Improved Source, Improved Quality? Demand for Drinking Water Quality in Rural India

Katrina Jessoe*

Abstract

This paper tests the hypothesis that the expansion of improved drinking water supplies in rural India reduced household expenditure on water quality, offsetting some of the quality benefits from source protection. I estimate demand for in-home treatment using geological characteristics to predict a household's drinking water source. The probability of treatment and in particular boiling reduces by 18 to 27 percentage points in response to source protection, offsetting 4% of the water quality gains and saving households 0.5 to 1% in monthly expenditure. Behavioral choices partly counteract the water quality gains from source protection.

JEL: I18, O13, Q51, Q53, Q56.

Keywords: Drinking water quality; averting expenditure; compensating behavior

*Assistant Professor, University of California, Davis; email: kkjessoe@ucdavis.edu. I would like to thank Sheila Olmstead, Michael Boozer, Robert Mendelsohn, Pierre Merel and Erin Mansur for their time and comments as well as Achyuta Adhvaryu, Reena Badiani, Colin Carter, Rahul Deb, Joshua Lustig, Bob Repetto, Maher Said and Steven Wallander. Kapil Narula and T.S. Panwar offered invaluable suggestions regarding the collection and analysis of field data. Vidur Chopra, Priyanaka Chatterjee and Amrita Singh provided excellent assistance in field data collection. I thank Mr. Swaroop for assistance in accessing and interpreting the groundwater data. This has been funded in part by the United States Environmental Protection Agency (EPA) under the Science to Achieve Results (STAR) Graduate Fellowship Program (FP 91690401), the Heinz Fellowship for Environmental Scholars and the V. Kann Rasmussen Foundation. All remaining errors are my own.

1 Introduction

Diarrheal disease, mostly attributable to unsafe water supplies, the absence of sanitation infrastructure, and a lack of hygiene, is a leading cause of child morbidity and mortality in developing countries (WHO 2004). Development aid has prioritized investment in public health infrastructure, especially drinking water, as a means to counteract this problem (Black and Talbot 2005, Butterworth and Soussan 2001). Economists have demonstrated the links between government policies to improve health and safety, and compensating behavior that offsets their intended gains in many contexts including seat belt laws (Peltzman 1975), CAFE standards (Small and Van Dender 2005), and sexual education programs (Oettinger 1999). This paper investigates whether public investments in safe drinking water sources induce a similar behavioral reaction and quantifies the net impact of this behavioral response on expenditure and water quality.

The intensity and scope of drinking water supply programs in rural India make it an ideal setting to measure the effect of source protection on averting expenditure. In India, a government report cites that on average every child under the age of 5 experiences 2 to 3 episodes of diarrhea yearly (Planning Commission 2002). In part to combat this, the government invested heavily in the expansion of improved drinking water sources - defined as taps, tube wells and hand pumps - which enable households to access groundwater protected from surface contamination, thereby lowering exposure to waterborne pathogens. Between 1989 and 1998 the percentage of rural households with access to improved sources increased from 55 to 70 percent (NSSO 1999).

While improved drinking water sources may significantly reduce the occurrence of waterborne pathogens at the source, there is mixed evidence for whether source protection causes measurable reductions in waterborne disease.[1] Households engage in a complex set of behaviors having to do with drinking water collection, storage, treatment, and withdrawal for drinking, all of which may influence disease incidence. To better understand the mechanisms underlying the relationship

[1]Meta-analyses assessing the impacts of water, sanitation, and hygiene improvements on diarrhea in developing countries suggest that safe drinking water supplies do reduce the incidence of diarrhea (Esrey and Habicht 1986, Esrey et al. 1991, Fewtrell et al. 2005). Similarly, in the U.S., the proliferation of piped drinking water supplies in urban areas in the late 1800s and early 1900s caused a rapid reduction in mortality rates (Cutler and Miller 2005). By contrast, other studies find that health benefits from access to clean drinking water only occur when bundled with sanitation, hygiene, or education programs (Brick et al. 2004, Checkley et al. 2004).

between source protection and childhood diarrhea, economists have begun to explore the role of information, education and income, as well as behavioral responses to source protection (Bennett 2012, Kremer et al. 2011b, Jalan and Ravallion 2003).[2] A randomized controlled trial in Kenya found that spring source protection increased drinking water quality by 62 percent and led to a 25 percent reduction in childhood diarrhea (Kremer et al. 2011b). However, other studies find limited or no health benefits from source protection. In India, piped water supplies appear to reduce childhood diarrhea only in high-income households (Jalan and Ravallion 2003). Of particular relevance is work in the urban Philippines that suggests substitution between piped drinking water and sanitation offsets the health gains from the expansion of piped water (Bennett 2012). The links between source protection, drinking water quality and childhood diarrhea remain unclear.

Using cross-sectional survey data from 1998, I investigate whether trade-offs between improvements in source water quality and in-home treatment "crowd-out" some of the quality benefits from source protection. The ideal study would exogenously introduce improved sources and collect household data both before and after their construction. And while one study has done this (Kremer et al. 2011b), most rely on observational data confronting empirical challenges in the placement of improved sources and the availability of household panel data. In India, taps and tube wells were intentionally placed in villages with few health services or little public infrastructure. Further households choosing to use improved sources likely differ from those who do not in education, health and income. In this study, spatial data on rock type are used to predict a household's drinking water source. These data are fixed over time and vary at the tehsil (U.S. equivalent of a civil township). A further difficulty when using observational data in India (e.g. NFHS, NSS) is obtaining locational identifiers. This challenge is magnified when using geographical characteristics as instruments. The 1998 NSS household data, the primary data set used in the analysis, can be spatially identified at the sub-district, thereby allowing me to control for district (the U.S. equivalent of a county) unobservables.[3]

[2] Related work evaluates the impact of information about arsenic levels, another contaminant found in groundwater, on a household's choice of water source (Bennear et al. 2013, Madajewicz et al. 2007).

[3] Ideally, the study would also control for uobservable cross-sectional heterogeneity. While the NFHS and NSS collected information on in-home treatment in more recent surveys, this information was not collected in the NFHS

Results support the hypothesis that demand for in-home treatment declines with an improvement in source water quality. Controlling for the endogenous choice of drinking water source and district unobservables, households with an improved source are 25 to 27 percentage points less likely to engage in in-home treatment. These results are robust to the inclusion of household, village and sub-district observables that might be systematically correlated with both in-home treatment and hydrogeological characteristics. I also consider demand for each mode of in-home treatment. Treatment technologies vary in the water quality they produce, the purchase cost, and the time cost to clean drinking water. Source protection appears to reduce demand for boiling, the most effective yet most time-intensive treatment technology by 18 percent.

Data collected during field work in India measure source water quality and pollution abatement supplied by in-home treatment technologies, providing a starting point to estimate the quality gains from source protection. This analysis relies on several simplifying assumptions and should be viewed as a back of the envelope calculation. Results suggest that changes in treatment offset the quality gains from improved sources by 3.6 percent overall. However, "crowding out" should only impact those households that were predicted to treat unimproved sources. For these households, which comprise 38 percent of the sample, changes in treatment offset 27 percent of the quality gains.

In addition to the potential health benefits from source protection, the reduction in expenditure on avoidance behavior also provides a benefit to households.[4] On average, I find that the per capita gain from source protection amounts to 0.5 to 1 percent of total expenditure. Of course, other avoidance behaviors may also be responsive to source protection.

Along multiple dimensions this paper contributes to recent discussions about the provision of clean drinking water, and more generally environmental infrastructure, in developing countries. Similar in spirit to Bennett (2012), this work shows that behavioral choices may compromise some of the water quality gains from source protection. In doing so, it adds another critical data point

1992-1993 survey or in earlier rounds conducted by the NSS. If it were the case that treatment behavior was surveyed in both rounds, then I could construct district panel data and estimate the effect of district source protection on district demand for treatment, controlling for district fixed effects. However, this specification would not control for the endogeneity of source protection.

[4] Other studies utilizing averting expenditure to measure willingness to pay for drinking water quality in developing countries include Dasgupta (2004), McConnell and Rosado (2000) and Pattanayak et al. (2005).

to questions about the impact of source protection, analyzing this question in a rural as opposed to an urban setting. Second, many studies assume that source protection provides water quality gains and that offsetting behavior compromises water quality. This study is among the first in south Asia to quantify the relationship between source protection, behavioral choices and drinking water quality. Translating source protection and behavioral choices into drinking water quality is essential to understand the relationship between source protection and human health. Lastly, this study provides a partial measure of willingness to pay for improvements in source water quality in rural India, extending earlier work by controlling for the endogeneity of source protection.

2 Context and related literature

Between 1981 and 2001, the number of rural households drinking from improved water sources nearly tripled from 27 to 73 percent, where this increase was largely fueled by government investment in source protection (Black and Talbot 2005). The GOI defines safe or improved sources to include taps, tube wells and hand pumps (Planning Commission 2002).[5] Improved sources enable households to access groundwater that is protected from surface contamination, thus lowering exposure to fecal coliform and other water-borne pathogens. By contrast, unimproved sources such as surface water or open dug wells are exposed to the surface and susceptible to pathogen contamination from free-flowing sewage (Black and Talbot 2005). The expansion of groundwater sources thus shifted drinking water supplies from surface water to protected groundwater, and in general provided households with a safer supply in terms of coliform bacteria.[6]

Still, improved sources may contain high levels of microbes or become contaminated during the transport and storage of the water supply.[7] In India, despite the existence of a microbial

[5]According to the World Health Organization (WHO) an improved source must adhere to the following criteria: (i) a significant increase in the probability that the water is safe; (ii) a more accessible source; and (iii) sufficient measures taken to protect the water source from contamination. In contrast to India, WHO extends the scope of improved sources to include protected dug wells, protected springs and rainwater collection (WHO 2000).

[6]Coliform bacteria describe a broad class of bacteria that are common in the environment. Though coliform bacteria are generally not harmful, the presence of coliforms serves as an indicator for potentially harmful pathogens and bacteria. *E. coli* and fecal coliform are bacteria that may originate from human or animal wastes. Microbes found in *E. coli* can cause short-term health effects, such as diarrhea, cramps, nausea, headaches, or other symptoms.

[7]Some ground water sources expose households to the separate health risk of unsafe concentrations of arsenic (Chaterjee et al. 1995, Chowdhury et al. 2000), though concentrations are typically higher and more of a concern in shallow wells.

standard for drinking water, both taps and tube wells as well as unimproved sources have been shown to contain high levels of coliform and other bacteria (Dasgupta 2004, Islam et al. 2007, Planning Commission 2002).[8] Further, significant contamination can occur during the transport of the water supply from the source to the household (Wright et al. 2004). Storage duration and storage material also influence the cleanliness of drinking water (Brick et al. 2004, Checkley et al. 2004).

One promising approach to improve drinking water quality is through increased uptake of in-home treatment. Studies suggest that treatment, as compared to source or storage improvements, is a more effective method to provide clean drinking water (Brick et al. 2004, Fewtrell et al. 2005). As such, recent empirical work has focused on understanding the drivers behind in-home treatment or the lack thereof (Ashraf et al. 2010, Hamoudi et al. 2011, Kremer et al. 2011a, Luoto et al. 2009).[9] Results from a randomized controlled trial point to price as a significant deterrent, finding that the adoption of chlorination significantly increases when households receive it free of charge (Kremer et al. 2011a). This work also highlights that salience, convenience, promotion and public provision can increase the adoption of chlorination. Other experimental work documents the importance of information (Jalan and Somanathan 2008) and the sharing of this information (Luoto et al. 2011). Of note is a study in urban India that informed households if their drinking water was clean or dirty; upon learning that it was contaminated, in-home treatment increased by 11 percentage points (Jalan and Somanathan 2008). Other work in rural India shows that information about drinking water quality induced households to purchase water from safer sources, but did not change their use of time intensive in-home treatment (Hamoudi et al. 2012).

In addition to this experimental work, studies relying on observational data have shown that price and information, as well household composition, employment, education and wealth influence averting behavior (Jalan et al. 2009, Mintz et al. 2001, Pattanayak et al. 2005, Quick et al. 1999). Some of these studies also evaluated the impact of source protection or perceived water quality

[8] A source is considered safe for human consumption if there are less than 10 coliform counts per 100 ml of drinking water (Dept. Drinking Water Supply 2007).

[9] See Ahuja et al. 2010 for an overview of randomized controlled trials. Somanathan (2010) provides a review of empirical studies assessing the impact of information on in-home treatment (among other things) and summarizes work in India evaluating the role of education, occupation, news media and information on water quality.

on coping behavior (Alberini et al. 1996, Dasgupta 2004, Larson and Gnedenko 1999, Um et al. 2002). This paper builds on this literature in two ways. First, it isolates the causal effect of source protection on averting behavior using an instrumental variables approach. Second, it translates changes in source protection and averting behavior into changes in drinking water quality as measured via coliform counts.

3 Conceptual framework

To explore the effect of improvements in source water quality on demand for treatment, I present a simple framework in which households produce health from expenditure on water quality, and derive utility from health as well as a composite good.[10] Consider a model in which household utility $U(H(E,T),Z)$ is derived from the consumption of health, H, and a composite good, Z. Utility is assumed to be quasiconcave and increasing in Z and H. Health is produced from source water quality, E, and the avoidance behavior, T, taken to improve source water quality. Health is increasing and concave. In this stylized model, source water quality is exogenous to the household and avoidance behavior, defined as in-home treatment, is modeled as a continuous time variable.

The household solves

$$\underset{T,Z}{Max}\ U(H(E,T),Z) \quad \text{subject to} \quad Z \leq \bar{Y} - wT \tag{1}$$

where the price of the composite good is normalized to 1 and w denotes the wage rate. $\overline{Y}$ is full income and can be thought as including an exogenous component and the value of the household's time endowment. Plugging the budget constraint in for Z, the utility maximization problem can be written as,

$$\underset{T}{Max}\ U(H(E,T),\overline{Y} - wT) \tag{2}$$

where the only endogenous variable is T. Under differentiability, the household sets demand for

[10]This framework builds upon the defensive expenditure models in Courant and Porter (1981) and Harrington and Portney (1987), and applied to drinking water in developing countries in Pattanayak and Pfaff (2009).

in-home treatment and the composite good such that

$$\frac{U_Z}{U_H} = \frac{1}{w/\frac{\partial H}{\partial T}} \tag{3}$$

The term $w/\frac{\partial H}{\partial T}$ reflects the implicit price of health, where the price to produce a unit of health is increasing in T. Note also that given the budget constraint, a reduction in T is equivalent to an increase in Z.

Now, suppose that there is an exogenous improvement in source water quality. As shown in the appendix, the impact of this improvement on treatment can be decomposed into a substitution effect and an income effect. First consider the income effect; this improvement in drinking water quality provides an in-kind gift that shifts out the opportunity set. If the composite good is a normal good, then the increase in E will unambiguously induce households to increase consumption of Z and thereby decrease T. Now, consider the substitution effect. Whether improvements in source water quality increase or decrease treatment largely depends on the sign of $\frac{\partial^2 H}{\partial E \partial T}$. If the marginal impact of in-home treatment on health is decreasing in source water quality, $\frac{\partial^2 H}{\partial E \partial T} \leq 0$, then an increase in E will increase the relative shadow price of health. This will lead to an increase in Z and thus a reduction in T. In contrast, if $\frac{\partial^2 H}{\partial E \partial T} \geq 0$, then the substitution effect will work towards an increase in treatment. As such, the sign of the total effect of source protection on in-home treatment is ambiguous. Summarizing, if the composite good is a normal good and $\frac{\partial^2 H}{\partial E \partial T} \leq 0$, or if the income effect dominates then treatment will decrease with source protection; if not, it will increase.

4 Estimation strategy

This section sets forth an empirical strategy to test if households reduce or increase demand for in-home treatment in response to source protection, and then estimates which modes of treatment are responsive to source protection.

4.1 Demand for in-home treatment

Household i's demand for any mode of in-home treatment is estimated using a linear probability model (LPM) with district fixed effects and standard errors clustered at the sub-district[11],

$$T_i = \beta X_i + \alpha S_i + \eta Y_i + \gamma_j + u_i. \tag{4}$$

Demand for treatment depends upon observed household and village characteristics (X_i), a household's primary drinking water source (S_i), household wealth and village income (Y_i), and unobservable district characteristics (γ_j) which includes the rural wage rate (w_j). Measures of price include the village market price of fuel wood, which captures the price of boiling treatment, and the district wage rate. Y_i denotes a durables good index, which proxies for household wealth, and mean per capita expenditure in a village. Household characteristics (X_i) include social group, percentage of females, percentage of children under the age of 2 and median education in the village. District fixed effects γ_j control for unobserved heterogeneity at the district. A stochastic component u_i captures the idiosyncratic effect of unobserved factors.

4.2 Identification

This estimating equation will generate inconsistent coefficient estimates of source protection due to omitted variables bias; formally $Cov(S_i, u_i) \neq 0$. Households that use improved sources likely differ from those who do not in health endowments, education, income and access to health services. Additionally villages with improved sources will differ from those without in unobservables such as public infrastructure. These village and household unobservables will likely influence the probability of using an improved source as well as demand for in-home treatment.

If improved drinking water sources and risk averting behavior are both normal goods, then demand for each should increase with income. Similarly, if households have low health endowments,

[11] The data exploit household, village, tehsil, sub-district and district variation. A tehsil represents a unit of government that consists of a collection of villages and cities, with the U.S. equivalent of a civil township. A sub-district describes a region within a district (i.e. part of the county). A district is the U.S. equivalent of a county. A collection of districts comprise a sub-state (e.g. Northeast Uttar Pradesh), and 4-6 sub-states make up a state.

they may try to improve health by choosing an improved drinking water source and engaging in in-home treatment. In these two instances, unobservables will attenuate the estimated effect of source protection on in-home treatment. By contrast, since taps and tube wells were intentionally placed in locations characterized by poor health, few health services or little public infrastructure (Black and Talbot 2005) it may also be the case that the negative correlation between source protection and unobservables will overstate the effect of source protection.

The percent of soft rock in sub-district m is used as the main instrument to predict a household's drinking water source.[12,13] The rock type covering an aquifer impacts both the costs to explore, assess and construct an improved source, as well as the price to extract a unit of groundwater. In this study, rock type is defined as either hard rock, hilly or soft rock. Compared to soft rock, aquifers underlying hard rock are more difficult and costly to assess (Black and Talbot 2005). Second, due to differences in porosity, soft rock aquifers on average tend to hold more water than hard rock aquifers (Chilton 1996). This second feature also implies that the discharge rate or volume of water per minute may be higher in aquifers underlying soft rock. As the discharge rate increases, households can access a larger quantity of water per unit of effort, subsequently lowering the price per unit of water.

Depending on the specification, other instruments may include the minimum aquifer depth or the percent of a sub-district classified as hilly. Minimum aquifer depth measures the average minimum depth to reach the aquifer if a tube well or bore hole was to be constructed. As the minimum depth increases, the price to access a unit of drinking water from an improved source increases since individuals must manually pump water for a longer duration or use more electricity. One concern with aquifer depth is whether it is fixed over time.[14] In an alternative specification

[12] The intuition and details describing these instruments were obtained from extensive conversations with the CGWB in Faridabad, India in fall 2007.

[13] A sub-district defines a unique state, district, sub-round and sub-sample. To spatially evaluate what comprised a sub-district, I looked at the geography of tehsils within a district. Tehsils that constitute a sub-district are geographically clustered; within a sub-district, 80 percent of tehsils are adjacent. In a few instances, I also find that tehsils may be located in more than one sub-district. When tehsils are placed in multiple sub-districts I continue to find strong spatial patterns.

[14] A confined aquifer is one that is sandwiched between two impervious formations (Heath 2004) and as such the depth does not change over time. In contrast, the minimum depth to an unconfined aquifers is the depth to the water table, and varies seasonally and with extraction. Based on conversations with the CGWB, aquifers in this study describe confined, semi-confined and leaky aquifers.

I include the percent of a sub-district that is classified as hilly in place of aquifer depth. In this specification the two instruments are time invariant.

The impact of source protection on in-home treatment is obtained using standard two stage least squares. The first stage is given by,[15]

$$S_i = \rho G_m + \beta X_i + \eta Y_i + \gamma_j + v_i \tag{5}$$

where G_m denotes a vector of hydrogeological characteristics.

4.2.1 Validity concerns

In using geology to predict the impact of source protection on in-home treatment, we must assume that $E(u_i|G_m, X_i, Y_i, \gamma_j) = 0$. I now discuss potential validity concerns and the controls employed to address them.[16]

If geological characteristics determine the ease of constructing an improved drinking water source or the price to extract a unit of drinking water from a tube well, then they also describe the ease of constructing a tube well or extracting water for irrigation or industrial purposes. Because of this, hydrogeological charactersitsics will also likely affect the economic activity, agricultural profits, industrial composition and income of districts (which may also impact demand for in-home treatment) (Badiani 2009, Keskin 2009). To address this, I estimate a district fixed effects model and control for both individual wealth and village per capita expenditure. This specification controls for the possibility that hydrogeology may partly reflect the impact of rural poverty and district development (including economic activity, employment rates, wage rates, and individual wealth) on demand for in-home treatment.

Another validity concern is that hydrogeological characteristics may be correlated with household and village unobservables such as attitudes about drinking water quality, and that these unobservables may influence the decision to treat drinking water. While I cannot rule out the possibility that the instruments are systematically related to unobservables such as social norms,

[15] 2SLS should produce similar effects to those estimated using discrete choice models (Angrist and Pischke 2008).

[16] In the data section, I also explore if the instruments are balanced across observables. Lastly, in the results section I report the results from an overidentification test, where the null hypothesis is that the instruments are exogneous.

I explore the sensitivity of my results to the inclusion of village and household demographics. If attitudes or other unobservables are correlated with the instruments (and treatment), then the coefficient estimate on source protection should meaningfully vary with the inclusion or exclusion of demographic controls that are likely to be related to these unobservables.

In a separate specification, I also estimate the effect of source protection conditional on the composition of the labor force, literacy rates and the employment rates in a sub-district. This allows me to test the extent to which intra-district heterogeneity is systematically correlated with both in-home treatment, and rock type.

4.3 Mode of in-home treatment

Since modes of in-home treatment vary in abatement, it is of interest to estimate which modes of in-home treatment are responsive to changes in source protection. Household i's demand for treatment technology k is estimated using a multinomial logit model with sub-state fixed effects and standard errors clustered at the district.[17] To introduce source protection exogenously, I recharacterize the error term in equation 4 as

$$u_{ik} = \theta v_i + \epsilon_{ik} \tag{6}$$

where $Var(u) = 1$ and respecify demand for in-home treatment as

$$T_{ik} = \beta_k X_i + \alpha_k S_i + \eta_k Y_i + \psi_k D_j + \gamma_l + \theta v_i + \epsilon_{ik}. \tag{7}$$

where $\epsilon_{ik}\tilde{}N(0, 1-\rho^2)$ and $\rho = Corr(v_i, u_{ik})$ (Rivers and Vuong 1988).[18] D_j describes district controls including the composition of the labor force, the literacy rate, the employment rate and the wage rate of agricultural workers, and γ_l captures sub-state fixed effects. Demand for in-home

[17] I also estimate multinomial logit models using district fixed effects. Coefficient estimates in the sub-state fixed effects and district fixed effects specifications are qualitatively similar, however in the district fixed effects model, standard errors are not precisely estimated. Imprecise standard errors occur because for some in-home treatment technologies only a small sample reports using the technology and I have limited degrees of freedom.

[18] Though u_{ik} has a logistic distribution, it approximates a normal distribution, thus allowing the assumption that ϵ_{ik} is normally distributed and the use of the estimator described in Rivers and Vuong (1988).

treatment is estimated using a two-step procedure, where in the first step I estimate equation 5 saving the residuals from the first stage and in the second stage I estimate (Wooldridge 2002),

$$T_{ik} = \beta_k X_i + \alpha_k S_i + \eta_k Y_i + \psi_k D_j + \gamma_l + \theta\widehat{v_i} + \epsilon_{ik}.^{19} \tag{8}$$

If $\theta \neq 0$ then the coefficient estimates are only estimated up to a scale. Nonetheless, the average partial effect of source protection on demand for in-home treatment can be consistently estimated.

5 Data and descriptive results

The data comprise information collected from four sources - the National Sample Survey Organization (NSS), the Central Groundwater Board (CGWB), the 2001 Census, and the Department of Agriculture and Cooperation - and vary in resolution from the household to the district. Table 1 provides descriptive results; means are reported for all households and by primary drinking water source: unimproved, publicly owned improved, and privately owned improved. The household data represent cross-sectional data collected by the NSS in two surveys - Round 54 Schedules 1 and 31. They comprise 20,795 rural households from 1,307 villages in 5 states of rural India sampled between January and June 1998. Data obtained from the CGWB consist of time invariant teshil hydrological characteristics and were merged with the NSS data at the sub-district.[20]

5.1 NSS data

In-home treatment technologies, the outcome variable of interest, are grouped into five categories: no treatment, cloth filter, other filter types, boil and chemical treatment. In total, 18 percent of households engage in some form of in-home treatment, with cloth, other filters, boil, and chemical treatment constituting 76.4, 8.8, 11 and 3.8 percent of the market, respectively. Mean comparisons

[19] Demand for in-home treatment could be estimated using conditional maximum likelihood instead of the 2 step procedure. This approach is more efficient than the 2-step procedure and has the advantage of estimating the unscaled coefficient estimates; however it is computationally intensive.

[20] Data on rural district characteristics were obtained from the 2001 Census and district rural wage data from 1998 were collected by the Department of Agriculture and Cooperation. Variables include literacy rates, population density, employment rates, and the composition of the labor force.

reveal a statistically significant difference in treatment across households using publicly owned improved and unimproved sources; 20 percent of unimproved users treat drinking water, whereas only 15 percent of improved users treat their drinking source.

Approximately 72 percent of the sample selects an improved source, the covariate of interest, as the primary source of drinking water. In the sample, 75 percent of households obtain drinking water from a public source. Since this analysis focuses on the impact of publicly provided improved drinking sources, in some specifications we restrict the sample to households that obtain water from public sources. In 86 percent of villages, households can choose from multiple publicly provided drinking water sources, where the median number of public tube wells in a village equals 4.[21]

Additional household data capture wealth and other demographic characteristics. Household weath as measured by a durable goods index is calculated using a principal components analysis of seven durable goods - bicycle, motorcycle, car, telephone, television, bathroom and latrine - solely owned by a household (Jalan et al. 2009). Information on household composition includes the percentage of children under the age of 2 and the ratio of females to males. Social group indicates whether a household belongs to a historically disadvantaged social group. Village per capita expenditure and village per capita education data were collected from the household expenditure survey (Schedule 1) and were merged with the household drinking water data at the village.[22] Mean monthly expenditure amounts to 428 Rs and median education is literate with some primary schooling.

A comparison of unconditional means indicates that many covariates are unbalanced across source type. In particular, improved users are wealthier, travel a shorter distance to the drinking water source, have a larger share of females in the household and are less likely to be in a disadvantaged social group. Additionally, these households are more likely to reside in villages with a higher median schooling level. These comparisons suggest that improved sources were not randomly assigned to households and that a simple OLS model will likely understate the effect of

[21]The price charged for drinking water from both improved and unimproved sources tends to be zero; of the villages surveyed, only 16 percent have a municipality that charges a flat or fixed tariff.

[22]These two variables are calculated as the mean monthly expenditure and median schooling level of all sampled households in a village, where the mean number of surveyed households in a village equals 16. Household data on education and expenditure were not collected in Schedule 31, so individual information on expenditure and education are not available.

source protection on treatment. Private users are wealthier, less likely to live in a disadvantaged social group and reside in relatively higher income villages than households using either publicly improved or unimproved sources, yet they are less likely to treat the water supply. Thus including private users in the sample may further attenuate the treatment effect.

5.2 Rock type and aquifer depth

Data collected by the CGWB are used to identify a household's primary drinking water source. These data spatially characterize fixed geological characteristics in the states of Uttar Pradesh, Madhya Pradesh, Maharashtra, Tamil Nadu and Rajasthan.[23] The instruments used in this study include the percentage of a sub-district covered by soft rock, hills and the minimum aquifer depth. Variation in these data, as well as locational identifiers, are at the tehsil. And while drinking water sources and in-home treatment vary at the household, these data can only be geographically identified at the sub-district.

Figure 2 illustrates the spatial variation in rock type across Tamil Nadu, one of the 5 states included in the analysis. District boundaries are demarcated in black. As shown in Figure 2, nearly all districts contain within district variation in rock type (as well as minimum aquifer depth). Table 1 highlights that minimum aquifer depth and the percentage of soft rock significantly differ across households with improved and unimproved sources, with improved users residing in areas with a shallower aquifer and a higher percentage of soft rock.

The empirical strategy hinges on the assumption that hydrogeological characteristics are valid instruments for source protection. To examine the plausibility of this assumption, I explore the extent to which geological characteristics directly impact other regressors used in the main estimating equation (i.e. durable goods index). And while I control for the possibility that hydrogeological characteristics may be correlated with household and village demographic characteristics, one indication that the instruments may be systematically correlated with unobservables is if they are systematically correlated with observables. To test for this, I regress each covariate used in the main estimating equation on hydrogeological characteristics controlling for district fixed ef-

[23]Due to confidentiality concerns, the CGWB would only provide these data for 5 randomly selected states.

fects. Columns 7-9 of Table 1 report coefficient estimates for the two instruments, as well as the F-statistic for their joint significance. Hydrogeological characteristics are not individually nor jointly significant in explaining any of the household or village observables. While I cannot state with certainty that the instruments are balanced across unobservables, these results suggest that conditional on district fixed effects they are balanced across observables.[24]

6 Estimation results

I begin by estimating equation 4, a LPM on whether a household chooses in-home treatment or not controlling for district fixed effects. As shown in column 1 of Table 2 source protection reduces demand for in-home treatment by 5 percentage points from a probability of 0.22 to 0.17. As anticipated, demand for in-home treatment increases with wealth and education. Treatment increases by 5 and 2 percentage points with a 1 standard deviation (sd) increase in the durable goods index and 2 percentage points if the median household in a village completes primary school.

As mentioned in the estimation strategy, a household's primary drinking water source may be correlated with unobservables that also impact demand for in-home treatment. To generate consistent estimates on the effect of source protection on in-home treatment, I use hydrogeological characteristics as instruments for a household's primary drinking water source and estimate 2SLS controlling for district unobservables. Table 3 reports estimates from the first stage regression, where results from the preferred specification are reported in column 1.

The percentage of soft rock is strong in predicting source protection, where a 10 percent increase in the percentage of land covered by soft rock is predicted to increase the probability of using an improved source by 2.5 percent. This relationship is expected since the cost to find and construct improved sources is less in soft rock areas (as compared to hard rock). In contrast, aquifer depth provides little explanatory power. The negative coefficient estimate, at least in terms of sign, suggests the probability of source protection decreases as the price to access an improved source increases. The F-statistic for the joint significance of the instruments is 25, indicating

[24] If these regressions do not control for district unobservables, the instruments are jointly significant in explaining many of the regressors. This points to the importance of controlling for district unobservables.

that the instruments are strong, though identifcation of the effect of source protection on in-home treatment comes almost entirely from variation in soft rock. The reported Hansen J-statistic from an over-identification test is 0.280 (P=0.9695) and I fail to reject the null that the instruments are valid. This statistic should be interpreted with the caveat that since aquifer depth is weak in predicting source protection, one can view source protection as almost perfectly identified.

Results from 2SLS are presented in columns 2-7 of Table 2 where column 2 displays results from the preferred specification in which I condition on household and village characteristics, as well as district fixed effects. The presence of an improved source reduces the probability of in-home treatment by 27 percentage points, suggesting that unobservables are positively correlated with both the probability of using an improved sources and the probability of in-home treatment. Coefficient estimates on income and education mirror those reported in column 1. I now explore the sensitivity of my results to the exclusion of households with privately owned sources or surface water users, as well as the validity of my instruments.

6.1 Robustness

To reduce the incidence of diarrheal disease, the government invested in publicly available improved drinking water sources. The sample in column 2 includes households using either publicly owned or privately owned sources. One concern is that households with privately owned sources are driving the results. As shown in Table 1, households with a privatley owned source are less likely to engage in treatment, perhaps because they use an in-premise source, thereby making the probability of contamination during transport low. In column 4, the sample excludes households drinking from privately owned sources. After their exclusion, we continue to find a significant reduction of 26 percentage points in demand for in-home treatment.

If households using taps are driving the difference in treatment between improved and unimproved users, then the limited water quality gains from groundwater expansion cannot be explained by behavior. This is because tap water can either come from surface or groundwater sources Column 5 excludes households with in-home taps and shows that tap users are not driving the treatment effect. Demand for in-home treatment again reduces by 27 percentage points.

One validity concern is that the instruments are systematically correlated with household and village unobservables such as social norms that also impact demand for in-home treatment. While I cannot directly test for this, I examine the sensitivity of my results to the exclusion of household and village demographic characteristics. If the magnitude of the coefficient estimate on source protection meaningfully varies with the inclusion or exclusion of these controls, this provides a signal that household and village unobservables may also be systematically correlated with both the instruments and treatment. In column 3 of Table 2, I estimate a district fixed effects model excluding all household and village regressors aside from source protection. While the exclusion of covariates decreases the precision of the standard errors, the relationship between source protection and in-home treatment is unchanged, suggesting that these socioeconomic characteristics are not systematically correlated with geology.

The inclusion of district fixed effects addresses concerns about the relationship between a district's industrial and employment profile and hydrogeological characteristics. Still one can imagine that the development of a sub-district will also depend on hydrogeological characteristics. As a robustness test, in column 6 of Table 2, I control for the population, literacy rate and composition of the rural labor force in a sub-district. It should be noted that the household and village data (NSS) were collected in 1998, while these data were collected in the 2001 Census. Similar to the results in our preferred specification, the probability of in-home treatment reduces by 28 percent if a household drinks from an improved source.

Depending on the type of aquifer, aquifer depth can change over time. If what I define as confined aquifers are in fact unconfined aquifers, then aquifer depth will change over time depending on weather and extraction, and potentially confound the treatment effect. And while (conditional on soft rock), aquifer depth is not statistically significant in predicting source protection, I test the robustness of the results to its exclusion and the inclusion of the percentage of a sub-district classified as hilly. Similar to the results shown in column 2, source protection reduces the probability of in-home treatment by 30 percentage points.

Together, my results suggest that the presence of an improved source reduces the probability of in-home treatment by 25 to 28 percentage points, where this result is robust to the exclusion household and village socioeconomic characteristics, to the exclusion of households with privately owned sources, to the exclusion of in-home taps and to the inclusion of rural sub-district controls.

6.2 Mode of in-home treatment

Modes of in-home treatment vary in the water quality provided by the technology, the market cost to purchase the good, and the time cost to filter the water. In Table 4, I present results from the multinomial logit described in equation 8. In-home treatment is defined as no treatment, cloth filter, other filter or boiling treatment. Note that since so few households engage in chemical treatment, I exclude chemical users from the sample. Results are reported as the log odds ratio of choosing technology k relative to no treatment.

The presence of an improved source significantly impacts a household's decision to boil drinking water, an effective but time intensive treatment technology. On average the presence of an improved source reduces the probability that a household boils water by 18 percentage points from a probability of 0.20 to 0.02.[25] This qualitative result holds if I include chemical treatment users in the analysis.

7 Valuing the gains from source improvements

Recently, the economics literature has begun to use water quality tests to quantify the water quality gains from source protection, as well as those offset from behavioral choices (Kremer et al. 2011b).[26] Translating source protection and behavioral choices into drinking water quality is an essential link in understanding the relationship between source protection and human health. Making use of data collected in Madhya Pradesh, this study is one of the first to evaluate the effect of behavioral choices on water quality in south Asia. However, as will be discussed, the results that follow should be viewed as a back-of-the-envelope measure of the quality gains offset from behavioral choices (as well as those gained from source protection).

[25] One might be concerned that these results are sensitive to the choice of modeling framework. Similar results were found when using nested logit and multinomial probit models.

[26] Studies have also relied on these tests to examine the impact of information about drinking water quality on behavior (Hamoudi et al. 2011, Jalan and Somanathan 2011).

Source protection also provides households with the benefit of reduced expenditure on mitigating behaviors. Using both field data and NSS data on the cost of in-home treatment technologies, I measure the expenditure savings attributable to the reduction in treatment. Cast differently, the reduction in averting expenditure provides a partial estimate of household willingness to pay for source protection.

7.1 Water quality data

Before describing the data, I want to highlight their limitations. First, the water quality data are not representative of source water quality throughout rural India, but rather are indicative of water quality in two villages. Second, I quantify the abatement supplied by in-home treatment technologies in a laboratory rather than in the field. In doing this, I do not account for the contamination that likely occurs during the transport and storage of water from the source to the household. I also assume that the effectiveness of treatment technologies in the lab mimics abatement in the field. The small sample size and laboratory setting limit the generalization of these data, and should be interpreted within these bounds.

Drinking water samples were collected from all drinking water sources - 6 hand pumps and four open wells - in two rural villages in the district of Jhansi. These samples were tested for the presence or absence of coliform in a laboratory.[27] All four open drinking water wells tested positive for total coliform, fecal coliform and *E. coli*. Of the six hand pumps, three tested positive for total coliform, fecal coliform and *E. coli*.[28] The rate of contamination in improved sources is high but comparable to the rate of fecal contamination detected in Jalan and Somanathan (2008) and McKenzie and Ray (2004). To classify the effectiveness of in-home treatment technologies, coliform counts in one improved and one unimproved source were measured before and after the application of nine in-home treatment technologies - a malmal cloth filter, a candle filter, boiling, chemical treatment and combinations of these.[29]

[27]These tests used the Colilert reagent and the Colilert P/A Test procedure. Colilert is a test certified by the U.S. EPA and used by U.S. drinking water suppliers for compliance with the Safe Drinking Water Act.

[28]Coliform tests performed by the NGO, Development Alternatives, in the same villages confirm my results.

[29]The Colilert reagent and the 15-Tube Most Probable Number Dilution Procedure were used to quantify coliform concentrations.

In Figure 1, I graph the quantity of coliform abatement and market cost of each in-home treatment technology by source. The horizontal axis measures coliform abatement, where a value of 0 indicates that no abatement occurs and 1600 is the maximum quantity of abatement. To collect market cost data on each risk averting technology, I visited one semi-urban and two urban markets in Delhi, surveyed all vendors on the choice set and price of treatment technologies, and purchased all available in-home treatment technologies. I verified the market price of each technology in rural villages in the district of Jhansi. Market costs, defined as the monthly cost of in-home treatment technologies, vary over technology and range from 0 to 95 Rs. It should be noted that neither the costs nor treatment technologies are representaive of treatment markets in India since the sample is confined to a handful of markets.[30]

Five central results emerge from this figure. First, source protection provides sizable water quality gains, though bacterial contamination still exceeds the drinking water standards set by the government. Assuming that no in-home treatment occurs, source protection produces an 82.5 percent or 1320 count reduction in coliforms. Second, each mode of in-home treatment leads to reductions in coliform contamination. Third, more expensive technologies supply larger reductions. Cloth treatment supplies the smallest absolute reduction in coliform counts, with coliform abatement of 700 counts in the unimproved source and 200 counts in the improved source. Boiling and chemical treatment eliminate all coliform from the drinking water samples. Fourth, these coliform tests highlight that the marginal product of in-home treatment is greater for unimproved than improved sources, suggesting that in-home treatment and source protection are substitutes in the production of water quality. Lastly, this figure implies that a reduction in demand for in-home treatment will offset some of the water quality gains from source protection.

7.2 Quality gains from source protection

To quantify the abatement offset by the reduction in demand for in-home treatment, I measure the quality gains of switching every household in the sample from an unimproved to an improved

[30]Some of the cost measures are similar to those reported by others (Clasen et al. 2008, Jalan and Somanathan 2008). There are however differences between chemical costs in this study and the one in Jalan and Somanathan (2008).

drinking water source. Table 5 reports result. Column 1 measures the coliform abatement under the assumption that each household drinks from an unimproved source. Column 2 reports coliform abatement under the assumptions that each household drinks from an improved source and households do not substitute source quality for in-home treatment (α=0). In column 3, households reduce in-home treatment in response to source protection (where coefficients estimates from column 2 of Table 2 are used.) Columns 4 and 5 measure the percentage change in coliform abatement from source protection under the two scenarios.[31] Column 6 quantifies the percent of coliform abatement offset by behavioral choices.

In the first half of the table, I measure the coliform abatement from source protection using the entire sample, and in the second half of the table, I only include households who were predicted to treat an unimproved source.[32] For the entire sample, source protection generates a 64 percent reduction in coliform contamination, and the reduction in demand for in-home treatment offsets coliform abatement by 3.6 percent. The relatively small role behavioral choices play can be largely explained by the 62 percent of households that choose to forgo treatment regardless of source type. Crowding out of coliform abatement can only occur for the households that were predicted to treat an unimproved source. For these households, the reduction in demand for in-home treatment offsets 26.5 percent of the anticipated quality gains from source protection. Nonetheless, source protection increases coliform abatement by more than 15 percent.

In the multinomial model, I can disentangle demand for in-home treatment by technology and assign technology specific abatement. In the row labeled “Boil Treatment”, I measure the per capita change in water quality for households that were predicted to engage in boiling when drinking from an unimproved source. On average, source protection lowers drinking water quality. The estimated 16.4 percent reduction in abatement occurs because boiling eliminates coliform and households that respond to source protection by shifting to no treatment will consume higher concentrations of coliform in drinking water.

[31] Percent change is measured as the change in abatement divided by 1600, since coliform counts can vary from 0 counts to 1600 counts per 100 ml.

[32] To derive coliform abatement estimates for the binary models, I weight the technology-specific coliform abatement by the observed frequency of the technology in the sample population. For example, since 76 percent of the households engaging in in-home treatment use cloth filters, I weight the coliform abatement provided by cloth treatment by 0.76.

Source protection provided substantial water quality gains, especially for households that chose to forgo treatment regardless of their drinking water source. Still, coliform counts in drinking water far exceed the 10 counts per 100 ml standard set by the government. For households that shift from treatment to no treatment in response to source protection, on average over a quarter of the water quality gains are offset. Depending on the treatment method employed, water quality for some households may actually decline in response to source protection.

7.3 Willingness to pay for source protection

Table 6 reports the change in averting expenditure from source protection in 1998 Rs using results from the 2SLS and multinomial logit models. For the entire sample the reduction in averting expenditure provides a welfare gain of 0.5 to 1% of annual expenditure. For households that switch from treatment to no treatment in response to source protection, this gain amounts to 2%.

In column 1 prices are measured as a weighted average of the monthly cost of treatment technologies (collected in the market survey) and the village price of fuel wood (NSS data), assuming that each individual requires 3 liters of drinking water per day (World Bank 2008).[33] In column 2 expenditure is calculated using the opportunity cost of time (OCT), where the value of time is assumed to equal one-third the female district wage rate for field labor (Englin and Shonkwiler 1995).[34] Based upon qualitative surveys, I assume that cloth treatment requires half the amount of time as boiling treatment, and ceramic and chemical treatment require neglible amounts.

As shown in row 1, the per capita change in expenditure on in-home treatment from an improvement in source water quality totals at 1.9 Rs or 2.1 Rs per month, depending on whether in-home treatment is evaluated at the market price or the OCT. This amounts to a 0.52 percent savings in monthly expenditure when price is evaluated using retail prices, and a 0.54 percent savings when price is evaluated at the OCT. The reduction in expenditure on boiling is 0.5 percent or 1.1 percent, when price is measured using market cost data and the OCT, respectively. In

[33]When measuring the price to boil water, I rely on a study in semi-urban India that reports it takes 100 grams of fuel wood and 6.42 minutes to boil a liter of water (Clasen et al. 2008).

[34]Opportunity cost of time ranges between 1/4 and 1/3 wage rate, though recently Kremer et al. (2011b) find the value of time equals 6.2% of the wage rate. Since female wage data for most districts are missing, I use the male wage rate and impose a male-female wage differential of 1.3 (Bhan 2001).

addition to the water quality benefits from source protection, the reduction in expenditure on mitigating behaviors also provides a real benefit to households.

8 Conclusion

This paper finds that households decrease expenditure on in-home treatment and in particular boiling treatment, a costly but effective mode of treatment in response to source protection. The empirical approach uses rock type to account for the endogenous placement of source protection. Given that the analysis relies on one cross-section, the results may be sensitive to unobservable cross-sectional heterogeneity. More generally, despite the validity checks and robustness tests, there may still be selection on unobservables. As such the behavioral, water quality and expenditure results should be interpreted with these caveats in mind.

Using field data on drinking water quality, I show that reductions in demand for in-home treatment crowd out more than a quarter of the coliform abatement from source protection (for households that were predicted to treat an unimproved source). Still, results suggest that source protection leads to gains in drinking water quality with the largest gains accruing to households that engage in no treatment regardless of source. Depending on the relationship between health and water quality, substitution between source quality improvements and private expenditure on water quality may offset some of the health gains from source protection.

While behavioral choices indeed offset the water quality gains, under certain assumptions this compensating behavior is welfare enhancing. The introduction of an improved source will cause utility maximizing households to reallocate time and money from water quality to other welfare-enhancing activities. In addition to the water quality benefits, I find that the per capita savings from source protection amounts to roughly 0.5 to 1 percent of monthly expenditure.

References

[1] **Alberini, Anna, G. Eskeland, A. Krupnick, and G. McGranahan.** 1996. "Determinants of Diarrheal Disease in Jakarta." *Water Resources Research*, 32(7): 2259-2269.

[2] **Angrist, Joshua, and Pischke, Jorn-Steffen**. 2008. *Mostly Harmless Econometrics: An Empiricists Companion*. Princeton, NJ: Princeton University Press.

[3] **Ashraf, Nava, James Berry, and Jesse M. Shapiro**. 2010. "Can Higher Prices Stimulate Product Use? Evidence from a Field Experiment in Zambia." *American Economic Review*, 100(5).

[4] **Ahuja, Amrita, Michael Kremer and Alix Peterson Zwane.** 2010. "Providing Safe Water: Evidence from Randomized Evaluations." *Annual Review of Resource Economics*, 2: 237-256.

[5] **Badiani, Reena**. 2009. "Industrial Location, Growth and Mineral Resources in India."

[6] **Bennear, Lori, Alessandro Tarozzi, Alexander Pfaff, Soumya Balasubramanya, Kazi Matin Ahmed and Alexander Van Geen.** 2013. "Impact of a Randomized Controlled Trial in Arsenic Risk Communication on Household Water-Source Choices in Bangladesh." *Journal of Environmental Economics and Management*, 65: 225-240.

[7] **Bennett, Daniel**. 2012. "Clean Water Makes You Dirty: Water Supply and Sanitation Behavior in the Philippines." *Journal of Human Resources*, 47(1): 146-173.

[8] **Bhan, Gautam**. 2001. *India Gender Profile, Report 62.* Sussex, U.K.: Institute of Development Studies.

[9] **Black, Maggie and Rupert Talbot**. 2005. *Water: A Matter of Life and Health.* New Delhi, India: Oxford University Press.

[10] **Brick, Thomas, B. Primrose, R. Chandrasekhar, S. Roy, and J. Muliyil**. 2004. "Water Contamination in Urban South India: Household Storage Practices and their Implications for Water Safety and Enteric Infections." *International Journal of Hygiene and Environmental Health*, 207: 473-480.

[11] **Butterworth, John and John Soussan**. 2001. "Water Supply and Sanitation and Integrated Water Resources Management: Why Seek Better Integration?" WHIRL Project Working Paper 2. Paper presented at the WHIRL project workshop on 'Water Supply and sanitation and Watershed Development: positive and negative interactions' Andhra Pradesh, India.

[12] **Census of India**. 2001. *India Map.* New Delhi, India: ML InfoMap Pvt. Ltd..

[13] **Chaterjee, A., D. Das, B.K. Mandal, T.R. Chowdhury, G. Samanta, and D. Chakraborti**. 1995. "Arsenic in Ground Water in Six districts of West Bengal, India: the Biggest Arsenic Calamity in the World." *Analyst*, 120: 643-650.

[14] **Checkley, William, Robert, Gilman, Robert Black, Leonardo Epstein, Lilia Cabrera, Charles Sterling, and Lawrence Moulton**. 2004. "Effect of Water and Sanitation on Childhood Health in a Poor Peruvian Peri-Urban Community." *The Lancet*, 363: 112-118.

[15] **Chilton, J.**1996."Groundwater." *Water Quality Assessments - A Guide to Use of Biota, Sediments and Water in Environmental Monitoring.* Ed. D. Chapman on behalf of WHO, UNESCO and UNDP. Chapman and Hall: London.

[16] **Chowdhury, Uttam, B. Biswas, T. Chowdhury, G. Samanta, B. Mandal, G. Basu, C. Chanda, D. Lodh, K. Saha, S. Mukherjee, S. Roy, S. Kabir, Q. Quamruzzaman, D. Chakraborti**. 2000. "Groundwater Arsenic Contamination in Bangladesh and West Bengal, India." *Environmental Health Perspectives*, 108(5): 393-397.

[17] **Clasen, Thomas, Catherine McLaughlin, Neeru Nayaar, Sophie Boisson, Romesh Gupta, Dolly Desai, and Nimish Shah**. 2008. "Microbiological Effectiveness and Cost of Disinfecting Water by Boiling in Semi-urban India." *American Journal Tropical Medicine and Hygiene*, 79(3): 407-413.

[18] **Courant, Paul and Richard Porter**. 1981. "Averting Expenditure and the Cost of Pollution." *Journal of Environmental Economics and Management*, 8:321-329.

[19] **Cutler, David and Grant Miller**. 2005. "The Role of Public Health Improvements in the Health Advances: The Twentieth-Century United States." *Demography*, 42(1): 1-22.

[20] **Dasgupta, Purnamita**. 2004. "Valuing Health Damages from Water Pollution in Urban Delhi, India: A Health Production Function Approach." *Environment and Development Economics*, 9: 83-106.

[21] **Dept. of Drinking Water Supply**. 2007. "Guidelines for the Preparation of Legislation for Framing Drinking Water Regulations.". Draft. New Delhi, India: Ministry of Rural Development, Government of India.

[22] **Englin, J. and J.S. Shonkwiler**. 1995. "Modeling Recreation Demand in the Presence of Unobservable Travel Costs: Towards a Travel Price Model." *Journal of Environmental Economics and Management*, 29: 308-377.

[23] **Esrey, Steven A. and Jean-Pierre Habicht**. 1986. "Epidemiologic Evidence for Health Benefits from Improved Water and Sanitation in Developing Countries." *Epidemiological Review*, 8(1):117-128.

[24] **Esrey, Steven A., J. B. Potash, L. Roberts, L. and C. Shiff**. 1991. "Effects of Improved Water Supply and Sanitation on Ascariasis, Diarrhoea, Dracunculiasis, Hookworm Infection, Schistosomiasis, and Trachoma." *Bulletin World Health Organization*, 69(5): 609-621.

[25] **Fewtrell, Lorna, R. Kaufmann, D. Kay, W. Enanoria, L. Haller and J. Colford Jr.**. 2005. "Water, Sanitation, and Hygiene Interventions to Reduce Diarrhoea in Less Developed Countries: a Systematic Review and Meta-Analysis." *Lancet Infectious Diseases*, 5 :42-52.

[26] **Hamoudi, A., M. Jeuland, S. Lombardo, S. Patil, S. Pattanayak, and S. Rai**. 2012. "The Effect of Water Quality Testing on Household Behavior: Evidence from an Experiment in Rural India." *American Journal Tropical Medicine and Hygiene*," 87(1): 18-22.

[27] **Harrington, Winston and Paul Portney**. 1987. "Valuing the Benefits of Health and Safety Regulation." *Journal of Urban Economics*,22:101-112.

[28] **Islam, M. Sirajul, A. Brooks, M.S. Kabir, I.K. Jahid, Shafiqul M. Islam, D. Goswami, G.B. Nair, C. Larson, W. Yukiko, and S. and Luby**. 2007. "Faecal Contamination of Drinking Water Sources of Dhaka City during the 2004 Flood in Bangladesh and Use of Disinfectants for Water Treatment." *Journal of Applied Microbiology*, 103(1):80-87.

[29] **Jalan, Jyotsna, E. Somanathan, and S. Chaudhuri**. 2009. “Awareness and the Demand for Environmental Quality: Survey Evidence on Drinking Water in Urban India.” *Environment and Development Economics*, 14(6): 665-692.

[30] **Jalan, Jyotsna and E. Somanathan**. 2008. “The Importance of Being Informed: Experimental Evidence on Demand for Environmental Quality.” *Journal of Development Economics*, 87: 14-28.

[31] **Jalan, Jyotsna and Martin Ravallion**. 2003. “Does Piped Water Reduce Diarrhea for Children in rural India?” *Journal of Econometrics*, 112: 153-173.

[32] **Harrington, Winston and Paul Portney**. 1987. “Valuing the Benefits of Health and Safety Regulation.” *Journal of Urban Economics*, 22:101-112.

[33] **Keskin, Pinar**. 2009. “Thirsty Factories, Hungry Farmers: Intersectoral Impacts of Industrial Water Demand.” http://www.pinarkeskin.com/Keskin–ThirstyFactories.pdf.

[34] **Kremer, Michael, Edward Miguel, Sendhil Mullainathan, Clair Null and Alix Zwane**. 2011. “Social Engineering: Evidence from a Suite of Take-up Experiments in Kenya” *Working Paper.*

[35] **Kremer, Michael, Jessica Leino, Edward Miguel and Alix Zwane**. 2011. “Spring Cleaning: A Randomized Evaluation of Source Water Quality Improvement” *Quarterly Journal of Economics*, 126(1): 145-205.

[36] **Larson, Bruce and Ekaterina Gnedenko**. 1999. “Avoiding Health Risks from Drinking Water in Moscow: An Empirical Analysis.” *Environment and Development Economics*, 4: 565-581.

[37] **Luoto, Jill, David Levine and Jeffrey Albert**. 2011. “Information Persuasion: Achieving Safe Water Behaviors in Kenya.” *RAND Working Paper 885.*.

[38] **Madajewicz , Malgosia, Alexander Pfaff, Alexander van Geen, Joseph Graziano, Iftikhar Hussein, Hasina Momotaj, Roksana Syliv and Habibul Ahsan, Habibul**. 2007. “Can Information Alone Change Behavior? Arsenic contamination of groundwater in Bangladesh.” *Journal of Development Economics*, 84: 731-754.

[39] **Mcconnell, Kenneth and Marcia Rosado**. 2000. “Valuing Discrete Improvements in Drinking Wter Quality Revealed Preferences.” *Water Resources Research,* 36: 1575-1582.

[40] **McKenzie, Dave and I. Ray**. 2004. “Household Water Delivery Options in Urban and Rural India.” Paper presented at the 5th Stanford Conference on Indian Economic Development, CA.

[41] **Ministry of Agriculture**. 2000. *Agricultural Wages in India 1997-1998.* New Delhi, India: Government of India, Department of Agriculture and Cooperation.

[42] **Mintz E., F. Reiff, and R. Tauxe**. 1995 “Safe Water Treatment and Storage in the Home: A Practical New Strategy to Prevent Waterborne Diseases.” *Journal of the American Medical Association*, 273: 948-953.

[43] **National Sample Survey Organization**. 1999. "Drinking Water, Sanitation and Hygiene in India: NSS 54th Round." Report 449.

[44] **Pattanayak, Subhrendu K., Jui-Chen Yang, Dale Whittington and K.C. Bal Kumar.** 2005. "Coping with Unreliable Public Water Supplies: Averting Expenditures by Households in Kathmandu, Nepal." *Water Resources Research*, 41(2).

[45] **Pattanayak, Subhrendu K. and Alexander Pfaff**. 2009. "Behavior, Environment and Health in Developing Countries: Evaluation and Valuation." *Annual Review of Resource Economics*, 1:183-222.

[46] **Peltzman, Sam**. 1975. "The Effect of Automobile Safety Regulations." *Journal of Political Economy*: 677-726.

[47] **Planning Commission**. 2002. "India Assessment 2002, Water supply and Sanitation." A WHO/UNICEF Sponsored Study.

[48] **Oettinger, Gerald S.**. 1999. "The Effects of Sex Education on Teen Sexual Activity and Teen Pregnancy." *The Journal of Political Economy*, 107(3):606-644.

[49] **Quick, R.E., L.V. Venczel, E.D. Mintz, L. Soleto, J. Aparicio, M. Gironaz, L. Hutwagner, K. Greene, C. Bopp, K. Maloney, D. Chavez, M. Sobsey, and R.V. Tauxe**. 1999. "Diarrhoea Prevention in Bolivia through Point-of-Use Water Treatment and Safe Storage: a Promising New Strategy." *Epidemiological Infect.*, 122:83-90.

[50] **Rivers, D. and Q. Vuong**. 1988. "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models." *Journal of Econometrics*, 39: 347-366.

[51] **Small, Kenneth and Kurt Van Dender**. 2005. "The Effect of Improved Fuel Economy on Vehicle Miles Traveled: Estimating the Rebound Effect Using U.S. State Data." UC Irvine Economics Working Paper 05-06-03.

[52] **Somanathan, E**. 2010. "Effects of Information on Environmental Quality in Developing Countries." *Review of Environmental Economics and Policy*, 4(2): 275-292.

[53] **Um, Mi-Jung, Seung-Jun Kwak and Tai-Yoo Kim**. 2002. "Estimating Willingness to Pay for Improved Drinking Water Quality Using Averting Behavior Method with Perception Measure." *Environmental and Resource Economics*, 21: 287-302.

[54] **World Health Organization(WHO)**. 2000. "World Health Organization. Global water supply and sanitation assessment 2000 report." WHO: Geneva.

[55] **World Health Organization(WHO)**. 2004. "Water, Sanitation, and Hygiene links to Health: Facts and Figures." WHO: Geneva.

[56] **Wooldridge, Jeffrey M.**. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

[57] **World Bank**. 2008. "World Bank Study on Review of Effectiveness of Rural Water Supply Schemes in India." World Bank Policy Paper, New Delhi: World Bank.

[58] **Wright, Jim, Stephen Gundry, and Ronan Conroy**. 2004. "Household Drinking water in Developing Countries: a Systematic Review of Microbiological Contamination between Source and Point-of-use." *Tropical Medicine and International Health*, 9(1):106-117.

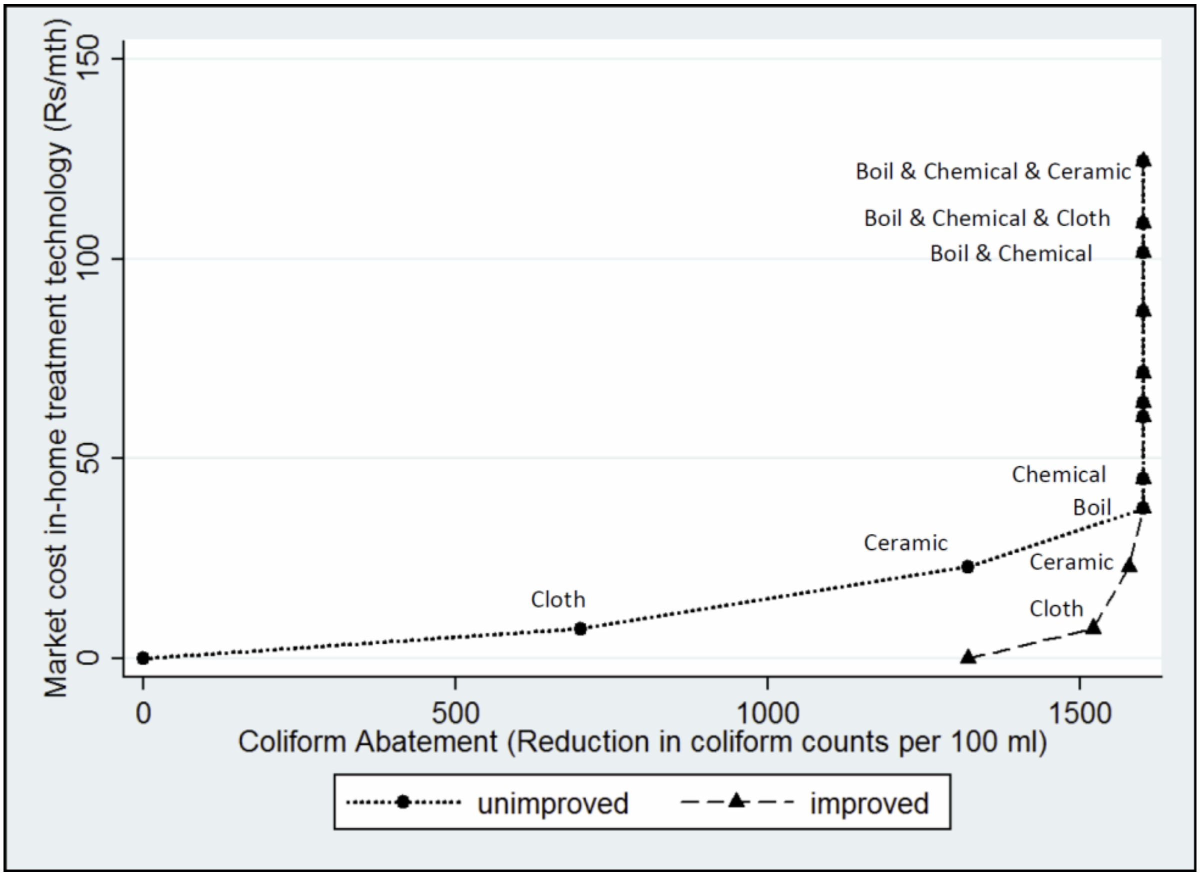

Figure 1: Price and Abatement Supplied by Treatment Technologies

Table 1: Summary Statistics

	Obs	Mean	SD	Mean by source type			OLS of regressors on instruments		
				Unimp	Imp	Private	Min depth	% soft rock	F-stat
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Household Variables (Source: NSS Round 54 Schedule 31)									
Improved source (1=yes)	20795	0.721	0.448						
In-home tap (1=yes)	20789	0.042	0.200	0	0	0.282			
Distance to source (km)	18226	0.163	0.229	0.213***	0.132				
Treatment (1=yes)	20795	0.182	0.386	0.196***	0.153	0.104			
Plain cloth (1=yes)	20794	0.140	0.347	0.153***	0.115	0.063			
Other filter	20794	0.016	0.124	0.017**	0.013	0.019			
Boil (1=yes)	20794	0.020	0.139	0.018	0.018	0.016			
Chem treatment(1=yes)	20794	0.007	0.084	0.0081*	0.007	0.006			
Durables index	20795	0.286	0.434	0.232***	0.252	0.531	-0.050	0.028	0.16
Social group (0= disadvantaged)	20795	0.348	0.476	0.346***	0.335	0.541	-0.047	-0.004	0.24
Share children < 2	20795	0.036	0.079	0.036	0.035	0.040	7.59e-03	-0.0104	1.14
Share females	20795	0.478	0.195	0.471***	0.481	0.474	-6.03e-03	-0.018	0.86
Village variables (Source: NSS Round 54 Schedules 1 and 31)									
Price fuelwood reported	1307	0.770	0.421	0.775	0.789	0.514	-0.185	0.175	0.24
Price fuelwood (Rs/kg)	1007	0.916	0.419	0.944	0.912	0.818	0.0501	0.012	0.08
Per cap expend (Rs/mth)	1307	428.3	183.3	422.9	426.9	469.5	5.05	10.01	0.64
Per cap education (school)	1307	2.14	1.13	2.01***	2.21	2.16	0.061	-0.076	0.04
District variables (Source: Central Groundwater Board)									
Min aquifer depth (m)	477	29.8	19.33	35.8***	28.2	31.2			
Percent soft rock	477	0.254	0.282	0.141***	0.274	0.494			
Percent hilly	477	0.076	0.134	0.082	0.076	0.001			

Notes: This table shows means and standard deviations for household, village and subdistrict covariates.
Columns 4-6 report means by publicly owned improved, publicly owned unimproved and private users. Asterisks
in column 4 indicate significant differences in means across improved and unimproved users; *** $p<0.01$, ** $p<0.05$, * $p<0.1$
Columns 7-9 show coefficients from regressions of each covariate on groundwater characteristics controlling for district fixed effects.
Standard errors are clustered at the sub-district. Column 9 reports the F-stat for joint significance of the coefficient estimates.

Table 2: Binary Models of Demand for In-Home Treatment

	OLS	2SLS					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	All	All	All	No private	No i/s tap	All	All
Improved source	-0 .0468***	-0.267*	-0.253	-0.258*	-0.268*	-0.267*	-0.296*
	(0.0107)	(0.158)	(0.157)	(0.145)	(0.151)	(0.158)	(0.160)
Durables index	0.111***	0.122***		0.122***	0.110***	0.122***	0.124***
	(0.0112)	(0.0135)		(0.0128)	(0.0118)	(0.0135)	(0.0139)
Social group	0.0237***	0.0273***		0.0309***	0.0298***	0.0273***	0.0278***
	(0.00721)	(0.00804)		(0.00818)	(0.00783)	(0.00804)	(0.00811)
Price fuelwood reported	0.0104	0.0142		0.0221	0.0151	0.0142	0.0147
	(0.0206)	(0.0221)		(0.0231)	(0.0221)	(0.0221)	(0.0223)
Price fuelwood	0.0237	0.0249		0.0297	0.0277	0.0249	0.0251
	(0.0188)	(0.0189)		(0.0197)	(0.0190)	(0.0189)	(0.0191)
Village expend(1000 Rs)	0.0509	0.0469		0.0437	0.0436	0.0469	0.0463
	(0.0324)	(0.0357)		(0.0375e-05)	(0.0353)	(0.0357)	(0.0363e)
Village education	0.0108	0.0160*		0.0165*	0.0160*	0.0160*	0.0167**
	(0.00680)	(0.00822)		(0.00878)	(0.00818)	(0.00822)	(0.00825)
Fixed Effects	district	district	district	district	district	district	district
Subdistrict controls	no	no	no	no	no	yes	no
Observations	20795	20795	20795	18241	19928	20795	20795

Notes: The dependent variable is whether (1) a household engages in any in-home treatment or not (0). Column 1 reports results from a linear probability mode with standard errors clustered at the subdistrict. Columns 2-6 report results from 2SLS. In columns 2-6 instruments are the percent soft rock and minimum aquifer depth. In column 7 instruments are the percent soft rock and the percent hilly. Additional variables in cols.1-2 and 4-6 are the share children < 2 and share of females; col 6. also includes the literacy rate, employment rate, excludes population and composition of the labor force in a sub-district. Column 4 excludes hh with private sources. and col. 5 in-home tap users Asterisks indicate statistical significance; *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Demand for Improved Drinking Water Sources (First Stage)

	(1) All	(2) All	(3) No private	(4) No i/s tap	(5) All	(6) All
Percent soft rock	0.251***	0.252***	0.295***	0.261***	0.251***	0.207**
	(0.0934)	(0.0946)	(0.106)	(0.0962)	(0.0936)	(0.0828)
Min aquifer depth	-0.00102	-0.00107	-0.00115	-0.00100	-0.00102	
	(0.00104)	(0.00110)	(0.00117)	(0.00105)	(0.00104)	
Percent hilly						0.175
						(0.211)
Durables index	0.0513***		0.00573	0.0313**	0.0512***	0.0515***
	(0.0127)		(0.0147)	(0.0135)	(0.0127)	(0.0127)
Social group	0.0165		0.00163	0.0126	0.0165	0.0163
	(0.0114)		(0.0131)	(0.0119)	(0.0114)	(0.0114)
Village expend(1000 Rs)	-0.0206		-0.0387	-0.0312	-0.0208	-0.0226
	(0.0556)		(0.0633)	(0.0586)	(0.0566)	(0.0556)
Village education	0.0237***		0.0257**	0.0228**	0.0236***	0.0238***
	(0.00884)		(0.00984)	(0.00926)	(0.00885)	(0.00884)
Fixed Effects	district	district	district	district	district	district
Observations	20795	20795	18247	19935	20795	20795
R^2	0.190	0.183	0.181	0.191	0.196	0.189
F-stat	25.24	25.06	26.96	26.33	46.86	24.74

Notes: The dependent variable is whether (1) or not (0) a household chooses an improved drinking water source. Columns 1-6 report estimates from linear probability models. Standard errors are clustered at the subdistrict. Col 3 excludes hh with privately owned sources and col. 4 excludes in-home tap users. Additional variables included in cols. 1 and 3-6 are the household share < 2, share females, price of fuel wood, price fuel wood reported and the max school year. Column 5 also includes the literacy rate, employment rate, population and employment type in a sub-district. *** p<0.01, ** p<0.05, * p<0.1. The Kleibergen-Paap Wald F-statistic is reported for the instruments.

Table 4: IV Multinomial Logit Model of In-Home Treatment

Variable	(1) Cloth	(2) Ceramic	(3) Boil
Improved source	-5.65	0.471	-6.63*
	(3.54)	(3.91)	(3.71)
Durables index	1.00***	0.884***	1.77***
	(0.231)	(0.263)	(0.223)
Social group	0.577***	0.146	0.0125
	(0.137)	(0.185)	(0.149)
Price fuelwood reported	0.789**	-0.233	0.401
	(0.378)	(0.617)	(0.418)
Price fuelwood	-0.00275	0.594	-0.147
	(0.216)	(0.365)	(0.209)
Village expend (100 Rs)	0.0198	0.115**	-0.0433
	(0.0381)	(0.0500)	(0.0387)
Village education	0.249**	-0.122	0.104
	(0.107)	(0.149)	(0.111)

Notes: The dependent variable is the mode of in-home treatment where results for cloth filters, ceramic filters and boiling are reported in Columns 1, 2 and 3, respectively. Columns 1-3 report log odds ratios from a multinomial logit model using sub-state fixed effects and district controls. Robust standard errors clustered at the district are in parentheses. The base outcome in all columns is no treatment. Instruments are sub-district groundwater characteristics. Additional variables included are share children < 2 and share of females. The number of observations is 20,951 households. Asterisks denote significance; *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 5: Abatement Offset by Change in In-home Treatment

	(1)	(2)	(3)	(4)	(5)	(6)
	Unimproved	Improved ($\alpha = 0$)	Improved ($\alpha \neq 0$)	Change Quality ($\alpha = 0$)	Change Quality ($\alpha \neq 0$)	Offset Quality
Entire Sample						
Any Mode of Treatment	323	1400	1343	67.3%	63.8%	3.6%
Households predicted to treat an unimproved source						
Any Mode of Treatment	861	1534	1106	42.1%	15.3%	26.8%
Boil Treatment	1600	1600	1341	0.0%	-16.2%	-16.2%

Notes: Quality and expenditure estimates are calculated using estimates from the 2SLS and IV multinomial logit models. Columns 1, 2 and 3 describe abatement in the unimproved and improved source. In columns 2 and 4, the constrained model, the coefficient on source is zero. In columns 3 and 5, the choices model, in-home treatment depends on a household's water source. Columns 4-5 describe the difference between cols. 2 and 1, and 3 and 1 as a percentage

Table 6: Change in Averting Expenditure from Improved Sources

	(1)	(2)
	Market cost	OCT
2SLS (col. 2 of Table 2)		
Change expend (Rs/mth)	1.88	2.03
Change expend/Month expend	0.514%	0.537%
IV Multinomial logit (col. 3 of Table 4)		
Change boiling (Rs/mth)	1.65	4.06
Change boiling/Month expend	0.437%	1.06%

Notes: The change in expenditure is calculated using a weighted sum of technology prices and the price of fuel wood in column 1 and a fraction of the district wage in column 2. The change in expenditure is reported in absolute terms, and as a percentage of monthly expenditure.

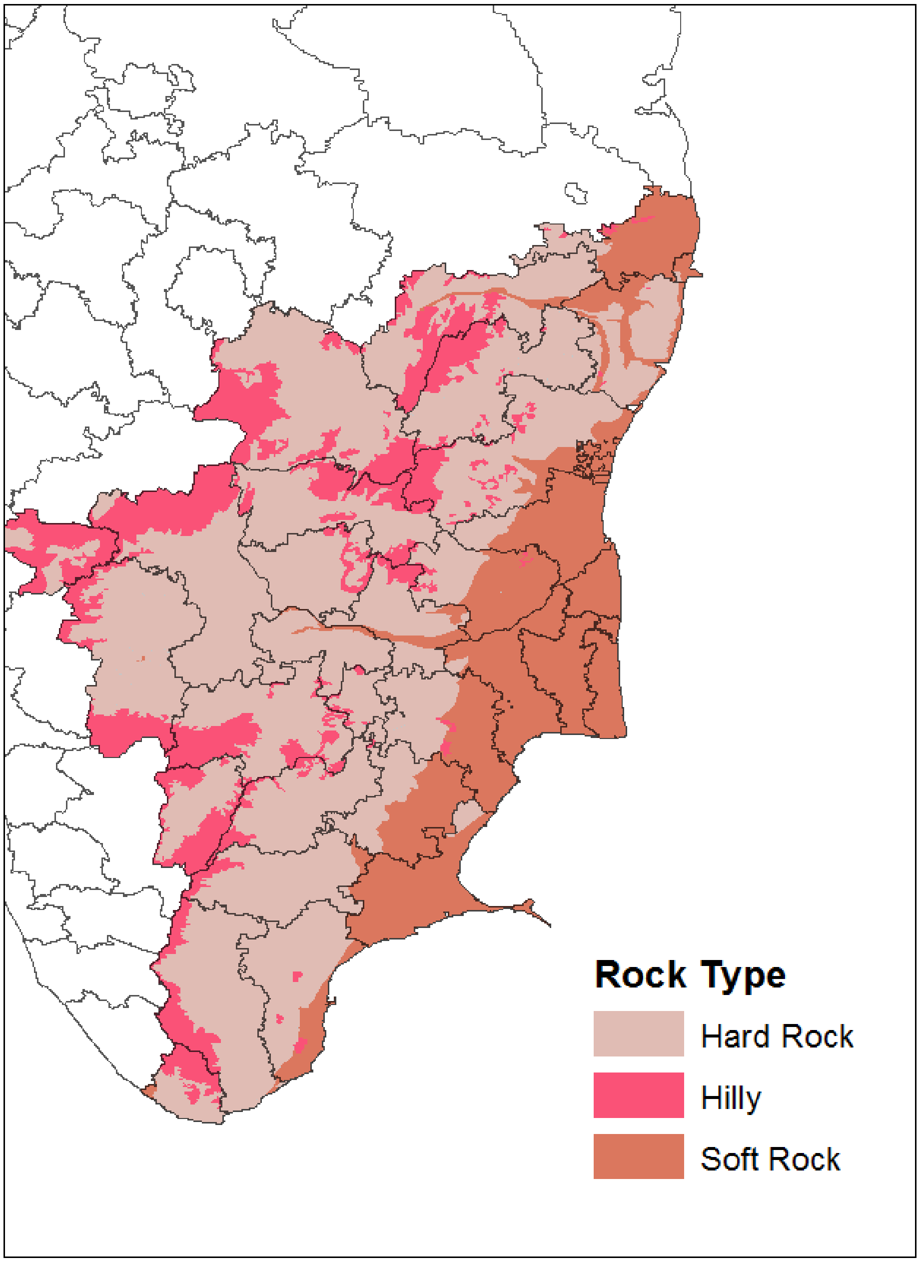

Figure 2: Spatial distribution of rock type in Tamil Nadu