Behaviors exposed: Panel data evidence on environmental health externalities in rural India

Sumeet R. Patil and Subhrendu K. Pattanayak[1]

Environmental policies can improve human health only if households cooperate. This is especially true for microbiological contamination of water that causes infections such as diarrhea, which kills over 1.5 million children annually and limits child growth. Rigorous econometric analyses of source water quality, household behaviors, and microbiological contamination remain scarce. We use a unique panel data of approximately 37,000 observations from 240 rural villages spanning four seasons in rural India, the world leader in diarrhea deaths. We combine survey data on household behaviors and community infrastructure with laboratory tests of microbial contamination. Applying a household production framework, we specify a structural panel model of E.coli exposure as a function of averting behaviors (which respond to community infrastructure), source water quality, and other exogenous factors. We find that in addition to averting behaviors, source water quality, community sanitation and village diarrhea prevalence have direct and indirect impacts (by changing behaviors) on water quality. We deliver among the first credible estimates of environmental health externalities and prevalence elasticities, and show how to estimate total effects for policy evaluation. We argue that any attempt to reduce diarrhea through water quality must combine infrastructure with information and education for behavior change.

Keywords: environmental health, microbial exposure, e.coli, diarrhea, two-stage estimation, Maharashtra, India

---

[1] Address correspondence to Subhrendu K. Pattanayak, Associate Professor, Duke University (subhrendu.pattanayak@duke.edu). Sumeet Patil is the Executive Director of NEERMAN, Mumbai, India. The study was funded by the World Bank. We are grateful to seminar participants at the Institute of Economic Growth (Delhi), RTI International, University of California (Berkeley), The World Bank, and the Indian Council for Medical Research for comments on initial results..

1. Introduction

Environmental protection can improve human health. Yet such gains are feasible only if households cooperate and complement environmental programs individually (e.g., preventing and treating infectious diseases) and collectively (e.g., safe handling of water by many households to reduce community contamination). This is especially true for microbiological contamination of water because of the complex web of fecal-oral exposure pathways for infectious disease such as diarrhea [3, 20]. Diarrhea is the poster child for environmental health both because it kills over 1.5 million children worldwide annually [26] and is caused primarily by environmental contamination. Although progress in treating diarrhea (e.g., discovery of oral rehydration salts) has reduced diarrhea mortality, we have made little progress in reducing morbidity and therefore in improving children's short-term human capital and long term wellbeing. Even though inadequate water and sanitation infrastructure and unsafe behaviors related to handling of food and water are collectively blamed for high rates of diarrhea[11, 26], we do not know enough about the specific pathways and the cost-effectiveness of environmental interventions [17-19]. Despite early attempts to study these linkages [6, 7, 27, 29], rigorous empirical evaluations of the linkages between source water quality, household behaviors, and microbiological contamination of in-house drinking water (DW) remain scarce. Thus, we return to this old yet persistent problem by analyzing a rich ($k > 15$) large ($n > 9500$ household) panel (4 seasons) data set from rural India, a country which leads the world with four hundred thousand child diarrhea deaths annually.

DW quality improvements or reductions in in-house microbial contamination can substantially reduce diarrhea – estimates suggest 6% to 50% reductions in diarrhea due to DW quality interventions [23]. However, DW quality depends on at least two factors: (a) source water quality [29, 7] and household behavior [17]. The thin empirical literature has focused mainly on water safety behaviors, without adequately considering source water contamination or how behaviors modify the impact of water and sanitation infrastructure [13, 16]. Other studies find that household with safe water management practices have lower diarrhea [12, 14, 21, 23]. A meta-analysis finds that household behaviors such as chemical or solar disinfection, boiling, filtration, and safe storage at point-of-use are very effective in reducing diarrhea [8].

Overall, the literature is more focused on role of averting behaviors in reducing diarrhea, but not on clarifying the exposure pathways, for example, by identifying which behavior or set of behaviors have significant impacts on DW quality. Such clarifications can help improve policy and program implementation by targeting key exposure risks. While some researchers have examined one or two behaviors related to water storage and treatment, fewer studied the combined effect of water, sanitation and hygiene (WSH) related behaviors. Also, the literature is thin on the effect of community WSH behaviors on DW quality at home.

It is unclear whether source water quality or household behavior is more important. Many suggest that in-house DW quality improvements (and thus household practices) are more effective than source quality improvement. For example, a metal analysis finds that household level or point-of-use water quality interventions are more effective in reducing diarrhea than source water supply improvements [2]. These finding challenge an older claim that source water quality is more important than household water quality because the “type” of pathogens introduced at the household level through cross contamination may not be as harmful as the type of pathogens existing in the source water [27]. Muddying the waters is the argument that these comparisons are incorrect because community level studies often exclude household behavior in determining DW quality [2].

Even if we accept that community based water, sanitation and hygiene (WSH) programs including those improving source water quality are effective [2, 6, 27], the role of household behaviors remains key in determining the success of these community WSH interventions. For example, studies find that there are no significant health benefits from improving public water supply [11, 28]. This could be because complementary household behaviors were excluded from these studies [14, 18, 23].

In section 2, we present our conceptual model, the structural specification, and panel data estimation methods. We follow the early averting behavior literature by using a household production framework (HPF) to model DW quality production as a two stage process. First we examine household choice of averting behaviors (A) in response to community WSH infrastructure, prevailing illness, and source water quality. Pattanayak and Pfaff [17] argue that households adopt their behaviors in response to exogenous interventions and can substitute or compensate the gains from such interventions. Second we model the microbial contamination of DW (Q) as a function of A, R and other exogenous inputs. We can estimate unbiased *partial effects* of these production inputs in the second stage because we model the potentially *endogenous* A in the first stage. Finally, we can estimate the *total effect* or *mutatis mutandis effect* of WSH intervention and other exogenous factors.[2] The total effect of WSH interventions is relevant because it not only demonstrates the true success of the intervention, but also captures how households can substitute or complement the policy.

The panel data (n = approximately 37,000) comes from an evaluation of community-led water and sanitation programs in Maharashtra, India [18,19]. Section 3 describes data collection and the key variables related to A, R, Q and various control variables. We surveyed

---

[2] Total effects of an input is the sum of: (1) the partial effect of that input on Q; and (2) the product of the partial effects of that input on A <u>and</u> the partial effect of A on Q. Please see Section 2 (Equation 3) for details.

approximately 50 households from each of 240 rural villages in two seasons (dry and rainy) during 2005 and 2007. In addition to rich information on household behaviors related to water quality – e.g., boiling, chemical treatment, storage, handling, handwashing, and toilet use – and community infrastructure – e.g., water source, garbage and wastewater disposal, we also tested water samples in laboratories for microbial contamination of DW quality (E. coli). To our knowledge, it is unique to have a large sized panel data set on socio-demographic, behavioral and water quality variables. Sections 4 and 5 present the results and key conclusions.

2. Methods

2.1 The Household Production Model

We adopt parts of the health production function by Harrington and Portney [10], Dickie and Gerking [5] and Berman et al. [1] to produce Q as a function of A, R and other exposure factors. We assume that a household maximizes utility by allocating its limited time and income across leisure (*L*) and health (represented by the number of sick days, *S*) given household preferences proxied by socio-economic vector (α). We assume that *S* is a direct function of Q and other health related variables (γ). We neglect composite consumption good because the focus of this paper is not on valuing water quality in terms of composite good prices. Production function for Q is assumed to be twice differentiable, continuous, and convex. A has a cost in terms of time, material and money (P_A). The household time and budget constraint ensures that time value of *L* and *S* valued at the wage rate (ω), and expenditure on a unit of A (P_A) is lower than exogenous income transfers (I) and value of total time endowment (T) at wage rate (ω). The Lagrangian for this model is presented below.

$$Max\, \mathcal{L} = U\{L, S[\mathrm{Q}(\mathrm{A}, \mathrm{R}, \beta), \gamma], \alpha\} + \lambda\{\mathrm{I} + \omega\mathrm{T} - \omega\mathrm{L} - \omega S[\mathrm{Q}(\mathrm{A}, \mathrm{R}, \beta), \gamma] + \mathrm{P_A}\}$$

Assuming that $U_L>0$, $U_Q<0$, $U_{LL} < 0$, $U_{QQ} > 0$, $Q_A < 0 < Q_R$, $Q_{AA}>0$, and $Q_{RR} = 0$, we solve the Lagrangian to obtain the first order conditions (FOC). Reduced form characterization of FOC is:

$$U_S S_Q Q_A - U_L S_Q Q_A - \frac{U_L}{\omega} P_A = 0 \tag{1}$$

To analyze the total effects of an exogenous factor on Q, we totally differentiate Equation (1) to obtain Equation (2) after rearranging terms. We have taken example of R in equations below, but the mathematics remains the same for other exogenous variables. Equation (3) is a simplified representation after accounting for expected signs of terms in Equation (2).

$$\frac{dQ}{dR} = \frac{\left\{\begin{array}{l} \left\langle \left[(U_{SQ} - U_{LS} S_Q).S_Q + (U_S - U_L) S_{QQ} \right] Q_A - \frac{P_A}{\omega} U_{LS} S_Q \right\rangle . Q_R + \\ \left\langle \left[(U_{SQ} - U_{LS} S_Q).S_Q + (U_S - U_L) S_{QQ} \right] Q_A - \frac{P_A}{\omega} U_{LS} S_Q \right\rangle . Q_A A_R \end{array}\right\}}{\left\langle \left[(U_{SQ} - U_{LS} S_Q).S_Q + (U_S - U_L) S_{QQ} \right] Q_A - \frac{P_A}{\omega} U_{LS} S_Q \right\rangle} \tag{2}$$

$$\frac{dQ}{dR} = Q_R^{+} + Q_A^{-} A_R^{?} \tag{3}$$

Equation (3) shows that the effect of R on Q depends upon two components. The first effect is the direct partial effect of R on Q. The second effect is the indirect effect of R on A, and effect of A on Q. Most analyses exclude the second term so that the results represent change in Q when R is changed *ceteris Paribus* (when all other factors are held constant). However, for effective policy and implementation, we must consider *all structural changes possible in the system of equations* to deduce total or *mutatis muntandis* effect of a policy change [9]. For example, improving R need not improve Q if the second term in Equation (3) is negative and larger than the first term. On the contrary, if the second term in Equation (3) is positive, A compliment the gains of improvement in R.

2.2 Structural Models

We specify the following structural models for A and Q as:

$$A = \alpha_0 + \alpha_1 R + \alpha_2 \text{ exsafety} + \alpha_3 \text{ hhsize} + \alpha_4 \text{ headedu} + \alpha_5 \text{ scst} + \alpha_6 \text{ ln\_totexpd} + \alpha_7 \text{ num\_farm} + \alpha_8 \text{ htype} + \alpha_9 \text{ wsh\_need} + \alpha_{10} \text{ wsh\_disease} + \alpha_{11} \text{ wsh\_pollu} + \alpha_{12} \text{ idcause} + \alpha_{13} \text{ iec\_msg} + \alpha_{14} \text{ HHdiarr2} + \alpha_{15} \text{ villdiarr} + \alpha_{16} \text{ paywater} + \alpha_{17} \text{ improvedw} + \alpha_{18} \text{ lpcd2} + \alpha_{19} \text{ JS} + \alpha_{20} \text{ pct\_od} + \alpha_{21} \text{ pct\_sw} + \alpha_{22} \text{ pct\_ww} + \alpha_{23} \text{ drain} + \alpha_{24} \text{ villwqprob} + \alpha_{25} \text{ vwsc} + \alpha_{26} \text{ villpgm} + \alpha_{27} \text{ villgrps} + \alpha_{28} \text{ round2} + \alpha_{29} \text{ round3} + \alpha_{30} \text{ round4} + u_1 \quad (4)$$

$$Q = \beta_0 + \beta_1 R + \beta_2 A + \beta_3 \text{ hhsize} + \beta_4 \text{ improvedw} + \beta_5 \text{ lpcd2} + \beta_6 \text{ pct\_od} + \beta_7 \text{ pct\_sw} + \beta_8 \text{ pct\_ww} + \beta_9 \text{ drain} + \beta_{10} \text{ villdiarr} + \beta_{11} \text{ round2} + \beta_{12} \text{ round3} + \beta_{13} \text{ round4} + u_2 \quad (5)$$

Where A = different WSH related averting behavior indices; R = source water e. coli contamination measured as $\log_{10}$ of CFU/ml (ln_villecoli); and Q = household DW e. coli contamination in $\log_{10}$ of CFU/ml (ln_ecoli) or as a dummy indicating contamination (HHecoli). Data used to estimate about models is a panel dataset collected in 4 rounds over two years. To capture the effect of time and season we use the following dummies: round2 = monsoon of 2005, round 3 = summer of 2007; and round4 = monsoon of 2007. Round1 = summer of 2005 is implicit and not specified in the model. Table I explains all variables used in the model and their descriptive statistics.

2.3 Estimation of the models

We use ordered probit with clustered standard errors (SE) and their random effects specification (RE) to model different indices for A related to water safety, hand washing, and sanitation behaviors (each as 4-point index) and a combined index for all these WSH indices (a 10-point index). We adjust SE for clustering at the village level because the households are clustered at a village level and most WSH interventions are at the village level. REM will estimate SE more accurately by accounting for autocorrelation across multiple rounds of data collection. To estimate the production model for Q, we use OLS with clustered SE and their RE when we use continuous *ln_ecoli* as an indicator for Q. We used probit with clusted SE and their RE when indicator for Q is a dummy variable (HHecoli). We have also estimated fixed effects model for each stage and found that these produce qualitative similar results (available from authors upon request).

To estimate Equations (4) and (5) by above models, the error terms u_1 and u_2 should not be correlated with explanatory variables; otherwise we face the problem of endogeneity. Endogeneity can mainly arise if A and Q are simultaneously determined or if omitted variables from any of the models for Q or A are correlated with the error term of the other model.

Although Equation (5) includes A (which is endogenous), Q and A are most likely *not* simultaneously determined (that is Q and u_2 are not correlated to cause endogeneity). In typical *health production* models, behaviors and health are simultaneously determined because households can observe health and modify behaviors accordingly [22, 4]. However, microbial water quality is not observable to the households to modify their behaviors in response to changing Q[3]. If we argue that household diarrhea episode is a proxy for Q, then they may increase A in response to diarrhea episodes. If true, then A and Q will be simultaneously determined causing endogeneity bias. However, if Q is not affected by diarrhea episodes in the household *or* if Q is not affected almost immediately in response to household diarrhea - a more plausible scenario - then Q and A are not simultaneously determined. Therefore, we can estimate models for Q and A independently even if households become aware of relationship between Q and diarrhea over time.

For sake of robustness of results, we relax above argument that Q and A can be independently estimated, and model Equations (4) and (5) using instrumental variables (IV) or two stage least square methods (2SLS) methods. 2SLS model also deals with endogeneity concerns when omitted / unobservable factors included in u_2 are correlated with A or omitted / unobservable factors included in u_1 are correlated with Q. To estimate 2SLS model, explanatory variables that are included in Equation (4), but not in Equation (5) serve as IV for A. We confirm validity of 2SLS results by conducting test for endogeneity and overidentification.

We calculate total or *mutatis muntandis* effect of exogenous factors using the estimated coefficients as shown in Equation (6) for the example of R using these coefficients estimated from Equations (4) and (5). *mutatis muntandis* effect of other exogenous exposure factors can be similarly calculated.

$$\frac{dQ^*}{dR} = \hat{\beta}_1 + \hat{\beta}_2\hat{\alpha}_1 \qquad (6)$$

We can also obtained the *mutatis muntandis* effect for R using a reduced form model where A in Equation (5) is replaced explanatory variables from Equation (4). As argued by Ford

---

[3] Households can identify physical and chemical contamination through taste, smell, color, but microbial contamination is not observable.

and Jackson [9], the reduced form estimates should be the same as those calculated as per Equation (6) if Equations (4) and (5) are independent. We prefer calculation of mutatis muntandis effect over reduced form estimates because they are more useful from policy and implementation perspective. For example, we can analyze the magnitude of direct versus indirect effect of improvement in R on Q. And, the partial effects of exogenous factors on A (Equation 4) can inform behavior change communication strategies. However, for academic interest we compare calculated and reduced form estimates of *mutatis muntandis* effects. We estimate all above models in STATA [24].

3.0 Data and Variables

3.1 Data Collection

We use data from a health impact evaluation study of WSH interventions in Maharashtra, India. Maharashtra is among the largest and more progressive Indian states with a population of 97 million living in approximately 44,000 villages from 33 districts. Our study area is confined to Buldana, Nashik, Osmanabad, and Sangli districts from four geographically different regions of Maharashtra. Our panel consists of approximately 9,300 households from 242 villages surveyed in 2005 and 2007 in summer and monsoon season (5 week data collection in each season). Of these, we have household water quality results (Q) for 50% households (approx 4,500). In 2007, 80 villages had received a community demand driven government program (*Jalswarajya*) to provide adequate and safe water and sanitation services and to promote safe sanitation and hygiene behaviors.

We collected data using three survey instruments. Main household survey was administered to the primary care giver (PCG) of under 5 year (U5) child. The survey instrument collected information on priorities, knowledge, perceptions, socio-economics, health, WSH facilities, and WSH related averting behaviors. Community survey was administered to the formal or informal village head or other key informants. These key informants are typically village *sarpanch* (31%), *panchayat* members (15%), and informal leaders (13%). Almost all respondents had at least primary schooling. We collect information on village socioeconomics and demographics, land use, village institutions and infrastructure, and WSH facilities and schemes. Water quality survey was randomly administered in approximately 50% of the surveyed households. We lab tested DW samples from these households as well in-use sources of DW in the village. The samples were collected in sterilized containers and transported to a nationally accredited lab within 24 hours of collection in ice boxes. The lab used USEPA approved spread plate method using CHROMagar media to enumerate total coliform and e. coli concentrations (CFU/ml). We also tried to link the source water quality with specific households wherever we can identify the exact source from which the households collected their DW. Details about the study design, survey, implementation, and WSH interventions and results of the evaluation are published elsewhere [19, 18].

3.2 Water Quality Indicators

We use e. coli as water quality indicators because it is a universally accepted indicator for fecal contamination of DW; and hence, a health risk indicators. E. coli is preferred over total coliform - a heterogeneous coliform group which need not indicate health risk - for untreated rural water supply [15]. Results for Q are available for approximately 19,000 households from 4 rounds of survey. Results for R are available for 2,900 sources from 242 villages in 4 rounds (on average 3 sources per village per round). For majority of the households where we tested for Q and those who use public water sources, we can identify the source used; and thus, use the specific value for R. For other households, we use average e. coli contamination by the type of main DW source as an indicator for R. In a few cases (<1%) where source water quality measure by type of DW source is not available, we used overall village e. coli average as R. Above substitution for the indicator of R allows us to use our complete sample of over 39,000 households in modeling Equation (4). However, to model Equation (5), we can use only a subset of 19,000+ households where we have results for Q.

We log transformed ($\log_{10}$) e. coli results for R (*ln_villecoli*) and Q (*ln_ecoli*) to reduce the effect of disproportionately high tail-end values[4]. For Q, we also create a dummy that indicartes fecal / e. coli contamination or not (*HHecoli*). We find that fecal contamination of household DW is slightly lower than that of the source water (see Table I). In general fecal contamination of water (household and source) is higher in monsoon than summer as expected, except Monsoon of 2007.

3.3 Averting Behavior Indices

We considered the following WSH related behaviors in our analysis.

Water Safety Behavior: (1) whether DW storage has narrow mouth (2) whether DW storage is covered; (3) whether DW is used with a ladle or tap (indirect handling); and (4) whether DW is stored at least 3 ft height.

---

[4] In taking log of e. coli results reported in CFU/mL, we face a mathematical issue of taking a log of 0 CFU/ml. To deal with this issue, we multiply the CFU/ml count by 100 – that is, convert the units to CFU/100 ml – and then take a log of that number. For samples with 0 CFU/100ml count, we assign them a loge value of 0 (i.e., 1 CFU/100ml) which is almost negligible compared to the next higher concentration level of 100 CFU/100ml.

Hand washing behaviors: (1) whether PCG washes hands at least 4 of 5 critical times[5]; (2) whether U5 children' hands are washed at 2 critical times[4]; (3) whether children palms and nails are clean; and (4) whether soap or ash is observed at hand washing place.

Sanitation behaviors: (1) whether household uses toilets (no open defecation [OD]); (2) whether solid waste disposed off safely; and (3) whether waste water is disposed off safely.

We estimated models shown in Equations (4) and (5) with each individual averting behavior listed above. However, most behavioral coefficients were insignificant because of multicollinearity between individual behavior variables. To avoid this problem, use indices of behaviors in our models. We constructed water safety behavior index (*watsafe*), hand washing index (*hw*), and sanitation index (*san*) by summing the binary (0 1) values for the constituent behaviors listed above. We truncated the *watsafe* and *hw* indices values to 3 so that they are commensurate with *san* index. All three indices have values from 0 to 3; 0 means no averting behavior and 3 means best possible averting behaviors. To simplify the models (especially 2SLS estimation and mutatis muntandis caculations) and to model the combined effect of averting behaviors, we add *watsafe*, *hw* and *san* indices together to create a composite behavior index (*behindex)* which ranges from 0 to 9.[6] On average, *watsafe* and *hw* index is between 1 and 2 whereas san index is between 0 and 1 (Table I). Behindex average is between 3 and 4. We dont find any systematic difference across time and seasons (survey rounds).

---

[5] 5 critical hand washing times for adults are: before preparing food or cooking, before eating, before feeding children, after changing baby/handling child's feces, and after defecation. 2 critical times for U5 children are: before eating and after defecation.

[6] In addition to the mathematical indices *watsafe*, *hw*, *san*, and *behindex*, we created principal component (PC) based indices as well. We used the first PC as the index and estimate all models presented in this paper. The results are practically the same as those for the mathematical indices presented here. Therefore, we are not reporting results of analysis using PC based behavioral indices.

In addition to above averting behaviors, we also use a dummy variable to represent whether households filter DW or not. In Maharashtra, filtering is done using cloth or simple net filter which may in fact contaminate DW as found in our exploratory analysis[7]. As reported in Table I, approximately 60% households filter DW at home.

3.4 Other Variables

To model the choice of A, we use several explanatory variables listed in Equation (4). We find that majority of households perceive their source water to be very safe, household head are typically illiterate or with some primary education. Households own 2-3 farm assets and incur monthly expenditure of Rs. 4,500 on average (we have $\log_{10}$ transformed expenditure). Majority of households identify WSH related problems and diseases as most important to tackle. Approximately 40% household had good knowledge of causes of diarrhea. Households reported receiving 1-2 behavior change messages (IEC messages) related to WSH in 2005 and 2-3 IEC messages in 2007 (when Jalswarajya was implemented). 25-30% Households report diarrhea cases in 2 week recall period except Monsoon 2007 when 20% households reported diarrhea. Village level diarrhea prevalence is approximately 30% with lower prevalence in 2007. Majority of households report paying for water through fees or taxes and accessing improved water sources[8]. These percentages are higher in year 2007. Water availability also jumped in year 2007 to 40 LPCD from average 25 LPCD levels in 2005. Above benefits may be because Jalswarajya program was implemented in about 40% of the villages in 2007. Village level WSH safe practices remain low throughout. Over 80% households reported open defecation, only 10%

---

[7] We tested the difference in household e coli levels for households which practice or don't practice several averting behaviors using t-test. We could find only weak but positive effect of covered DW storage, storage of DW at height, soap or ash at hand washing station, no open defecation and proper waster water management on reduced level of e. coli at home. However, for filtering (mostly with cloth or net), we found strong and consistent negative effect.

[8] Improved sources are: public or private taps and bore wells, but not surface water or dug wells.

disposed solid waste safely and 30% disposed wastewater safely. In 2007, we see some improvement in open defecation practice (73% households). Over 50% of the villages reported proper drainage system. Less than 20% of the community survey respondents reported any knowledge of microbial contamination of source water in recent past (3 months prior to the survey). Majority (60%+) villages reported existence of a village water and sanitation committee (VWSC)[9]. In 2005, villages reported on average no or one government program or scheme (any type of assistance) to be active in the village whereas in 2007 we see slight improvement; may be because of Jalswarajya program. Number of cooperative groups, associations, clubs also increase from approximately 2 in 2005 to 5 in 2007.

To model Q, we use only exposure related explanatory variables besides A and R. These include the number of members in the household which determines the human capital available for water quality production as well as higher possibility of contamination at home. Improved type of water sources and amount of water will also determine Q. Community WSH practices and village diarrhea prevalence determine the pathogens load in the environment; and thus, can directly affect Q. In addition, we use dummies to evaluate the effects of seasons and time on Q.

4. *Results*

We model the choice of *watsafe*, *hw*, *san* and *behindex* (the composite index) using different methods such as ordered probit, probit, OLS and RE (8 models) as reported in Table II.[10] Overall, we specify 12 models for Q in Table III. We model Q as a function of *watsafe*, *hw*

---

[9] Government of India, as per its reform agenda recommend setting up of VWSC in villages to take ownership of water sources and sanitation situation in the village. This body is independent from the elected *Gram Panchayat*.

[10] We also model A and Q using nine individual averting behaviors that are used to create different behavior indices using OLS, probit and their RE models. We could not find consistent results because of multicollinearity between the behaviors. These and other alternative specifications are available from the authors upon request.

.

and *san* behavior indices together *or* using only *behindex*. We consider both the continuous and binary descriptions of Q. Thus, we use 3 alternate methods to model ln_ecoli (continuous Q): OLS, OLS with RE and 2SLS and we use probit, probit with RE and 2SLS when Q is binary (HHecoli). In Table IV, we report the *mutatis muntandis* effects using the results from Tables II for *behindex* (using ordered probit) and *filter* (using probit) and from Table III for ln_ecoli (using OLS) and HHecoli (probit). In these calculation, all statistically insignificant coefficients (at $\alpha >$ 0.1 level) are assumed to be zero. We compare these calculated *mutatis muntandis* effects with those estimated from reduced form models in Table IV as well.

In all tables, we report the coefficient and its SE below it, except in Table IV where we report only coefficients. The significance levels are indicated as: # for $\alpha <= 0.01$, & for $\alpha <= 0.05$, and * for $\alpha <= 0.1$, and all coefficients and their SE are bold faced. We report R^2 for all models except 2SLS models where R^2 values do not provide an accurate measure of models predictive power. For 2SLS models, we have confirmed validity of results using test of endogeneity and overidentification (results available from authors upon request).

4.1 Choice of Averting Behaviors (Table II)

The choice of A is positively associated with education of household head, caste and monthly expenditure. Other household socio-economic related variables such as house construction type, number of farm assets have mixed and weak results across different models. As number of IEC messages related to WSH increase the households' likelihood of adopting averting behaviors also increase. However, the likelihood of averting behaviors is lower, except for filtering, if households rate WSH related improvement as the most important one needed in their village. Other attitude or disease knowledge related variables are not consistent and significant.

Household level diarrhea prevalence is positively associated with hand washing index and negatively with water safety and sanitation indices. Households that pay for water are strongly and significantly more likely to engage in averting behaviors.   Effect of amount of water available to households is significant but weak. Households that use improved water sources filter less (strong and significant substitution effect). Similarly, households that perceive the DW sources to be very safe are less likely to filter their water.

We find a strong and significant effect of Jalswarajya program on filtering, but weak effect on other behaviors. As open defecation practice in the village increases, the households are less significantly likely to engage in averting behaviors. However, households own sanitation behaviors improve as the community waste management behaviors worsen. Number of government programs or schemes, and number of community groups, cooperatives and associations active in the village have weak effect on behaviors except filtering. Filtering is less in households that belong with villages with more number of programs or groups. The dummy for seasons and years indicate that compared to summer of 2005, water safety and sanitation behaviors worsen over time and in monsoon whereas hand washing behaviors improve. Averting behaviors are better in monsoon of 2005, but worse in monsoon of 2007. R has no effect on averting behaviors except filtering which is more in households with higher source water contamination. Since filtering itself contaminates DW (as explained next), total effect of source water contamination on Q is further exacerbated.

4.2 Household Water Quality Production (Table III)

The results for OLS or probit models and their RE models are practically the same. The results for 2SLS models show more number of significant explanatory variables and theie coefficients are also larger than those estimated by other methods. However, the direction of effect is consistent regardless of model specification. We mainly use results from OLS and probit regressions in our discussions.

Source water quality has consistent, strong and significant effect on household DW quality. If source water contamination doubles (100% increase) then the household water quality will worsen by 16% or the probability of e coli contamination of DW increases by approximately 8-9%. 2SLS results are similar.

Averting behaviors can lower DW contamination at home as expected. However, we find this effect only for water safety related behaviors. Each additional *watsafe* behavior lowers the probability of contaminated DW by 3-4%. Thus, if a household that did not use any *watsafe* behaviors can reduce likelihood of contaminated DW by 10-11% by engaging in at least 3 *watsafe* related behaviors. *behidex* has similar direction of effect, but the coefficient is smaller because we include insignificant *hw* and *san* indices in *behindex* along with *watsafe*. As per 2SLS results, *hw* can substantially and significantly lower contamination at home.

Filtering deteriorates the DW quality at home. Unless cloth and net filters are regularly cleaned and disinfected, they may in fact harbor and breed e. coli. This negative effect is significant and strong in all our models (including those which are not presented in this paper). In households that filter their DW, e. coli levels are 13-14% higher or they are 10% more likely to have e. coli in their DW compared to the households which don't filter DW.

Other exposure factors also are key in producing Q. Households using improved water source have 33% less contamination level or they are 20% less likely to have contaminated DW. Village level safe disposal of solid waste is associated with 40% lower contamination levels or 30% less likelihood of e. coli contamination. Village level diarrhea prevalence is also a key determinant of household water quality. For a10% increase in diarrhea prevalence in the village, household e. coli contamination increases by approximately 7% and the likelihood of e. coli in DW increases by 4-5%. Season also affects Q. Compared to summer, monsoon DW contamination can be 78-79% higher in 2005. In 2007, level of e. coli contamination does not substantially vary by season, but the likelihood of contamination in monsoon is somewhat lower (~20%) than that in summer (~30%). This may be the effect of improved water and sanitation infrastructure in 2007 compared to 2005 due to government programs such as Jalswarajya or other time variant factor[11].

4.3 Mutatis Muntandis Effect (Table IV)

Exogenous variables, such as R, affect Q directly ($\partial Q/\partial R$) as well as through change in behaviors ($\partial Q/\partial A * \partial A/\partial R$) as shown in Equation (6). True impact of an intervention is best captured by *mutatis muntandis* effect than *ceteris paribus* effect. We find that R and improved type of DW source have higher magnitude of total effects whereas percentage of village households that safely dispose of solid waste has lower total effect than the partial effects reported in Table III. Dummies for season and time are have higher total effects. More importantly, indirect effects of paying for water and participation in Jalswarajya program on

---

[11] We verified this by including JS dummy in production of water quality models. The magnitude of coefficients for round2, round3 and round4 dummy substantially lowers and JS dummy is strong and substantial. This indicate that along with Jalswarajya there could be other time and season variant factors that affect household water quality.

improving Q are weak, but significant which demonstrates that how individual household exposure is modified through public interventions.

5. Conclusions

Our main finding is that behaviors influence the in-house production of water quality. The effect of endogenous household averting behaviors on DW quality is significant, but weak compared to other exogenous 'policy' variables. We find that the effectiveness of hand washing and water safety behaviors is sensitive to the functional form. However, our overall behavior index is robust to model specification and suggests that good behaviors improve in-house water quality, with one exception. We find that traditional filtration actually increases e. coli contamination of in-house drinking water, presumably because cloth or net filters are not disinfected frequently. Given that over 60% households in our sample filter DW, this message has important public health implications.

While reductions in source water microbial contamination improve DW quality, these effects are greater for improved sources. Further, community sanitation (no open defecation and better waste management) directly improves in-house water quality and indirectly influences averting behaviors; thus, the total effect of these community behaviors is large. To our knowledge, this estimated relationship between community behaviors and household water quality represent one of the first credible estimates of environmental health externalities – whose size and significance is often a critical parameter in justifying public health interventions [17]. We find that village level diarrhea prevalence is strongly positively related to both household adoption choice and DW quality – underlining the fact that diarrhea is environmental and thus preventable. The first finding indicates that demand for averting behaviors is prevalence elastic

[17] and the second finding implies that public health mitigation (e.g., treatment and case management) can in turn improve household water quality and lower diarrhea.

Previously, we have made the case that using only ceteris paribus coefficients - of a complex model of fecal-oral contamination – can be significantly biased. Instead, we suggest that *mutatis muntandis* effects – which consider total effect of policy intervention on endogenous variables *considering all necessary changes* – are needed to evaluate policy impacts. In our analysis, several exposure (and thus policy) factors also affect Q *indirectly* by improving or worsening A. Sometimes households substitute (lower A) and sometimes they compliment (increase A) the gains from an exogenous improvement. For example, households that use improved source reduce *watsafe* and *filter*. Higher diarrhea levels result in better *hw*, but not *watsafe* and *san*. Better community waste management may improve *san*, but not *watsafe* and *hw*. On the other hand, households who pay for water better manage DW water at home as well as improved *san* and *hw*. Lower open defecation in the community results in better *watsafe*, *h*w and *san* at household level. In sum, any attempts to reduce diarrhea deaths and illness through water quality improvements, must package infrastructure delivery with education and information to induce behavior change.

6. References

1. Berman P., C. Kendall, and K. Bhattacharyya. The household production of health: Integrating social science perspectives on micro-level health determinants. *Social Science and Medicine*, 38(2):205-215 (1994).
2. Clasen T., W. Schmidt, T. Rabie, I. Roberts, and S. Cairncross. Interventions to improve water quality for preventing diarrhea: systematic review and meta-analysis. *British Medical Journal*, doi:10.1136/bmj.39118.489931.BE (2007).
3. Curtis V., and S. Cairncross. Effect of washing hands with soap on diarrhea risk in the community: a systematic review. *The Lancet of Infectious Diseases*, 3(5):275-281 (2003).
4. Dickie M. Defensive behavior and damage cost methods IN "A primer on nonmarket valuation" (P.A. Champ, K.J. Boyle, and T.C. Brwon, Eds), Kluwer Academic Publishers: Boston (2003).
5. Dickie M., and S. Gerking. Valuing morbidity: a household production approach. *Southern Economic Journal*, 57(3):690-702 (1991).
6. Esrey S.A., J.B. Potash, L. Roberts, and C. Shiff. Effects of improved water supply and sanitation on ascaris, diarrhea, dracunculiasis, hookworm infection, schistosomiasis, and trachoma. *Bulletin of World Health Organization*, 69(5):609-621 (1991).
7. Esrey S.A., J.P. Habicht, M.C. Latham, D.G. Sisler, and G. Casella. Drinking water source, diarrheal morbidity, and child growth in villages with both traditional and improved water supplies in rural Lesotho, Southern Africa. *American Journal of Public Health*, 78(11):1451-1455 (1988).
8. Fewtrell L., R. Kaufmann, D. Kay, W. Enanoria, L. Haller, and J. Colford. Water, sanitation, and hygiene interventions to reduce diarrhea in less developed countries: a systematic review and meta-analysis. *Lancet Infectious Diseases*, 5(1):42-52 (2005).

9. Ford G., and J. Jackson. On the interpretation of policy effects from estimates of simultaneous systems of equations. *Applied Economics*, 30(1):995-999 (1998).

10. Harrington W., and P. Portney. Valuing the benefits of health and safety regulations. *Journal of Urban Economics*, 22(1):101-112 (1987).

11. Hughes G., K. Lvovsky, and M. Dunleavy. Environment Health In India: Priorities in Andhra Pradesh. South Asia Environment and Social Development Unit, The World Bank (2001).

12. Huttly S., D. Blum, B. Kirkwood, R. Emeh, N. Okeke, M. Ajala, G. Smith, D. Caron, O. Dosunmu-Ogunbi, and R. Feachem. The Imo State (Nigeria) drinking water supply and sanitation project 2. impact on dracunculiasis, diarrhoea, and nutritional status. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 84(2): 316-321 (1990).

13. Jagals P., W. Grabow, and E. Williams. The effects of supplied water quality on human health in an urban development with limited basic subsistence facilities. *Water SA*, 23(4):373-378 (1997).

14. Jensen P., J. Ensink, G. Jayasinghe, W. van der Hoek, S. Cairncross, and A. Dalsgaard. Domestic transmission routes of pathogens: the problem of in-house contamination of drinking water during storage in developing countries. *Tropical Medicine and International Health*, 7(7):604-609 (2002).

15. Kravitz J., M. Nyaphisi, R. Mandel, and E. Petersen. Quantitative bacterial examination of domestic water supplies in the Lesotho highlands: water quality sanitation and village health. *Bulletin of the World Health Organization*, 77(10):829-836 (1999).

16. Nanan D., F. White, I. Azam, H. Afsar, and S. Hozhabri. Evaluation of a water, sanitation, and hygiene education intervention on diarrhea in northern Pakistan. *Bulletin of the World Health Organization*, 81(3):160-165 (2003).

17. Pattanayak S., and A. Pfaff. Behavior, environment and health in developing countries: evaluation and valuation. *Annual Review of Resource Economics*, 1: 183-222 doi:10.1146/annurev.resource.050708.144053 (2009).

18. Pattanayak S., C. Poulos, J-C Yang, and S. Patil. Forthcoming. How valuable are environmental health interventions? economic evaluation of a water and sanitation project in Maharashtra, India. *Bulletin of the World Health Organization*. In Publication MS BLT/2009/06605.

19. Pattanayak S., C. Poulos, J-C. Yang, S. Patil, and K. Wendland. Of taps and toilets: quasi-experimental protocols for evaluating community-demand driven projects. *Journal of Water and Health*, 7(3):434-451 (2009).

20. Prüss A., D. Kay, L. Fewtrell, and J. Bartram. Estimating the burden of deases from water, sanitation, and hygiene at a global level. *Environmental Health Perspectives*, 110(5):537-542 (2002).

21. Quick R., A. Kimura, A. Thevos, M. Tembo, I. Shamputa, L. Hutwagner, and E. Mintz. Diarrhea prevention through household-level water disinfection and safe storage in Zambia. *American Journal of Tropical Medicine and Hygiene*, 66(5):584-589 (2002).

22. Rosenzweig M., and T. Schultz. Estimating a household production function: heterogeneity, the demand for health inputs and their effect on birth weight. *Journal of Political Economy*, 91(5):723-746 (1983).

23. Sobsey, M. Managing Water in the Home: Accelerated Health Gains from Improved Water Supply. WHO/SDE/WSH/02.02, World Health Organization: Geneva (2002).

24. StataCorp. Stata Statistical Software: Release 9. StataCorp LP: College Station, TX (2005).

25. The World Health Organization (WHO). Guidelines for Drinking Water Quality, 3rd Ed. Volume 1. WHO: Geneva (2000).

26. The United Nations Children's Fund (UNICEF) and the World Health Organization (WHO). Diarrhoea: Why Children are Still Dying and What can be Done. Division of Communication (UNICEF): New York, and WHO Press: Geneva (2009).

27. VanDerslice J., and J. Briscoe. All coliform are not created equal: A comparison of the effects of water source and in-house water contamination on infantile diarrheal disease. *Water resources Research*, 29(7):1983-95 (1993).

28. Wang L. Health Outcomes in Low-income Countries and Policy Implications: Empirical Findings from Demographic and Health Surveys. Policy Research Working Paper #2831, The World Bank: Washington DC (2002).

29. Young B., and J. Briscoe. A case-control study of the effect of environmental sanitation on diarrhoea morbidity in Malawi. *Journal of Epidemiology and Community Health*, 42(1):83-88 (1988).

Table I. Description of Variables

Variables	**Description**	**Type**	**Summer '05**	**Monsoon '07**	**Summer '07**	**Monsoon '07**
Sample Size						
	All Households		10,194	9,307	9,329	9,303
	HH with data on Q		4,222	4,753	5,530	5,228
watsafe	water safety additive index score	Scale (0-3)	1.47 (0.76)	1.31 (0.72)	1.32 (0.66)	1.26 (0.7)
hw	hand washing additive index score	Scale (0-3)	1.51 (0.95)	1.88 (0.89)	1.86 (0.97)	1.75 (0.96)
san	Sanitation additive index score	Scale (0-3)	0.51 (0.69)	0.56 (0.75)	0.67 (0.83)	0.69 (0.8)
behindex	Combined WSH additive index	Scale (0-9)	3.49 (1.47)	3.76 (1.43)	3.85 (1.48)	3.71 (1.46)
filter	Does household filter DW?	Dummy	57%	66%	64%	65%
ln_ecoli	Log10 of HH e. coli contamination	Continuous	0.82 (1.68)	1.78 (2.47)	1.13 (1.9)	1.09 (2.01)
HHecoli	Is E. coli is present in hh DW?	Dummy	22%	40%	30%	26%
ln_villecoli	Log10 of Source e. coli contamination	Continuous	1.49 (2.03)	2.19 (2.72)	1.49 (2.08)	1.52 (2.13)
exsafety	Does HH perceive source water very safe?	Dummy	71%	55%	61%	63%
hhsize	Number of members in the HH	Continuous	6.43 (2.39)	6.62 (2.45)	7.06 (2.59)	7.15 (2.63)
headedu	Education level of HH head	Scale (1-6)				

	No Education		45%	41%	39%	37%
	Primary		19%	18%	22%	21%
	Secondary		27%	32%	32%	34%
	Higher Secondary		6%	6%	5%	5%
	Some College		2%	2%	2%	2%
	Graduate		1%	1%	1%	1%
scst	Does HH belong to (SC ST) backward caste?	Dummy	38%	37%	49%	48%
ln_totexpd		Continuous	8.17 (0.75)	8.13 (0.53)	8.4 (0.5)	8.34 (0.49)
num_farm		Continuous	2.1 (1.71)	0.06 (0.31)	2.44 (1.61)	0.07 (0.33)
htype	Type of house construction	Scale (0-2)				
	Kuccha		22%	21%	13%	13%
	Semi Pucca		40%	40%	35%	38%
	Pucca		38%	39%	52%	50%
wsh_need	Does HH id WSH as most important need?	Dummy	59%	62%	58%	52%
wsh_disease	Does HH id WSH related diseases as most important?	Dummy	52%	71%	69%	69%
wsh_pollu	Does HH id WSH related pollution as most important?	Dummy	43%	49%	46%	45%

idcause	Can HH list at least 5 of 8 causes of diarrhea?	Dummy	32%	34%	39%	48%
iec_msg	No of IEC messages received by HH out of 4	Continuous	1.76 (1.8)	1.8 (1.22)	2.23 (1.25)	2.61 (1.2)
HHdiarr2	Does anyone in HH has diarrhea in past 15 days except yesterday?	Dummy	26%	28%	24%	19%
villdiarr	% of HH in the village that have diarrhea in past 2 weeks	Continuous	32% (14%)	32% (12%)	27% (13%)	22% (12%)
paywater	Does HH pay for water?	Dummy	56%	70%	70%	68%
improvedw	Does HH use improved water source?	Dummy	66%	41%	72%	76%
lpcd2	Liters per capita per day water available to HH	Continuous	28.13 (25.43)	25.71 (18.69)	39.69 (26)	38.07 (19.72)
JS	Does village participate in Jalswarajya WSH program?	Dummy	0%	0%	36%	36%
pct_od	% of surveyed HH in the village that openly defecate	Continuous	0.86 (0.2)	0.84 (0.23)	0.73 (0.31)	0.73 (0.3)
pct_sw	% of surveyed HH in the village properly dispose Solid Waste	Continuous	0.09 (0.09)	0.11 (0.12)	0.1 (0.14)	0.11 (0.13)
pct_ww	% of surveyed HH in the village that properly dispose waste water	Continuous	0.28 (0.17)	0.3 (0.22)	0.3 (0.26)	0.31 (0.25)
drain	Does village has pucca organized drainage?	Dummy	53%	51%	61%	56%
villwqprob	Did community respondents report microbial	Dummy	16%	14%	18%	18%

	contamination of source in past?					
vwsc	Is water and sanitation committee present in the village?	Dummy	60%	47%	67%	71%
villpgm	No govt programs / assistances in the village	Continuous	0.56 (0.47)	0.57 (0.47)	1.73 (1.07)	0.88 (0.82)
villgrps	No of co-op groups, committee etc in village	Continuous	2.19 (1.84)	2.22 (1.86)	5.11 (1.66)	4.9 (1.62)

Table II. Estimation of behaviors using e. coli contamination for different estimation methods

	watsafe		***hw***		***san***		***behindex***		**filter**	
	O. Probit	**Reg RE**	**O. Probit**	**Reg RE**	**O. Probit**	**Reg RE**	**O. Probit**	**Reg RE**	**Probit**	**Probit RE**
ln_villecoli	-0.002	-0.001	0.006	0.006	0	-0.001	0.002	0.004	**0.021**	**0.019**
	0.005	0.003	0.006	0.006	0.003	0.001	0.005	0.006	**0.010****	**0.004*****
exsafety	0.025	0.015	**0.042**	**0.037**	**-0.033**	**-0.02**	0.025	0.032	**-0.049**	**-0.062**
	0.02	0.012	**0.020**$^{\#}$	**0.017**$^{\#}$	**0.017**$^{\#}$	**0.008**$^{\#}$	0.017	0.021	**0.023****	**0.016*****
hhsize	0.001	0	**0.022**	**0.021**	**0.011**	**0.007**	**0.02**	**0.029**	**-0.009**	**-0.01**
	0.003	0.002	**0.004**$^{\&}$	**0.003**$^{\&}$	**0.004**$^{\&}$	**0.002**$^{\&}$	**0.003**$^{\&}$	**0.004**$^{\&}$	**0.005****	**0.004****
headedu	**0.026**	**0.015**	**0.056**	**0.047**	**0.089**	**0.045**	**0.085**	**0.107**	**0.066**	**0.065**
	0.007$^{\&}$	**0.004**$^{\&}$	**0.007**$^{\&}$	**0.006**$^{\&}$	**0.007**$^{\&}$	**0.004**$^{\&}$	**0.007**$^{\&}$	**0.009**$^{\&}$	**0.010*****	**0.008*****
scst	**-0.035**	**-0.023**	**-0.047**	**-0.036**	**-0.098**	**-0.041**	**-0.083**	**-0.101**	**-0.115**	**-0.101**
	0.019*	**0.012***	**0.021**$^{\#}$	**0.018**$^{\#}$	**0.021**$^{\&}$	**0.009**$^{\&}$	**0.018**$^{\&}$	**0.024**$^{\&}$	**0.035*****	**0.019*****
ln_totexpd	0.014	0.004	**0.139**	**0.117**	**0.194**	**0.078**	**0.169**	**0.201**	**0.068**	**0.068**
	0.018	0.011	**0.018**$^{\&}$	**0.015**$^{\&}$	**0.021**$^{\&}$	**0.010**$^{\&}$	**0.018**$^{\&}$	**0.022**$^{\&}$	**0.023*****	**0.016*****
num_farm	**-0.01**	**-0.006**	0.008	0.005	**0.038**	**0.019**	**0.016**	**0.018**	**0.021**	**0.027**

	0.005*	**0.003$^{\#}$**	0.006	0.005	**0.005$^{\&}$**	**0.003$^{\&}$**	**0.005$^{\&}$**	**0.007$^{\&}$**	**0.008*****	**0.005*****
htype	0.017	0.012	0.01	0.009	**0.078**	**0.031**	**0.042**	**0.051**	-0.036	**-0.045**
	0.013	0.008	0.014	0.012	**0.013$^{\&}$**	**0.006$^{\&}$**	**0.013$^{\&}$**	**0.016$^{\&}$**	0.023	**0.013*****
wsh_need	**-0.077**	**-0.049**	0.007	0.003	**-0.099**	**-0.05**	**-0.075**	**-0.1**	**0.096**	**0.098**
	0.019$^{\&}$	**0.012$^{\&}$**	0.02	0.017	**0.016$^{\&}$**	**0.008$^{\&}$**	**0.018$^{\&}$**	**0.023$^{\&}$**	**0.027*****	**0.016*****
wsh_diseas	0.011	0.008	**0.036**	0.026	**-0.032**	-0.006	0.026	0.023	-0.024	-0.017
	0.016	0.01	**0.019***	0.016	**0.018***	0.008	0.016	0.021	0.025	0.017
wsh_pollu	-0.004	0.002	0.002	0.002	0.013	-0.001	0.004	0.005	0.009	-0.005
	0.017	0.011	0.018	0.015	0.016	0.008	0.016	0.02	0.022	0.016
idcause	**-0.058**	**-0.036**	**-0.033**	-0.025	**0.047**	**0.028**	**-0.03**	-0.032	0.008	0.026
	0.022$^{\&}$	**0.014$^{\#}$**	**0.018***	0.015	**0.017$^{\&}$**	**0.009$^{\&}$**	**0.018***	0.023	0.024	0.017
iec_msg	0.002	0	**0.033**	**0.03**	**0.022**	**0.006**	**0.03**	**0.037**	**0.034**	**0.045**
	0.007	0.005	**0.008$^{\&}$**	**0.007$^{\&}$**	**0.006$^{\&}$**	**0.003$^{\#}$**	**0.007$^{\&}$**	**0.008$^{\&}$**	**0.009*****	**0.006*****
HHdiarr2	**-0.046**	**-0.029**	**0.089**	**0.072**	**-0.039**	**-0.015**	0.02	0.026	0.018	0.027
	0.014$^{\&}$	**0.009$^{\&}$**	**0.013$^{\&}$**	**0.011$^{\&}$**	**0.017$^{\#}$**	**0.008***	0.013	0.017	0.016	0.019
villdiarr	-0.066	-0.043	**0.23**	**0.219**	0.005	-0.001	0.123	0.188	0.224	**0.353**

	0.102	0.066	**0.132***	**0.113***	0.053	0.018	0.106	0.14	0.184	**0.068*****
paywater	**0.056**	**0.034**	**0.107**	**0.086**	**0.083**	**0.041**	**0.139**	**0.158**	**0.173**	**0.143**
	0.019$^{\&}$	**0.012**$^{\&}$	**0.025**$^{\&}$	**0.022**$^{\&}$	**0.019**$^{\&}$	**0.009**$^{\&}$	**0.022**$^{\&}$	**0.027**$^{\&}$	**0.034*****	**0.018*****
improved	**-0.052**	**-0.031**	**0.049**	0.038	**0.087**	**0.032**	0.032	0.038	**-0.183**	**-0.175**
	0.023$^{\#}$	**0.015**$^{\#}$	**0.029***	0.024	**0.021**$^{\&}$	**0.008**$^{\&}$	0.024	0.03	**0.041*****	**0.019*****
lpcd2	**-0.001**	**-0.001**	**0.002**	**0.001**	**0.001**	**0**	**0.001**	**0.001**	0	**0.001**
	0.000$^{\&}$	**0.000**$^{\&}$	**0.000**$^{\&}$	**0.000**$^{\&}$	**0.000**$^{\&}$	**0.000**$^{\#}$	**0.000**$^{\#}$	**0.000**$^{\#}$	0.001	**0.000***
JS	0.065	0.038	-0.007	-0.005	**0.097**	**0.017**	0.042	0.051	**0.156**	**0.165**
	0.047	0.029	0.049	0.042	**0.021**$^{\&}$	**0.006**$^{\&}$	0.043	0.056	**0.081***	**0.026*****
pct_od	**-0.235**	**-0.131**	**-0.393**	**-0.325**	**-1.541**	**-0.965**	**-1.092**	**-1.415**	-0.134	**-0.12**
	0.071$^{\&}$	**0.045**$^{\&}$	**0.080**$^{\&}$	**0.068**$^{\&}$	**0.038**$^{\&}$	**0.014**$^{\&}$	**0.063**$^{\&}$	**0.081**$^{\&}$	0.145	**0.043*****
pct_sw	-0.098	-0.024	**-0.459**	**-0.37**	**1.712**	**0.909**	**0.374**	**0.517**	**-0.345**	**-0.336**
	0.131	0.076	**0.125**$^{\&}$	**0.107**$^{\&}$	**0.048**$^{\&}$	**0.017**$^{\&}$	**0.089**$^{\&}$	**0.113**$^{\&}$	**0.200***	**0.067*****
pct_ww	**-0.159**	**-0.096**	**-0.431**	**-0.372**	**1.401**	**0.769**	**0.217**	**0.303**	0.177	**0.15**
	0.072$^{\#}$	**0.047**$^{\#}$	**0.099**$^{\&}$	**0.085**$^{\&}$	**0.039**$^{\&}$	**0.014**$^{\&}$	**0.077**$^{\&}$	**0.100**$^{\&}$	0.134	**0.051*****
drain	-0.038	-0.017	0.022	0.018	**0.044**	0.003	0.001	0.008	0.05	**0.05**

	0.037	0.024	0.036	0.032	**0.014**$^{\&}$	0.004	0.03	0.039	0.054	**0.019*** **
villwqprob	0.01	0.007	-0.043	-0.039	0.004	-0.003	-0.025	-0.034	**0.159**	**0.153**
	0.036	0.023	0.043	0.037	0.019	0.006	0.034	0.043	**0.055*** **	**0.022*** **
vwsc	0.005	0.004	0.037	0.027	**0.024**	0.004	0.028	0.034	0.019	0.015
	0.027	0.017	0.037	0.032	**0.013* **	0.004	0.027	0.034	0.045	0.017
villpgm	-0.03	-0.013	**0.057**	**0.05**	**-0.018**	**-0.011**	0.019	0.023	**-0.105**	**-0.108**
	0.028	0.017	**0.024**$^{\#}$	**0.020**$^{\#}$	**0.009* **	**0.002**$^{\&}$	0.022	0.028	**0.031*** **	**0.011*** **
villgrps	0.004	0.001	0.017	**0.018**	**-0.009**	-0.001	0.01	0.017	**-0.154**	**-0.155**
	0.01	0.007	0.012	**0.010* **	**0.005* **	0.002	0.009	0.012	**0.022*** **	**0.007*** **
round2	**-0.257**	**-0.165**	**0.411**	**0.364**	0.01	0.009	**0.164**	**0.209**	**0.171**	**0.221**
	0.040$^{\&}$	**0.026**$^{\&}$	**0.050**$^{\&}$	**0.042**$^{\&}$	0.017	0.006	**0.037**$^{\&}$	**0.047**$^{\&}$	**0.039*** **	**0.022*** **
round3	**-0.226**	**-0.151**	**0.148**	**0.122**	**-0.135**	**-0.029**	-0.04	-0.055	**0.677**	**0.713**
	0.044$^{\&}$	**0.028**$^{\&}$	**0.069**$^{\#}$	**0.059**$^{\#}$	**0.020**$^{\&}$	**0.006**$^{\&}$	0.051	0.065	**0.087*** **	**0.031*** **
round4	**-0.355**	**-0.218**	0.1	0.082	**-0.146**	**-0.042**	**-0.133**	**-0.177**	**0.568**	**0.605**
	0.058$^{\&}$	**0.037**$^{\&}$	0.065	0.056	**0.022**$^{\&}$	**0.008**$^{\&}$	**0.055**$^{\#}$	**0.071**$^{\#}$	**0.081*** **	**0.029*** **
Constant		**1.633**		**0.354**		**0.187**		**2.148**	-0.145	-0.154

		0.120$^{\&}$		**0.157**$^{\#}$		**0.085**$^{\#}$		**0.210**$^{\&}$	0.259	0.135
Obs	37727	37727	37727	37727	37727	37727	37727	37727	37619	37619
R^2	0.01	0.02	0.03	0.07	0.21	0.38	0.05	0.18	0.05	NA$^{\wedge}$

^ Log-likelihood ratio test is significant at $\alpha = 0.0001$.

Table III. Estimation of household exposure production technology using e. coli contamination and different behaviors and methods.

	ln_ecoli			**ln_ecoli**			**HHEcoli**			**HHEcoli**		
	OLS	REM	**2SLS**	**OLS**	**REM**	**2SLS**	**Probit**	**Probit REM**	**2SLS (Probit)**	**Probit**	**Probit REM**	**2SLS (Probit)**
ln_villecoli	**0.163**	**0.161**	**0.153**	**0.163**	**0.161**	**0.151**	**0.084**	**0.086**	**0.078**	**0.085**	**0.086**	**0.077**
	0.018$^{\&}$	**0.018**$^{\&}$	**0.018**$^{\&}$	**0.018**$^{\&}$	**0.018**$^{\&}$	**0.018**$^{\&}$	**0.010**$^{\&}$	**0.004**$^{\&}$	**0.005**$^{\&}$	**0.010**$^{\&}$	**0.004**$^{\&}$	**0.005**$^{\&}$
watsafe	**-0.051**	**-0.049**	-0.369				**-0.033**	**-0.034**	-0.294			
	0.023$^{\#}$	**0.023**$^{\#}$	0.518				**0.016**$^{\#}$	**0.014**$^{\#}$	0.24			
hw	-0.025	-0.023	**-0.681**				-0.012	-0.01	**-0.424**			
	0.022	0.022	**0.290**$^{\#}$				0.014	0.011	**0.117**$^{\&}$			
san	-0.013	-0.013	-0.343				-0.007	-0.007	-0.176			
	0.025	0.025	0.367				0.016	0.017	0.168			
behindex				**-0.03**	**-0.028**	**-0.524**				**-0.017**	**-0.016**	**-0.319**
				0.014$^{\#}$	**0.014**$^{\#}$	**0.118**$^{\&}$				**0.009***	**0.008**$^{\#}$	**0.047**$^{\&}$
filter	**0.137**	**0.133**	**1.889**	**0.139**	**0.136**	**1.949**	**0.093**	**0.093**	**1.155**	**0.094**	**0.095**	**1.202**
	0.042$^{\&}$	**0.041**$^{\&}$	**0.392**$^{\&}$	**0.042**$^{\&}$	**0.041**$^{\&}$	**0.382**$^{\&}$	**0.029**$^{\&}$	**0.021**$^{\&}$	**0.128**$^{\&}$	**0.029**$^{\&}$	**0.021**$^{\&}$	**0.116**$^{\&}$

hhsize	**-0.018**	**-0.018**	0.012	**-0.017**	**-0.017**	0.009	**-0.009**	**-0.009**	0.009	**-0.009**	**-0.009**	0.007
	0.007$^{\&}$	**0.007$^{\&}$**	0.011	**0.007$^{\#}$**	**0.007$^{\#}$**	0.009	**0.005$^{\#}$**	**0.004$^{\#}$**	0.006	**0.005***	**0.004$^{\#}$**	0.005
improved	**-0.339**	**-0.324**	-0.123	**-0.338**	**-0.322**	**-0.123**	**-0.2**	**-0.196**	**-0.071**	**-0.199**	**-0.195**	**-0.069**
	0.073$^{\&}$	**0.071$^{\&}$**	0.077	**0.073$^{\&}$**	**0.070$^{\&}$**	**0.074***	**0.044$^{\&}$**	**0.023$^{\&}$**	**0.029$^{\#}$**	**0.044$^{\&}$**	**0.023$^{\&}$**	**0.027$^{\#}$**
lpcd2	0	0	0.001	0	0	0.001	0	0	0.001	0	0	0
	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0	0.001	0.001	0	0.001
pct_od	0.261	0.259	-0.472	0.246	0.245	**-0.611**	**0.217**	**0.223**	-0.209	**0.209**	**0.214**	**-0.31**
	0.165	0.162	0.369	0.168	0.165	**0.271$^{\#}$**	**0.109$^{\#}$**	**0.049$^{\&}$**	0.14	**0.110***	**0.047$^{\&}$**	**0.089$^{\&}$**
pct_sw	**0.395**	**0.399**	0.588	**0.409**	**0.412**	**0.829**	**0.283**	**0.293**	**0.37**	**0.289**	**0.298**	**0.553**
	0.239*	**0.236***	0.521	**0.236***	**0.233***	**0.274$^{\&}$**	**0.163***	**0.086$^{\&}$**	**0.221***	**0.162***	**0.084$^{\&}$**	**0.095$^{\&}$**
pct_ww	-0.02	-0.029	0.04	-0.004	-0.015	0.225	-0.076	-0.085	-0.088	-0.067	-0.077	0.062
	0.194	0.191	0.395	0.193	0.19	0.196	0.126	0.06	0.185	0.125	0.058	0.067
drain	-0.101	-0.103	-0.022	-0.1	-0.102	-0.031	-0.062	**-0.066**	-0.015	-0.06	**-0.065**	-0.019
	0.072	0.072	0.091	0.072	0.072	0.089	0.047	**0.023$^{\&}$**	0.026	0.047	**0.023$^{\&}$**	0.025
villdiarr	**0.739**	**0.736**	**0.662**	**0.742**	**0.739**	**0.596**	**0.44**	**0.446**	**0.378**	**0.443**	**0.449**	**0.342**
	0.276$^{\&}$	**0.272$^{\&}$**	**0.328$^{\#}$**	**0.277$^{\&}$**	**0.272$^{\&}$**	**0.302$^{\#}$**	**0.179$^{\#}$**	**0.080$^{\&}$**	**0.097$^{\&}$**	**0.180$^{\#}$**	**0.080$^{\&}$**	**0.087$^{\&}$**

round2	**0.782**	**0.785**	**0.87**	**0.787**	**0.79**	**0.786**	**0.412**	**0.429**	**0.46**	**0.416**	**0.434**	**0.417**
	0.087$^{\&}$	**0.087**$^{\&}$	**0.169**$^{\&}$	**0.089**$^{\&}$	**0.088**$^{\&}$	**0.095**$^{\&}$	**0.058**$^{\&}$	**0.031**$^{\&}$	**0.067**$^{\&}$	**0.059**$^{\&}$	**0.030**$^{\&}$	**0.035**$^{\&}$
round3	**0.424**	**0.419**	**0.363**	**0.427**	**0.423**	**0.293**	**0.307**	**0.316**	**0.266**	**0.31**	**0.319**	**0.23**
	0.073$^{\&}$	**0.073**$^{\&}$	**0.146**$^{\#}$	**0.075**$^{\&}$	**0.074**$^{\&}$	**0.090**$^{\&}$	**0.054**$^{\&}$	**0.031**$^{\&}$	**0.060**$^{\&}$	**0.055**$^{\&}$	**0.031**$^{\&}$	**0.035**$^{\&}$
round4	**0.427**	**0.422**	**0.287**	**0.431**	**0.426**	**0.219**	**0.201**	**0.204**	0.105	**0.205**	**0.208**	**0.074**
	0.083$^{\&}$	**0.083**$^{\&}$	**0.151***	**0.085**$^{\&}$	**0.084**$^{\&}$	**0.097**$^{\#}$	**0.057**$^{\&}$	**0.032**$^{\&}$	0.064	**0.058**$^{\&}$	**0.032**$^{\&}$	**0.037**$^{\#}$
Constant	**0.527**	**0.522**	1.39	**0.51**	**0.505**	**1.533**	**-0.991**	**-1.027**	-0.354	**-1.005**	**-1.041**	**-0.378**
	0.227$^{\#}$	**0.223**$^{\#}$	0.864	**0.230**$^{\#}$	**0.226**$^{\#}$	**0.462**$^{\&}$	**0.151**$^{\&}$	**0.079**$^{\&}$	0.405	**0.153**$^{\&}$	**0.078**$^{\&}$	**0.198***
Obs	19460	19460	19460	19460	19460	19460	19460	19460	19460	19460	19460	19460
R^{2}	0.08	0.08	NA	0.08	0.8	NA	0.05	NA^	NA	0.05	NA^	NA

^ Log-likelihood ratio test is significant at $\alpha = 0.0001$

Table IV. Comparing partial and total effects– results

	beh index	**filter**	**ln_ecoli**			**HHecoli**		
			Partial	**Total**		**Partial**	**Total**	
	O. Probit	**Probit**	**OLS**	**Calculated**	**Reduced**	**Probit**	**Calculated**	**Reduced**
ln_villecoli		**0.021**	**0.163**	**0.166**	**0.159**$^{\&}$	**0.085**	**0.087**	**0.082**$^{\&}$
behindex			**-0.03**	**-0.030**		**-0.017**	**-0.017**	
filter			**0.139**	**0.139**		**0.094**	**0.094**	
hhsize	**0.02**	**-0.009**	**-0.017**	**-0.019**	**-0.015**$^{\#}$	**-0.009**	**-0.010**	-0.008
improvedw		**-0.183**	**-0.338**	**-0.363**	**-0.242**$^{\&}$	**-0.199**	**-0.216**	**-0.143**$^{\&}$
lpcd2	**0.001**			0.000	0.001		0.000	0.001
pct_od	**-1.092**			0.033	0.093	**0.209**	0.228	0.106
pct_sw	**0.374**	**-0.345**	**0.409**	**0.350**	0.312	**0.289**	**0.250**	**0.263***
pct_ww	**0.217**			-0.007	0.204		-0.004	0.046
drain				0.000	0.018		0.000	0.013
villdiarr			**0.742**	**0.742**	**0.65**$^{\#}$	**0.443**	**0.443**	**0.407***
round2	**0.164**	**0.171**	**0.787**	**0.806**	**0.825**$^{\&}$	**0.416**	**0.429**	**0.433**
round3		**0.677**	**0.427**	**0.521**	**0.695**$^{\&}$	**0.31**	**0.374**	**0.48**$^{\&}$
round4	**-0.133**	**0.568**	**0.431**	**0.514**	**0.65**$^{\&}$	**0.205**	**0.261**	**0.351**$^{\&}$
exsafety		**-0.049**		-0.007	-0.038		-0.005	-0.022
headedu	**0.085**	**0.066**		**0.007**	**-0.028***		**0.005**	**-0.017***
scst	**-0.083**	**-0.115**		-0.013	0.017		-0.009	0.009
ln_totexpd	**0.169**	**0.068**		0.004	0.036		0.004	0.025

num_farm	**0.016**	**0.021**		0.002	0.004		0.002	0.004
htype	**0.042**			-0.001	-0.025		-0.001	0
wsh_need	**-0.075**	**0.096**		0.016	0.034		**0.010**	**0.043***
wsh_disease				0.000	-0.038		0.000	-0.015
wsh_pollu				0.000	0.035		0.000	0.031
idcause	**-0.03**			0.001	-0.007		0.001	-0.003
iec_msg	**0.03**	**0.034**		0.004	-0.019		**0.003**	**-0.016***
HHdiarr2				**0.000**	**0.069***		0.000	0.033
paywater	**0.139**	**0.173**		**0.020**	**-0.112***		**0.014**	**-0.065**$^{\#}$
JS		**0.156**		**0.022**	**0.279**$^{\&}$		**0.015**	**0.186**$^{\&}$
villwqprob		**0.159**		**0.022**	**0.196**$^{\#}$		0.015	0.088
vwsc				0.000	-0.014		0.000	-0.03
villpgm		**-0.105**		-0.015	-0.047		-0.010	-0.022
villgrps		**-0.154**		**-0.021**	**-0.124**$^{\&}$		**-0.014**	**-0.083**$^{\&}$
Constant			**0.51**	0.510	0.585	**-1.005**	**-1.005**	**-0.974**$^{\&}$
R^{2}	0.05	0.05	0.08	NA	0.09	0.05	NA	0.06