Z^*, \Gamma_{-\Sigma^*}, y, T \sim IW\left(n + v, \left[\sum_{i=1}^{n}(Z^* - X\beta^*)(Z^* - X\beta^*)' + vR\right]^{-1}\right) \tag{10}$$

3. Draw z^* independently from conditional truncated normal

$$z_{iT}^*|z_{i2}^*, z_{i1}^*, z_{i0}^*, \Gamma^*, y, T \sim TN_{(\gamma_{T_v}^*, \gamma_{T_{v+1}}^*)}\left(\mu_{T_v}^{*c}, \sigma_{T_v}^{*c}\right) \tag{11}$$

where $\mu_{T_v}^{*c} = \left(w_v\beta_T^* + \Sigma_{Tj}^*\Sigma_{jj}^{*-1}(Z_j^* - X_j\beta_j^*)\right)$ and $\sigma_T^{*c} = \Sigma_{TT}^* - \Sigma_{Tj}^*\Sigma_{jj}^{*-1}\Sigma_{Tj}^{*'}$

4. Draw z_2^* independently from

$$z_{i2}^*|z_{iT}^*, z_{i1}^*, z_{i0}^*, \Gamma^*, y, T \sim \begin{cases} TN_{(\widetilde{\gamma}_{y_{iv}}^*, \widetilde{\gamma}_{y_{iv}+1}^*)}\left(\mu_{2iv}^{*c}, \sigma_{2v}^{*c}\right) & if \quad T_v = 2 \\ N\left(\mu_{2iv}^{*c}, \sigma_{2v}^{*c}\right) & if \quad T_v \neq 2 \end{cases} \tag{12}$$

where $\mu_{2iv}^{*c} = \left(x_{2iv}\beta_2^* + \Sigma_{2j}^*\Sigma_{jj}^{*-1}(Z^* - X_j\beta_j^*)\right)$ and $\sigma_2^{*c} = \Sigma_{22}^* - \Sigma_{2j}^*\Sigma_{jj}^{*-1}\Sigma_{2j}^{*'}$

5. Draw z_1^* independently from

$$z_{i1}^*|z_{iT}^*, z_{i2}^*, z_{i0}^*, \Gamma^*, y, T \sim \begin{cases} TN_{(\widetilde{\gamma}_{y_{iv}}^*, \widetilde{\gamma}_{y_{iv}+1}^*)}\left(\mu_{1iv}^{*c}, \sigma_{1v}^{*c}\right) & if \quad T_v = 1 \\ N\left(\mu_{1iv}^{*c}, \sigma_{1v}^{*c}\right) & if \quad T_v \neq 1 \end{cases} \tag{13}$$

where $\mu_{1iv}^{*c} = \left(x_{1iv}\beta_1^* + \Sigma_{1j}^*\Sigma_{jj}^{*-1}(Z^* - X_j\beta_j^*)\right)$ and $\sigma_1^{*c} = \Sigma_{11}^* - \Sigma_{1j}^*\Sigma_{jj}^{*-1}\Sigma_{1j}^{*'}$

6. Draw z_0^* independently from

$$z_{i0}^*|z_{iT}^*, z_{i2}^*, z_{i1}^*, \Gamma^*, y, T \sim \begin{cases} TN_{(\widetilde{\gamma}_{y_{iv}}^*, \widetilde{\gamma}_{y_{iv}+1}^*)}\left(\mu_{0iv}^{*c}, \sigma_{0v}^{*c}\right) & if \quad T_v = 0 \\ N\left(\mu_{0iv}^{*c}, \sigma_{0v}^{*c}\right) & if \quad T_v \neq 0 \end{cases} \tag{14}$$

where $\mu_{0iv}^{*c} = \left(x_{0iv}\beta_0^* + \Sigma_{0j}^*\Sigma_{jj}^{*-1}(Z^* - X_j\beta_j^*)\right)$ and $\sigma_0^{*c} = \Sigma_{00}^* - \Sigma_{0j}^*\Sigma_{jj}^{*-1}\Sigma_{0j}^{*'}$
Cycle through steps [1] - [6] until convergence.

Table 1: Descriptive statistics of fertility and child mortality by treatment status

	All	No Exposure	Exposed to 1 HEW	Exposed to 2 HEWs
			Mothers (Age: 14-45)	
No. of under-five children	1.13	1.20	1.08	1.06
	(0.77)	(0.80)	(0.72)	(0.75)
No. of live children	4.22	4.25	4.26	4.17
	(2.36)	(2.37)	(2.23)	(2.47)
No. of dead children	0.86	0.92	0.79	0.80
	(1.20)	(1.25)	(1.13)	(1.19)
Total no. of children	5.08	5.17	5.05	4.98
	(2.81)	(2.83)	(2.74)	(2.86)
No. of Observations	2490	1192	654	624
			Women (Age: 14-45)	
No. of under-five children	0.78	0.80	0.74	0.78
	(0.82)	(0.86)	(0.78)	(0.79)
No. of live children	2.82	2.75	2.81	2.98
	(2.77)	(2.78)	(2.71)	(2.81)
No. of dead children	-.-	-.-	-.-	-.-
Total no. of children	3.39	3.36	3.34	3.56
	(3.32)	(3.36)	(3.27)	(3.30)
No. of Observations	3708	1842	991	875
			Women (Age: 50+)	
No. of under-five children	0.07	0.08	0.03	0.08
	(0.27)	(0.30)	(0.18)	(0.27)
No. of live children	3.93	3.99	3.41	4.44
	(3.23)	(3.18)	(3.11)	(3.44)
No. of dead children	1.67	1.67	1.59	1.80
	(2.10)	(2.03)	(2.21)	(2.18)
Total no. of children	5.60	5.66	5.00	6.24
	(4.33)	(4.23)	(4.30)	(4.54)
No. of Observations	469	235	127	104

Table 2: Summary statistics of individual level characteristics (all women aged 14-45)

	HEW = 0 Mean (Std.Dev.)	HEW = 1 Mean (Std.Dev.)	HEW = 2+ Mean (Std.Dev.)
Age	26.8 (8.87)	27.2(8.71)	27.1 (8.56)
Married	65%	66%	70%
Household Head	5%	5%	5%
Spouse	64%	51%	80%
Adult Child	28%	41%	11%
No. of Female (14-45) members	1.80 (0.96)	1.73 (0.89)	1.63 (0.84)
Can Write and Read	29%	29%	28%
Domestic Work	70%	63%	74%
Student	24%	31%	8%
Orthodox Christian	49%	81%	55%
Muslim	37%	6%	30%
Wealth Index: Quintile 1	19%	13%	19%
Wealth Index: Quintile 2	14%	25%	19%
Wealth Index: Quintile 3	15%	29%	9%
Wealth Index: Quintile 4	23%	8%	18%
Wealth Index: Quintile 5	29%	25%	34%
Low-land (*kola*)	44%	19%	50%
Access to pipe water	5%	6%	4%
Distance to district Town (Km)	21.9 (15.7)	20.4 (15.4)	25.8 (20.3)
Distance to tarmac road (Km)	34.8 (44.5)	23.7 (17.7)	35.5 (27.3)
Village-level literacy rate	35%	40%	40%
Amhara Region	14%	40%	35%
Oromia Region	64%	28%	43%
SNNPR Region	23%	32%	22%
No. of Observations	1,842	991	906

Table 3: Summary statistics of village-level characteristics

	All Mean (SD)	Control Mean (SD)	Treatment Mean (SD)
Adult Basic Education centers	17%	18%	17%
Literacy Rate	38%	35%	41%
Primary School	78%	77%	80%
Distance to Tarmac Road	34.52 (40.75)	39.30 (50.23)	27.86 (22.35)
Low Land	37%	43%	32%
Access to Pipe water	4%	3%	5%
Water borne diseases	69%	65%	73%
Health Centers	8%	8%	8%
Amhara Region	28%	17%	36%
oromia Region	46%	58%	35%
SNNP Region	27%	25%	29%
Zeway Dugda district	13%	20%	8%
Gimbichu district	12%	13%	11%
Dedo district	6%	8%	3%
Sodo district	13%	12%	14%
Mana district	8%	8%	8%
Kuyu district	7%	8%	6%
Bule district	13%	5%	21%
Enemay district	7%	7%	8%
Bahirdar Zuria district	6%	7%	5%
Jawi district	15%	12%	18%
No. of villages	126	60	66

Table 4: Posterior means of parameters from ordered probit model of assigning 0, 1, 2+ HEWs

Variables	Fertility		Child Mortality
	Mothers	Women	Mothers
Adult basic education centers	0.15	0.22	0.19
	(0.08)	(0.07)	(0.09)
Literacy eate	0.11	0.13	0.12
	(0.02)	(0.02)	(0.02)
Primary school	0.09	0.15	0.06
	(0.07)	(0.06)	(0.07)
Distance to tarmac road	-0.01	-0.01	-0.01
	(0.00)	(0.00)	(0.00)
Lowland area	-0.12	-0.19	-0.11
	(0.10)	(0.08)	(0.10)
Access to pipe water	0.30	0.28	0.57
	(0.13)	(0.11)	(0.13)
Water borne diseases	0.44	0.47	0.58
	(0.06)	(0.05)	(0.07)
Health centers	0.02	-0.01	0.00
	(0.10)	(0.08)	(0.10)
Oromia region	0.02	0.00	0.01
	(0.44)	(0.45)	(0.45)
Amhara region	-0.02	-0.11	-0.06
	(0.26)	(0.22)	(0.26)
Intercept	-1.13	-1.27	-1.26
	(0.14)	(0.12)	(0.15)

Note: District dummies included but not reported here.

Table 5: Posterior means of PFOs and ATEs

	(i) Population		(ii) Not Exposure		(iii) Exposed to 1 HEW		(iv) Exposed to 2 HEW	
	Mothers	Women	Mothers	Women	Mothers	Women	Mothers	Women
Panel (a)								
PFO2	0.256	0.179	0.238	0.158	0.166	0.116	0.383	0.290
	(0.060)	(0.041)	(0.059)	(0.039)	(0.046)	(0.031)	(0.076)	(0.057)
PFO1	0.277	0.196	0.182	0.127	0.553	0.385	0.173	0.130
	(0.056)	(0.039)	(0.043)	(0.029)	(0.092)	(0.064)	(0.044)	(0.032)
PFO0	0.526	0.368	0.640	0.437	0.354	0.241	0.489	0.369
	(0.079)	(0.054)	(0.088)	(0.060)	(0.066)	(0.043)	(0.077)	(0.055)
Panel (b)								
ATE {2 vs 0}	-0.270	-0.189	-0.402	-0.279	-0.188	-0.125	-0.106	-0.079
	(0.109)	(0.073)	(0.115)	(0.077)	(0.088)	(0.058)	(0.117)	(0.084)
ATE {1 vs 0}	-0.249	-0.173	-0.458	-0.310	0.199	0.144	-0.316	-0.239
	(0.107)	(0.073)	(0.106)	(0.072)	(0.120)	(0.080)	(0.096)	(0.069)
ATE {2 vs 1}	-0.021	-0.017	0.056	0.031	-0.387	-0.268	0.210	0.160
	(0.091)	(0.063)	(0.083)	(0.055)	(0.105)	(0.074)	(0.093)	(0.069)

Note: PFO2 = Potential fertility outcome had a typical woman been exposed to the rHEP with 2 HEWs. PFO1 = Potential fertility outcome had a typical woman been exposed to the rHEP with 1 HEWs. PFO0 = Potential fertility outcome had a typical woman not been exposed to the rHEP.

Table 6: Posterior means of PFOs and ATEs

	(i) Population	(ii) No Exposuer	(iii) 1 HEW	(iv) 2 HEW
Panel (a)				
PCMO2	0.171	0.160	0.099	0.265
	(0.052)	(0.052)	(0.036)	(0.069)
PCMO1	0.184	0.121	0.373	0.107
	(0.050)	(0.038)	(0.084)	(0.037)
PCMO0	0.350	0.439	0.240	0.296
	(0.071)	(0.081)	(0.056)	(0.065)
Panel (b)				
ATE {2 vs 0}	-0.179	-0.279	-0.141	-0.031
	(0.094)	(0.103)	(0.072)	(0.100)
ATE {1 vs 0}	-0.167	-0.318	0.133	-0.189
	(0.095)	(0.097)	(0.107)	(0.080)
ATE {2 vs 1}	-0.012	0.039	-0.274	0.158
	(0.082)	(0.074)	(0.095)	(0.083)

Note: PCMO2 = Potential child mortality outcome had a typical woman been exposed to the rHEP with 2 HEWs. PCMO1 = Potential child mortality outcome had a typical woman been exposed to the rHEP with 1 HEWs. PCMO0 = Potential child mortality outcome had a typical woman not been exposed to the rHEP.

Table 7: Estimated correlations and cut-points of potential fertility outcomes equation

	Mothers	Women
(a) Correlations between HEWs assignment and		
PFO2 ($\widehat{\rho_{2T}}$)	0.52	0.40
	(0.02)	(0.02)
PFO1 ($\widehat{\rho_{1T}}$)	0.17	0.11
	(0.03)	(0.03)
PFO0 ($\widehat{\rho_{0T}}$)	-0.59	-0.46
	(0.02)	(0.02)
(b) Correlations among PFOs		
PFO2 and PFO1($\widehat{\rho_{21}}$)	0.09	0.05
	(0.04)	(0.03)
PFO2 and PFO0 ($\widehat{\rho_{20}}$)	-0.47	-0.32
	(0.03)	(0.02)
PFO1 and PFO0 ($\widehat{\rho_{10}}$)	-0.22	-0.14
	(0.03)	(0.03)
(c) Estimated cut-points		
rHEP assignment ($\widehat{\gamma_2}$)	0.11	0.11
	(0.00)	(0.00)
PFO2 ($\widehat{\widetilde{\gamma_2}}$)	0.10	0.10
	(0.00)	(0.00)
PFO1 ($\widehat{\widetilde{\gamma_1}}$)	0.11	0.10
	(0.00)	(0.00)
PFO0 ($\widehat{\widetilde{\gamma_0}}$)	0.10	0.10
	(0.00)	(0.00)

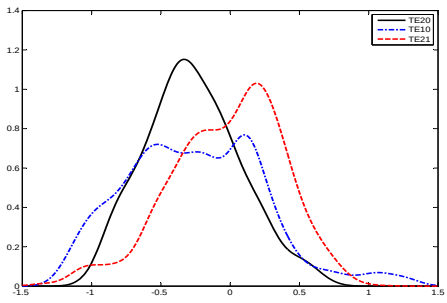
Note: PFO2 = Potential fertility outcome had a typical woman been exposed to the rHEP with 2 HEWs. PFO1 = Potential fertility outcome had a typical woman been exposed to the rHEP with 1 HEWs. PFO0 = Potential fertility outcome had a typical woman not been exposed to the rHEP.

Table 8: Estimated correlations and cut-points of potential child mortality outcomes equation

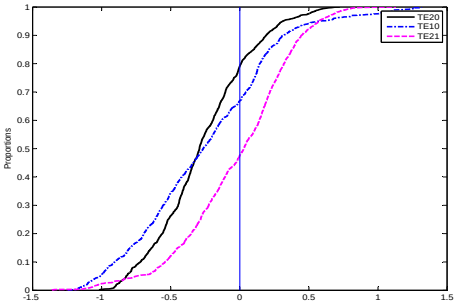
	Mothers
(a) Correlations between HEWs assignment	
PCMO2 ($\widehat{\rho_{2T}}$)	0.39
	(0.03)
PCMO1 ($\widehat{\rho_{1T}}$)	0.11
	(0.04)
PCMO0 ($\widehat{\rho_{0T}}$)	-0.46
	(0.03)
(b) Correlations among PFOs	
PCMO2 and PFO1 ($\widehat{\rho_{21}}$)	0.05
	(0.03)
PCMO2 and PFO0 ($\widehat{\rho_{21}}$)	-0.31
	(0.03)
PCMO1 and PFO0 ($\widehat{\rho_{21}}$)	-0.13
	(0.03)
(c) Estimated Cut-points	
rHEP assignment ($\widehat{\gamma_2}$)	0.11
	(0.00)
PCMO2 ($\widehat{\widetilde{\gamma}_2}$)	0.10
	(0.00)
PCMO1 ($\widehat{\widetilde{\gamma}_1}$)	0.10
	(0.00)
PCMO0 ($\widehat{\widetilde{\gamma}_0}$)	0.10
	(0.00)

Note: PCMO2 = Potential child mortality outcome had a typical woman been exposed to the rHEP with 2 HEWs. PCMO1 = Potential child mortality outcome had a typical woman been exposed to the rHEP with 1 HEWs. PCMO0 = Potential child mortality outcome had a typical woman not been exposed to the rHEP.

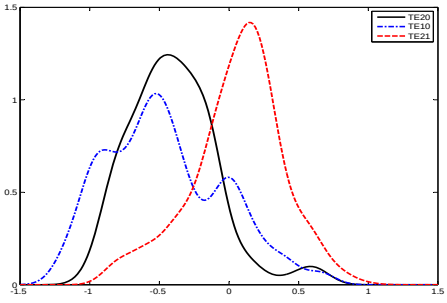
Figure 1: Density and Cumulative Density Plots of iTEs of rHEP on Fertility Outcomes of Mothers



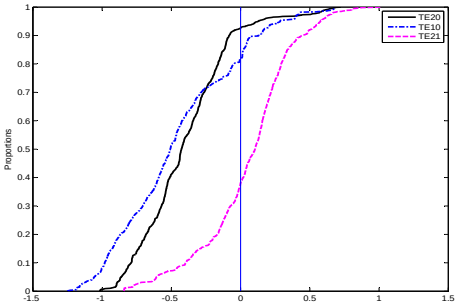
(a) Density: population



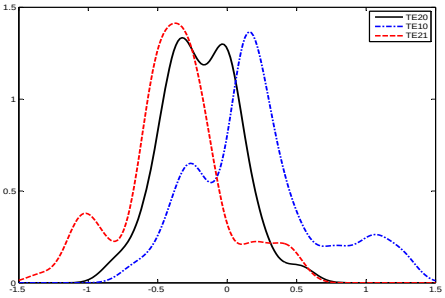
(b) Cumulative density: population



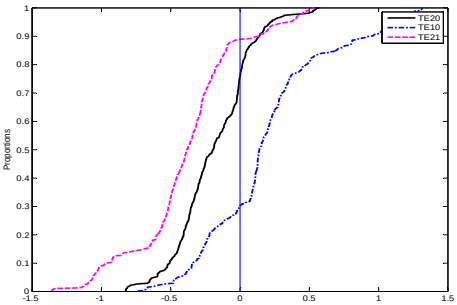
(c) Density: not exposed to HEWs



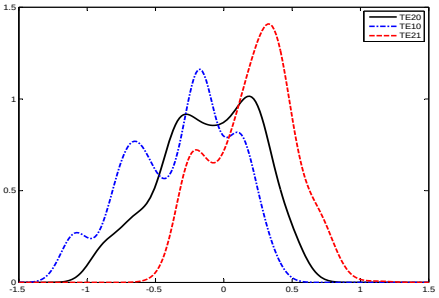
(d) Cumulative density: not exposed to HEWs



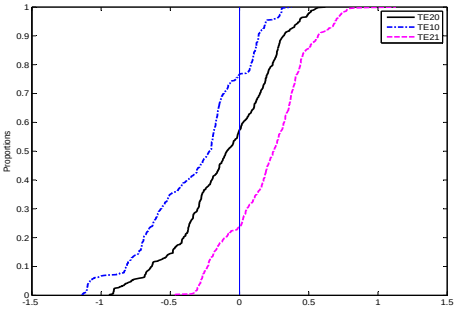
(e) Density: exposed to 1 HEW



(f) Cumulative density: exposed to 1 HEW

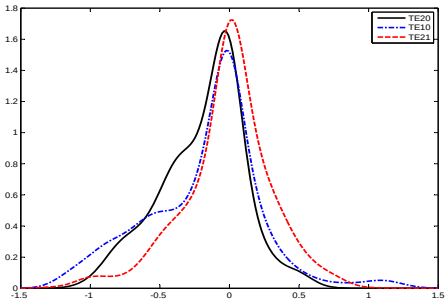


(g) Density: exposed to 2 or more HEWs

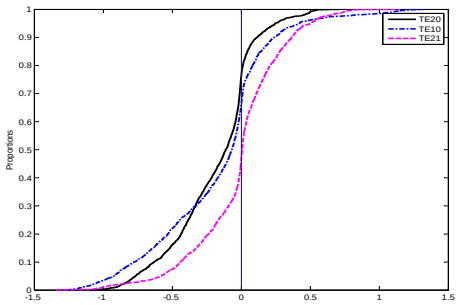


(h) Cumulative density: exposed to 2 or more HEWs

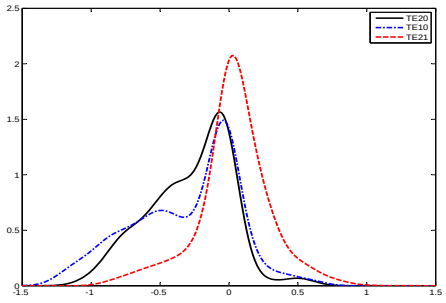
Figure 2: Density and Cumulative Density Plots of iTEs of rHEP on Fertility Outcomes of Reproductive-Age Women



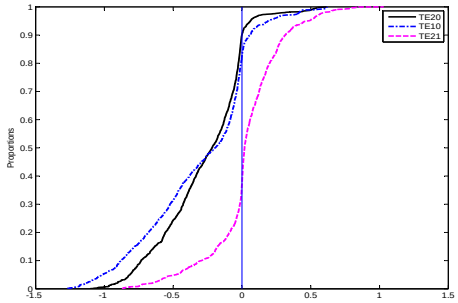
(a) Density: population



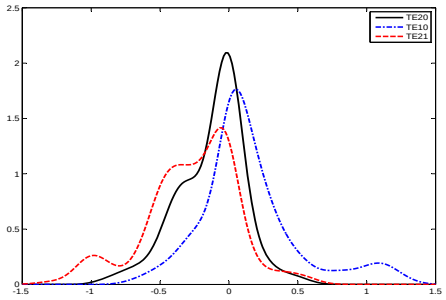
(b) Cumulative density: population



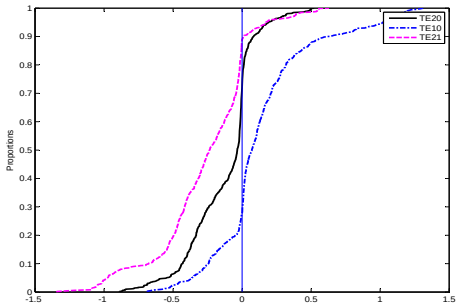
(c) Density: not exposed to HEWs



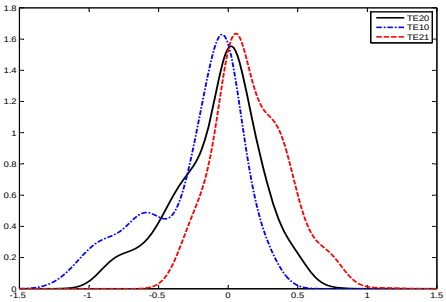
(d) Cumulative density: not exposed to HEWs



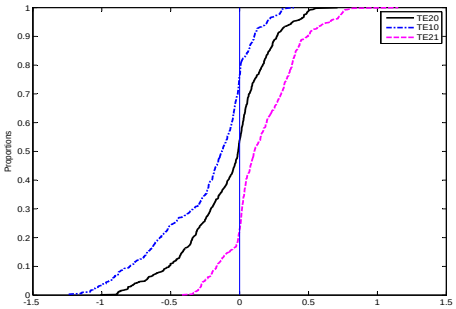
(e) Density: exposed to 1 HEW



(f) Cumulative density: exposed to 1 HEW

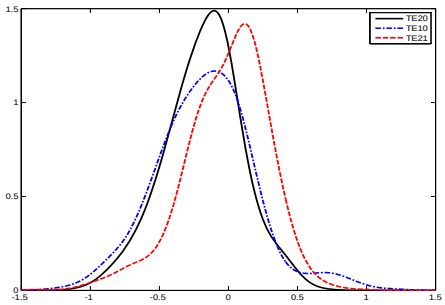


(g) Density: exposed to 2 or more HEWs

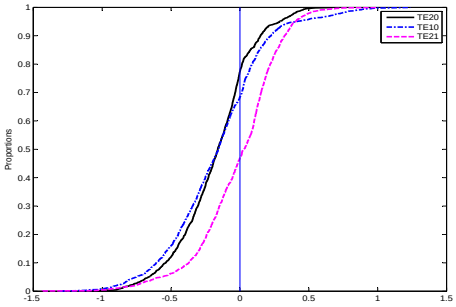


(h) Cumulative density: exposed to 2 or more HEWs

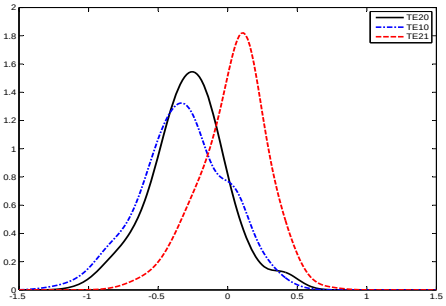
Figure 3: Density and Cumulative Density Plots of iTEs of rHEP on Child Mortality Outcomes



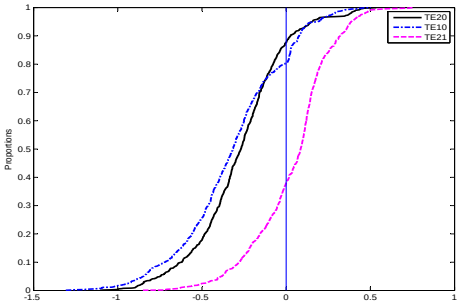
(a) Density: population



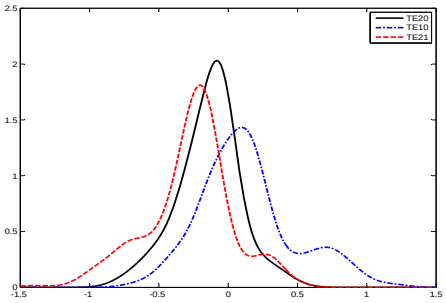
(b) Cumulative density: population



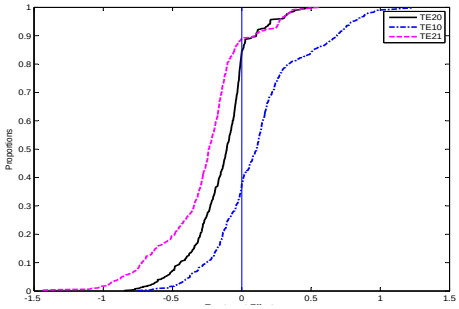
(c) Density: not exposed to HEWs



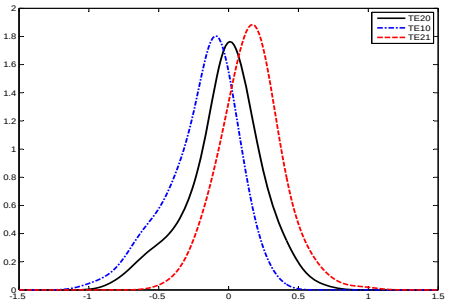
(d) Cumulative density: not exposed to HEWs



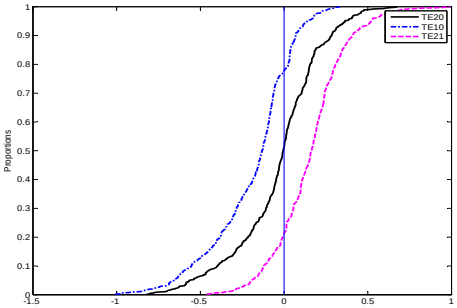
(e) Density: exposed to 1 HEW



(f) Cumulative density: exposed to 1 HEW



(g) Density: exposed to 2 or more HEWs



(h) Cumulative density: exposed to 2 or more HEWs