The World Bank Economic Review, 00(0), 2020, 1–29
doi: 10.1093/wber/lhaa009
Article

Demand for Information on Environmental Health Risk, Mode of Delivery, and Behavioral Change: Evidence from Sonargaon, Bangladesh

Alessandro Tarozzi, Ricardo Maertens, Kazi Matin Ahmed, and Alexander van Geen

Abstract

Millions of villagers in Bangladesh are exposed to arsenic by drinking contaminated water from private wells. Testing for arsenic can encourage switching from unsafe wells to safer sources. This study describes results from a cluster randomized controlled trial conducted in 112 villages in Bangladesh to evaluate the effectiveness of different test selling schemes at inducing switching from unsafe wells. At a price of about US0.60, only one in four households purchased a test. Sales were not increased by informal inter-household agreements to share water from wells found to be safe, or by visual reminders of well status in the form of metal placards mounted on the well pump. However, switching away from unsafe wells almost doubled in response to agreements or placards relative to the one in three proportion of households that switched away from an unsafe well with simple individual sales.

JEL classification: I12, I15, I18, Q53

Keywords: arsenic, Bangladesh, environmental health risk

Alessandro Tarozzi (corresponding author) is an Associate Professor at the Universitat Pompeu Fabra and the Barcelona Graduate School of Economics, Barcelona; his email address is alessandro.tarozzi@upf.edu. Ricardo Maertens is a Post-Doctoral Research Fellow at Harvard University; his email address is maertensodria@fas.harvard.edu. Kazi Matin Ahmed is a Professor at the University of Dhaka; his email address is kazimatin@yahoo.com. Alexander van Geen is Lamont Research Professor at Columbia University, New York; his email address is avangeen@ldeo.columbia.edu. We are very grateful to Prabhat Barnwal for the idea of forming groups to share risk, which started this project. We acknowledge financial support from the Earth Clinic at the Earth Institute, Columbia University (this is LDEO publication number 8384), NIEHS (grant P42 ES010349), NSF (grant ICER1414131), and the Ministerio de Economía y Competitividad of the Spanish Government (grant ECO2015-69869-R). We thank Dr. Zahed Masud of the Arsenic Mitigation and Research Foundation for obtaining the necessary approvals in Bangladesh and for managing the finances of the project. We are also grateful to the survey team, and in particular Ershad Bin Ahmed, for their work during the data collection, and to Mr. Saifur Rahman from the Department of Public Health Engineering for his insights throughout the project concerning arsenic mitigation. The paper benefited from constructive comments and suggestions from the editor and anonymous referees, as well as many colleagues at Lund University, the Helsinki Center of Economic Research, the Workshop on Health Economics and Health Policy (Heidelberg), Université Catholique de Louvain, University of Bristol, King's College (London), Venezia Cà Foscari, CEPREMAP India-China Conference (PSE, Paris), the VI Navarra Center for International Development (Fundación Ramón Areces, Madrid), Oxford (CSAE), Stanford University, Université de Genève, Uppsala, Università di Torino, Waikato University, and the 14th CEPAR Summer Workshop in the Economics of Health and Ageing at the University of New South Wales, Sydney. The study was approved

1. Introduction

Poor health stands out as a common feature of life in less developed countries (LDCs). Several factors contribute to the persistence of the problem, including poor availability and the high cost of good quality health care, insufficient investment in prevention, and the frequent reliance on ineffective and sometimes unnecessarily expensive treatments (see Dupas 2012; Dupas and Miguel 2017; Tarozzi 2016 for recent reviews). Information campaigns on health risks are sometimes seen as an appealing tool in environmental and health policy. This is because they can be relatively inexpensive to run when compared to other options such as investments in infrastructure or public health measures needed to eliminate the risk at its root. Some health conditions are in fact preventable if appropriate risk-mitigating behavior is adopted. However, governments in LDCs may lack the resources or the political will to carry out even simple information campaigns (let alone campaigns that provide reports specific to each household), and information alone is often not sufficient to promote positive changes in behavior.

This paper describes the results of a randomized controlled trial (RCT), carried out in the Sonargaon subdistrict of Bangladesh, to examine the impact of different ways of selling a contaminant test on risk-avoiding behavior. Households were offered tests that measure tube-well water contamination with arsenic, a common occurrence in the area. The primary objective was to determine whether a novel mode of test delivery, leveraging within-village solidarity networks, could increase health-protective behavioral responses relative to the standard delivery of private information to well users. In a first group of 49 randomly selected villages, field tests were offered at a (subsidized) price of BDT 45 (about USD 0.60 at current nominal exchange rates, close to the price of 1 kg of rice in Dhaka), an amount estimated to be just enough to cover the salary of the surveyors hired for the project.[1] In an additional subset of 48 villages, surveyors received incentives to offer tests—at the same price, BDT 45—to *groups* of buyers: group members were asked to sign an informal agreement according to which those with safe wells would share their well with others in the group whose well water was found to be unsafe. The agreement was not binding legally, but the prior was that it would increase rates of switching from unsafe sources through two mechanisms: first, by making sharing more likely through a form of soft-commitment and, second, by facilitating the spread of information about the safety of wells, thereby facilitating the identification of safe options within the village. While a large literature documents the importance of village networks to cope with shocks, including health shocks (see Fafchamps 2011 for a review), we are not aware of other work studying how informal networks can help in creating opportunities to reduce environmental health risk.[2]

The study also examines the impact of a second mode of information delivery, in the form of metal placards attached to the well spout to convey test results. Budget limitations, however, only allowed the inclusion of 15 villages in this experimental arm, reducing statistical power. In these villages, individuals who purchased a test at the same price of BDT 45 were also given a metal placard of a color depending on the arsenic level: blue for arsenic up to 10 ppb (parts per billion or micrograms per liter), the World Health Organization guideline for arsenic in drinking water, green if above 10 and up to 50 ppb, and red if “unsafe,” that is, above the national government standard of 50 ppb. Similar metal placards have been used before in some testing campaigns (Opar et al. 2007; van Geen et al. 2014), as a more durable alternative to the routine strategy—adopted also during a past nationwide testing campaign in Bangladesh—of applying to the well spout red or green paint that often becomes invisible within a year (Pfaff et al. 2017).

by the Columbia University IRB (Protocol AAAN9900) and the Government of Bangladesh NGO Affairs Bureau. All errors are our own. A supplementary online appendix is available with this article at the *World Bank Economic Review* website.

[1] Throughout the paper, Bangladesh taka (BDT) are converted into United States dollars (USD) using a nominal exchange rate of BDT 80/USD 1, and a Purchasing Power Parity (PPP) exchange rate of 23.145, as indicated in World Bank (2015, table 2.1).

[2] In broadly related work, Goldberg, Macis, and Chintagunta (2018) show that peer networks can be leveraged to improve screening for tuberculosis in Indian urban areas.

Such visible indicators are a reminder about the status of specific tube wells with respect to arsenic and can spread this information throughout the village. In different contexts, other researchers have found large impacts of reminders on health-related behavior, for instance through the use of SMS messages; see Pop-Eleches et al. (2011) and Raifman et al. (2014). However, the cost of the placards (about BDT 80) is high enough to increase significantly the total cost of testing. It was therefore important to determine whether the placards made any difference relative to the alternative solution (adopted in the two experimental arms described earlier) of informing the household by providing a less expensive laminated card indicating the test result and encouraging the household to keep the card in the house.

Despite much progress on numerous health indicators (Chowdhury et al. 2013), Bangladesh remains in the midst of a severe health crisis due to the widespread presence of naturally occurring arsenic (As) in shallow aquifers (see Ahmed et al. 2006; Johnston et al. 2014; Pfaff et al. 2017). The problem, due to the widespread presence in the country of geological conditions conducive to accumulation of arsenic in groundwater, is compounded by millions of households in rural areas relying on water from privately owned, unregulated shallow tube wells for drinking and cooking. Using nationwide data from 2009, Flanagan, Johnston, and Zheng (2012) estimated that, in a country of more than 150 million people, about 20 million were likely to have been exposed to arsenic levels above the official Bangladesh standard of 50 ppb, while almost one-third of the population was likely to have been exposed to levels above the significantly lower WHO guideline of 10 ppb.

The most visible health consequences of chronic exposure to arsenic from drinking tube-well water in South Asia, such as cancerous skin lesions and loss of limb, were recognized in the state of West Bengal, India in the mid-1980s (Smith, Lingas, and Rahman 2000). It has since then been shown on the basis of long-term studies in neighboring Bangladesh that arsenic exposure increases mortality due to cardiovascular disease, and may inhibit intellectual development in children and be detrimental for mental health (Wasserman et al. 2007; Argos et al. 2010; Rahman et al. 2010; Chen et al. 2011; Chowdhury, Krause, and Zimmermann 2016). These health effects are accompanied by significant economic impacts: exposure to arsenic has been estimated to reduce household labor supply by 8 percent (Carson, Koundouri, and Nauges 2011) and household income by 9 percent per every earner exposed (Pitt, Rosenzweig, and Hassan 2015), while Flanagan, Johnston, and Zheng (2012) calculated that a predicted arsenic-related mortality rate of 1 in every 18 adult deaths represents an additional economic burden of USD 13 billion in lost productivity alone over the next 20 years.

Piped water from regulated and monitored supplies is likely to be the most effective policy answer, but such a solution would require immense investments in infrastructure that may not be sustainable or cost effective for the foreseeable future, so that identifying short-term mitigation strategies remains essential. The consensus view is that household-level water treatment, dug wells, and rainwater harvesting are not viable alternatives for lowering arsenic exposure because of the cost and logistics of maintaining such systems in rural South Asia (Ahmed et al. 2006; Howard et al. 2006; Johnston et al. 2014; Sanchez et al. 2016). In contrast, despite being the main source of arsenic exposure, tube wells remain the most effective way of providing safe drinking water to the rural population of Bangladesh in the short to medium term (Krupoff, Mobarak, and van Geen 2020). With the exception of the most severely affected areas of the country, the spatial distribution of high- and low-arsenic wells is highly mixed, even over small distances. At the same time, whether a well is contaminated with arsenic or not rarely changes over time (van Geen et al. 2007; McArthur et al. 2010). Therefore, exposure among users of arsenic-contaminated wells can often be avoided by switching to a nearby safe well, be it a shallow private well or a deeper—which usually means safer—community well (van Geen et al. 2002; van Geen et al. 2003). Using data from Araihazar, a subdistrict bordering the location of this study, Jamil et al. (2019) estimate that blanket testing campaigns that inform households about the arsenic contamination of all private wells were significantly more cost effective at reducing arsenic exposure than the provision of piped water, or the construction of deep wells by the government.

Previous campaigns aimed at testing tube-well water for arsenic have only partially succeeded at promoting risk-avoiding behavior, highlighting the need to devise novel strategies to achieve this goal. Between 1999 and 2005, the Bangladesh Department of Public Health Engineering (DPHE), with support from the World Bank, DANIDA, UNICEF, and a number of non-governmental organizations (NGOs), coordinated the Bangladesh Arsenic Mitigation and Water Supply Program (BAMWSP) testing campaign. The campaign tested close to 5 million wells making use of field kits, and identified them as "safe" or "unsafe"—according to the Bangladesh standard of 50 ppb—by painting the well spout with green or red paint, respectively. Several studies have documented switching rates from an unsafe to a safe well after testing of between one-third and three-quarters, with higher switching rates in trials that provided repeat visits and information campaigns on arsenic health risks, in some cases with objective measures of exposure taken in the form of urine samples (Chen et al. 2007; Madajewicz et al. 2007; Opar et al. 2007; George et al. 2012; Bennear et al. 2013; Balasubramanya et al. 2014; Inauen, Tobias, and Mosler 2014; Pfaff et al. 2017). Despite these partial successes, a substantial fraction of households continue to use unsafe wells today and it is thus important to identify mechanisms to increase risk-mitigating responses. In addition, millions of new wells have sprouted in the country, and in most cases users do not know the arsenic level of the water, because campaigns such as BAMWSP have not been replicated, and a market for tests barely exists. There are a few commercial laboratories in Dhaka with the capability to test wells for arsenic, but few rural households are aware of these services.[3] The cost of well testing is greatly reduced and the logistics are greatly simplified by the use of field kits, which have become increasingly reliable and easy to use (George et al. 2012; van Geen et al. 2014), but even these tests are rarely available in the villages.

In the context of this study, only about one in four households purchased a test, regardless of the offer type, despite the low—and subsidized—sale price and widespread awareness about the arsenic problem, and despite little prior awareness about the safety status of individual wells. This is consistent with a growing literature that documents low demand for health-protecting technologies in developing countries for a variety of such products, ranging from insecticide-treated nets (Tarozzi et al. 2009; Cohen and Dupas 2010; Dupas 2014; Tarozzi et al. 2014) to de-worming drugs (Kremer and Miguel 2007) and water disinfectant (Ashraf, Berry, and Shapiro 2010). However, while the offer type barely affected demand, it did affect how households responded to the information. intent-to-treat (ITT) estimates (*not* conditional on purchase) show that while standard individual sales led 3.7 percent of households to switch water source, the fraction was 4.3 percent with group sales (14 percent higher, 95 percent C.I. [−0.0137, 0.0339], *p*-value 0.404) and 6.4 percent (75 percent higher, 95 percent C.I. [0.003, 0.072], *p*-value 0.031) when metal plates were attached to the well spout in the case of purchase, although the difference is only statistically significant in the latter case. In addition, the estimates show very substantial differences between arms in the response of households that receive "bad news" about the safety of their well water. Among households informed that their well is high in arsenic, switching rates almost doubled from 30 to 56 percent with group sales relative to standard individual sales (95 percent C.I. of the difference adjusted for baseline covariates is [0.011, 0.362]). Switching rates also more than doubled when metal plates were attached to the pump head (from 30 to 72 percent, 95 percent C.I. of adjusted difference [0.188, 0.609]).

This work complements the literature on the demand for health-protecting technologies by looking at demand for health-related *information* that can be exploited by households to reduce risk. The focus here is on the offer of information that is specific to the buyer (the test measures arsenic contamination in the water from a specific well), in contrast to general information (for instance, on the likelihood of arsenic contamination, or the health risks associated with unsafe water). While this article studies demand for information on environmental factors, earlier work has considered the demand for information on health

---

[3] In addition, the cost of the laboratory analysis is as high as USD 25–40, not including the cost of the kits necessary for the collection of the water sample.

status; see in particular Thornton (2008), Cohen, Dupas, and Schaner (2015), Bai et al. (2017), and Gong (2015). These studies suggest that even among households willing to pay for information, behavioral responses may not be optimal from a public health perspective, so that it is important to study whether the mode of delivery of information can help in achieving desirable policy objectives.

This article is related to Barnwal et al. (2017) who estimate a very steep demand curve for arsenic tests in Bihar, India, another location with a groundwater arsenic problem. This study found that uptake was 25 percent at a price of INR 40 (about BDT 49), which is about the same as this study estimates at a very similar price of BDT 45. Unlike Barnwal et al. (2017), this study does not analyze how demand changes with price, but it examines the role of non-price factors on demand and behavioral responses to information. The results show that demand was not sensitive to the introduction of informal agreements or the use of placards, but conditional on demand, these nudges led to large and significant increases in switching among users of unsafe wells relative to simpler, private sales.

The paper proceeds as follows. The next section provides additional background information on the extent of the arsenic problem in the study area and describe the experimental design. Section 3 describes the data collection protocol, presents selected summary statistics, and shows that by chance the means of some covariates were not balanced at baseline, highlighting the importance of controlling for baseline characteristics in our estimates (the adjusted and unadjusted estimates remain qualitatively similar). Section 4 presents the conceptual framework that guided the study design and the interpretation of the results provided in Section 5. The cost effectiveness of the interventions is evaluated in Section 6. Finally, Section 7 concludes and highlights limitations of the results.

2. Program Description and Study Design

This study was carried out in Sonargaon, a 171 km^2 sub-administrative unit (or *upazila*) of the Narayanganj district, located approximately 25 km south-east of the capital Dhaka. According to the 2011 Census of Bangladesh, Sonargaon had a population of about 400,000, and at the time of the study it included 365 villages. According to the BAMWSP blanket testing, conducted locally in 1999–2000, about 40 percent of villages in Sonargaon contained >90 percent of wells that did not meet the national standard of 50 ppb.[4]

2.1. Study Area and Program Description

The study area for this trial was initially formed by the list of all 128 villages in Sonargaon with more than 10 wells and with a 40–90 percent share of unsafe wells based on the BAMWSP testing conducted years earlier. A lower bound was chosen to focus on areas where a sizeable fraction of new untested wells were likely to be unsafe, while the upper bound was designed to avoid areas where switching to safe wells was not likely to be a viable option for most households.

Between January and June 2016, surveyors conducted home visits to identify all wells in the selected villages, regardless of whether they had been tested before. Privately owned wells were linked to the household that owned them, while public wells were linked to the main caretaker or user. Almost all wells (98.6 percent) were privately owned, and for simplicity in the rest of the paper the term “owner” will be used to refer to the household that owned the well, or to the household that was the primary user of a community well.

During the home visits, surveyors explained to an adult—typically the most senior woman—the risk of consuming arsenic-contaminated tube-well water and offered to test the well for a fee. When a test was purchased, tube-well water was tested in the field using the Arsenic Econo-Quick (EQ) test kit, which has been shown to be reliable, and can deliver results within ten minutes (see George et al. 2012 for details).

[4] Blanket testing in Sonargaon was carried out by BRAC, a partner NGO of BAMWSP. A total of 25,048 tube wells were tested for arsenic.

The result was immediately communicated to the buyer. The kit's test strip is evaluated visually and the result classified in the following sequence (in ppb) {0, 10, 25, 50, 100, 200, 300, 500, 1000}. The tests cost USD 0.30 for volume purchases, although the total cost per test is higher, at about USD 2.40 per test, based on a testing campaign that also covered the costs of trained personnel and metal placards displaying the test result and attached to the body of the hand pump (van Geen et al. 2014).

Surveyors also administered a short household questionnaire and distributed color-coded laminated cards with the hand-written test result (in case of purchase) and well identification number. All cards included the following information: (a) that arsenicosis is not a communicable disease, (b) that arsenic cannot be removed by boiling water, (c) that testing tube-well water for arsenic is important, and (d) that the Bangladesh safety standard for arsenic concentration in water is 50 ppb. Black cards with these messages were given to households that did not buy a test, while in the case of purchase the laminated card was blue if arsenic was up to 10 ppb, green if 25–50 ppb, and red if >50 ppb (i.e., unsafe according to the national standard). Owners of unsafe (red) wells were encouraged face-to-face to switch to a safe (blue or green) well, while owners of wells with concentrations up to 10 ppb were encouraged to share their well water with their neighbors. Owners of green wells were encouraged both to share their water and to switch to a safer (blue) well, if possible.

Experimental variation came from differences between three ways of selling arsenic tests across villages. In a first group of villages, referred to as group A, surveyors offered to test tube-well water for a fee of BDT 45 (about USD 0.60). This fee was expected to cover the salary of the testers and their supervisor. Of the BDT 45 charged per test, testers kept BDT 30 to cover their transportation expenses and salary, and handed over the remaining BDT 15 to their supervisor. The price was determined by assuming that a field worker would test about 15 wells/day for 20 days/month, leading to a monthly salary of BDT 9,000 (USD 112.50/month), which is roughly what village health workers were paid for blanket testing in the neighboring Araihazar in 2012–2013 (van Geen et al. 2014). According to the same scenario, a supervisor of 10–15 workers would earn BDT 45,000–67,500 (USD 563–844), a range that spans a supervisor's earnings during the testing in Araihazar in 2012–2013. Across all experimental arms, the cost of the field kits (USD 0.30/test) was covered by the project.[5]

In a second group of villages (B), surveyors were asked to sell the tests to *groups* of buyers, rather than to individual households. When a well owner was identified, surveyors would propose the formation of a group of buyers of at least 3 and up to 10 households, while individual sales were not allowed.[6] Surveyors would help group formation, for instance by proposing a sale to all households within the same compound (or *bari*), and then coordinating the inclusion of additional buyers via mobile phones. After a group was formed and an informal well-sharing agreement was reached by all group members, each household would pay BDT 45. The agreement had no legal standing, and was meant to serve as a soft commitment device. Our study design called for an agreement in writing, but in practice most buyers were uncomfortable about signing a document, so in a large majority of cases a verbal agreement took place instead. All members were informed of the test results for all wells within the group.

In a third group of villages (C), households were again assigned to receive individual test offers at BDT 45 (as in group A), but in the case of purchase a color-coded stainless steel placard was attached to the well's pump head. Placards displayed both in text and in color whether the arsenic concentration was up to 10 ppb (blue), 25 to 50 ppb (green), or above 50 ppb (red). As shown in fig. 1, the placards displayed two hands holding drinking cups, one hand holding a drinking cup, or a large cross over a hand holding a drinking cup, depending on the arsenic concentration.

---

[5] In practice, demand (and thus testers' compensation) was lower than expected; see footnote 15.

[6] Most groups included 7–10 buyers, although some had as few as 3 and some had 11.

Figure 1. Metal Placards

Source: Authors' photograph.

Note: The three pictures show examples of the stainless steel placards that, in the case of a test being purchased, were attached to tube-well spouts in arm C. The pictures show placards attached to tested wells that were found to be, from left to right, safe (blue, $As \leq 10$ ppb), marginally safe (green, $10 < As \leq 50$ ppb), and unsafe (red, $As > 50$ ppb), respectively.

The split of the test fee between the tester and supervisor in groups B and C was the same as in group A, but in B the project gave an additional bonus of BDT 12 per sale to testers to compensate for the additional effort necessary to coordinate group formation.[7]

2.2. Power Calculation and Study Design

The trial was not registered, and was not based on a pre-analysis plan. Comparison of demand for and responses to tests between arms A and B was the primary objective of the study. Data from earlier work in the neighboring Araihazar subdistrict (see Bennear et al. 2013) were used to estimate an intra-village correlation of $\rho = 0.09$ in switching decisions among owners of unsafe wells and a switching rate of about 30 percent, a rate at the lower end of what was observed in previous studies. It was determined that 50 clusters (villages) would be sufficient to detect meaningful impacts with high probability. In particular, assuming with individual sales (arm A) a 30 percent switching rate from unsafe wells, and even using a conservative 0.12 intra-village correlation in switching rates, and with only 5 unsafe wells identified through the test sales per village, 49 clusters per arm would have been necessary to have 80 percent power to detect a 15 percentage points difference in switching rates between arms for a two-sided test of equality at the 5 percent significance level (effect size $= 0.15/\sqrt{0.3 \times 0.7} = 0.33$). The same number of clusters would have achieved the same power for a smaller 10 percentage points difference (effect size $= 0.10/\sqrt{0.3 \times 0.7} = 0.22$) with a less conservative $\rho = 0.09$ and with higher demand leading to the discovery of 20 unsafe wells per cluster, while 51 clusters per arm and 40 unsafe wells per village would have been necessary with the more conservative $\rho = 0.12$.

---

[7] The experimental design also included two exploratory arms, with only 6 villages each, where tests were sold individually at a village-level price of either BDT 45 or BDT 90, but with payment required only in the case of "good news," that is, in the case of an arsenic level no higher than 50 ppb. The inclusion of these proof-of-concept conditional sales was motivated by the aversion expressed by several members of focus groups to the idea of "paying for bad news." Sales conditional on the results may have thus increased demand (a prediction strongly supported by the observed purchase rates), although the conditional payment also generates a reduction in the (expected) price and a different selection into purchase conditional on beliefs about the safety of the well water. Because of these confounding factors and because of the very small number of villages assigned to these sales, the results are not discussed in detail in this article but they are available upon request from the authors.

Given that the available funding allowed the inclusion of a larger number of villages, it was possible to include the additional experimental arm C, for which, however, the sample size was dictated by budget constraints rather than power calculations.

The assignment to treatment arms was done by the principal investigators using random assignment, using the statistical software Stata, after stratification. Strata were determined by whether the share of unsafe wells in the BAMWSP testing campaign carried out years earlier was below or above the median, and by union (an administrative unit).[8] There were two deviations from the experimental protocol. First, while programming the mobile application used for data collection, 27 villages were assigned by mistake to a treatment different from the original one. The partial reassignment of treatments was thus due to a programming error and not to the incorrect implementation of the protocol in the field. In addition, the checks for balance in covariates are very similar based on originally assigned or actual treatment (see below). For this reason, treatment status is defined as *actual* treatment in the analysis. The second deviation is that surveyors were unable to differentiate a village from the one adjacent to it in four cases. While data were collected from households in these four villages and the ones adjacent to them, only pairs of villages could be distinguished, and both villages in each pair received the same treatment. In the statistical analysis there are thus effectively 112 clusters divided into experimental arms A (49 clusters), B (48), and C (15). For simplicity, in the rest of the paper these clusters will be referred to as "villages."[9]

It should be emphasized that, while surveyors completed a census of *wells* in study villages, our data do not represent a census of *households*. The choice to survey only owners—who were anyway a majority—was due to budgetary constraints, but an implication is that one cannot study whether the choice of the primary source of drinking water changed also among non-owners.

3. Data

A team of testers who had at least completed secondary education was recruited locally. During the home visits when sale offers were made, testers also administered a short household baseline survey and recorded information on sales and, in the case of purchase, the result of the test. Additionally, surveyors recorded GPS coordinates of all wells and noted down whether there were any visible labels attached to the well indicating the status with respect to arsenic. The baseline questionnaire included a household roster, basic questions on socioeconomic status, detailed questions on the well, and a number of questions related to knowledge and practice in relation to drinking water and arsenic risk. Testers also recorded whether, according to the respondent, the arsenic status of the well water was safe, unsafe, or unknown. Information about a total of 12,606 wells was recorded and the household survey was completed for all but three well owners.

A limitation of the data is that, in the case of group sales in arm B, surveyors did not keep accurate records of who belonged to which group. In other words, while the data indicate for which wells a test was purchased, one cannot study the characteristics of buyers belonging to the same group, or to what extent well sharing was actually taking place *within each group* as a consequence of the test results.

The endline survey was completed between August 2016 and January 2017.[10] The average time between baseline and endline surveys was 7.7 months, and 86 percent of households had their follow-up interview between 7 and 9 months after the baseline interview. During the endline survey, surveyors were instructed to return to the wells identified during the test sales and record whether the household owning

---

[8] Unions are the third smallest administrative unit in Bangladesh, and are formed by several *mouzas*, which in turn are composed of 2 or 3 villages. The 128 study villages belong to 9 different unions.

[9] The supplementary online appendix shows that the randomization led to large spatial variation in treatments; see fig. S1.1.

[10] Unlike the baseline survey, where the wages of surveyors and the supervisor were covered mainly from test fees, the cost of the follow-up survey was paid for by the project.

the well was still using it as a primary source of water for cooking and drinking. In the case of a negative answer, the respondent was asked to accompany the surveyor to the actual source and would record the new GPS location and the presence of any visible indicators of arsenic status (for instance, one of the metal placards distributed during our intervention). The surveyor would also ask the respondent about the perceived safety of the new source, as well as about the primary reason for switching to a different source. Switching behavior was thus self-reported, but earlier work in the neighboring Araihazar subdistrict found that switching behavior recorded in a similar way to this study was actually consistent with urinary arsenic concentrations, an objective biomarker of exposure (Chen et al. 2007). Unfortunately, the records on the locations of the new sources of drinking water do not allow precise measurements of the extent to which switching was associated with a reduction in arsenic risk. The smartphone's GPS sensor—with a precision of 10 m at best—cannot uniquely identify a specific well among those surveyed due to the density of wells within a typical village in Bangladesh. The GPS data was still used to estimate the distance from the old to the new well used for drinking.

3.1. Summary Statistics at Baseline

Table 1 shows selected summary statistics measured at baseline. Throughout this paper, unless otherwise noted, the analysis is restricted to the large majority of households (91 percent) that used their own well at baseline, as this is the sample for which baseline water source and post-intervention switching can be determined. The baseline source was not recorded for non-users, and during the endline survey they were only asked again whether they used their own well for drinking and cooking.[11] All summary statistics except those on the first row of table 1 are thus calculated for households that used water from their own well for drinking and cooking.

Household heads had low levels of educational attainment on average, and most households were poor, with only 17 percent of the houses having a concrete roof (an indicator of wealth), while the rest had tin or (in rare cases) mud roofs.

The average well owner in our study area lived in a village where 75 percent of the wells tested by BAMWSP between 1999 and 2000 were unsafe with respect to arsenic. Despite the BAMWSP blanket testing campaign, a large majority of respondents (76 percent) did not know whether their well was safe or unsafe with respect to arsenic. Only 7 percent of them thought that their well was unsafe, while the remaining 17 percent reported believing that their well was safe. Although about a quarter of respondents indicated that they knew the status of their well, more than 99 percent of wells lacked any visible sign of their status with respect to arsenic, such as red or green paint. Information on the safety of the minority of wells that had been tested was thus not immediately observable by other households, although in principle knowledge could have been shared with others privately. Using geographic coordinates, we estimate that before our pay-for-test campaign the average well owner had about 0.02 wells labeled as safe within 50 m, out of an average of nearly 12 wells within that distance.

The immense public health challenge due to widespread arsenic contamination of well water has been widely discussed and advertised in Bangladesh, and this is reflected in the data. Almost all respondents replied yes to the question "Have you ever heard about arsenic in tube well water?" Similarly, all but a handful of respondents replied yes when asked "Are you aware of the health risks of drinking tube well water containing arsenic?"

On average, wells were shallow (179 ft., or 55 m) and about nine years old—consistent with many wells having been installed after the BAMWSP blanket testing. The average reported installation cost of wells in the sample was BDT 7,560, or about USD 100 (USD 323 using the PPP exchange rate from

---

[11] Wells not used for drinking were on average significantly shallower, and thus more likely to be contaminated with high levels of arsenic. Of the 1,193 wells not used for drinking, only 14 (about 1 percent) were believed to be safe by the owner.

Table 1. Baseline Summary Statistics and Balance across Treatment Arms

		Means by experimental arm			Tests of equality (*p*-values)
					H_0
	Obs. (1)	A (2)	B (3)	C (4)	A=B=C (5)
Drink from well at baseline	12,603	0.930	0.884	0.891	0.2520
Household head is male	11,410	0.843	0.850	0.860	0.9210
Household head wage worker	10,890	0.247	0.298	0.378	0.4260
Household head self-employed	10,890	0.440	0.438	0.285	0.3390
Household head no schooling	10,890	0.262	0.163	0.054	$<0.001^{***}$
Household head primary school	10,890	0.317	0.343	0.319	0.7900
Heard about As in well water	11,410	0.996	0.996	0.997	0.7430
Aware of health risks of As	11,410	0.997	0.998	0.999	0.2950
House has concrete roof	11,252	0.175	0.184	0.133	0.4060
Household members	11,045	3.630	3.570	3.590	0.8270
Number of children	11,045	1.480	1.430	1.470	0.6920
Well As status unknown (belief)	10,515	0.684	0.789	0.903	0.0005^{***}
Well As status unsafe (belief)	10,515	0.097	0.042	0.057	0.2590
Well As status safe (belief)	10,515	0.220	0.168	0.041	$<0.001^{***}$
Well labeled safe	10,515	0.003	0.001	0.001	0.1020
Wells within 50 m	10,260	10.400	14.500	12.100	0.2530
Wells within 50 m labeled safe	10,260	0.028	0.003	0.007	0.0302^{**}
Share unsafe wells (BAMWSP)	11,410	0.759	0.720	0.782	0.4120
Well is privately owned	11,410	0.980	0.991	0.992	0.1690
Well depth (×100 ft.)	11,410	1.830	1.770	1.730	0.8800
Well age (years)	11,410	8.860	9.490	9.000	0.0105^{**}
Well cost (×10,000 BDT)	11,410	0.763	0.764	0.707	0.7340
Persons drinking from well	11,343	8.880	8.490	9.750	0.5710
Attrition	11,410	0.094	0.089	0.063	0.3150
Lost after baseline	11,410	0.058	0.055	0.045	0.5990
Duplicate I.D. at baseline	11,410	0.036	0.033	0.018	0.4170

Source: Authors' calculations from baseline data (January to June 2016).

Note: The unit of observation is the primary household attached to a specific well. The number of clusters (villages) in the three arms are 49 (arm A, $n = 5{,}550$ wells), 48 (B, $n = 5{,}314$) and 15 (C, $n = 1{,}739$). Except for the first variable ("Drinks from well at baseline"), all variables are summarized for households that used the specific well for cooking and drinking at baseline. Differences in the numbers of observations across these variables are explained by missing entries during the data collection. The *p*-values in column (5) are for tests of the null of equal means across treatment arms (robust to intra-village correlation). Asterisks denote test significance: $^{***}p < .01$, $^{**}\ p < .05$, $^{*}\ p < .1$. Table S1.1 in the supplementary online appendix shows the detailed results for the normalized differences described in the text.

World Bank 2015). Well depth is a key predictor of installation costs: in the data, the elasticity of cost with respect to depth is 0.72 (s.e. 0.04). The BDT 45 price charged for the test in this study thus represented slightly more than one-half of one percent of the installation cost.

Well sharing was already common in the study area: while the average household had fewer than 4 members, the average number of individuals using water from a well for drinking was 8.8, and in more than half of the sample wells the number of users was larger than household size.[12]

3.2. Variable Balance across Experimental Arms

Column (5) of table 1 shows the *p*-value for the null hypothesis of equality of means across the three treatment arms. The null is rejected in 5 of 26 cases at the 10 percent level. The differences are due to

---

[12] Recall that households that did not own a well were not surveyed.

chance, and because baseline data were collected while offering wells tests, balance on variables measured at baseline could not be enforced through stratification or re-randomization. Some of the differences are substantively important. While overall 19 percent of household heads in the study area had no schooling, this number drops to 5 percent in treatment C and is close to 26 percent in arm A. The fraction of respondents who did not know the status of their well ranged from 68 percent in arm A to 90 percent in arm C, while the fraction of wells perceived as safe ranged from 4 percent in C to 22 percent in A. Both the group-specific means and the tests of significance are very similar if the estimation is repeated using treatment as initially randomized rather than actual treatment, sometimes not identical due to errors in coding the smartphones used for surveying.[13]

The overlap in the distribution of covariates between arms can be gauged more systematically following the approach described in Imbens and Rubin (2015, Ch. 14). First, for each covariate in table 1, differences between means for each pair of experimental arms normalized by a measure of variance are estimated; see the supplementary online appendix for details. While the usual t-statistics used to construct the tests for balance have a denominator that shrinks to zero when sample size grows large (because the standard *errors* become smaller), this is not the case for the normalized differences, where the denominator is the simple average of two arm-specific standard *deviations*. Imbens and Rubin argue that these latter statistics are more relevant than the t-statistics for assessing whether simple adjustment methods, such as controlling for covariates or matching estimators, can adequately remove bias due to covariance imbalance. One can also calculate, for each pair of arms, a “multivariate difference” estimated with a Mahalanobis distance that aggregates all the individual differences. Although Imbens and Rubin do not propose formal tests based on these statistics to gauge balance, they argue that balance is excellent in an empirical illustration where all standardized differences are smaller than 0.3 and the multivariate measure is 0.44. In contrast, simple regression adjustments are deemed likely to be inadequate to eliminate bias in cases where some standardized differences are larger than 0.50 and the multivariate measure is 1.5 or above.

The normalized differences, reported in supplementary online appendix table S1.1, show that overall there is good balance between arms A and B: there is no variable for which the standardized difference is larger than 0.3, and the aggregate measure of balance is 0.604. In contrast, lack of balance is more problematic when one compares either arm A or arm B to arm C. In comparing A and C, consistent with the formal tests of equality, the differences are particularly large for schooling of the head and beliefs about well safety, with standardized differences larger than 0.5 in absolute value. The multivariate difference is also relatively large and equal to 1.1. The comparisons between B and C also show that 4 of 22 standardized differences are larger than 0.3, with a multivariate difference equal to 0.720.

Because there is a lack of balance in some characteristics such as beliefs or schooling that may affect behavior, results that control for observed covariates will also be analyzed. The estimates are qualitatively robust to such inclusion, although the point estimates are in some cases affected, and the standardized differences described above suggest that some caution should be exercised, in particular when making comparisons that involve group C.

Attrition is analyzed at the bottom of table 1. Overall, 8.8 percent of households could not be matched to the endline data, either because of true attrition (6 percentage points) or because errors in identifiers—which appear as duplicates in the data—did not allow the match. The null of equality among the three arms is not rejected at conventional levels for any of the attrition measures.

---

[13] In terms of statistical significance, the only differences are that “wells within 50 m labeled safe” becomes significant at the 10 percent level (instead of 5) and the fraction of privately owned wells become significant at the 1 percent level, despite very similar means across arms, ranging from 97.8 to 99.7 percent. The full results are available upon request from the authors.

4. Conceptual Framework

Before discussing the results, it is useful to consider the main factors likely to influence purchase choices and, conditional on test results, risk-mitigating behavior. This section does not describe a formal model but rather offers a simple conceptual framework to interpret the results in terms of expected differences between experimental arms and in terms of mechanisms.

Willingness to pay for a test is likely to require the existence of three conditions: first, that the test provides new information; second, the perception that there are health and/or economic costs associated with continued use of arsenic-contaminated water; third, that in the case of "bad news" there will be mitigation strategies available (e.g., the possibility of switching to a nearby safer well). All these conditions were present in the empirical context of this study.

The first condition is satisfied by the large majority of respondents (76 percent) who did not know whether their well water was safe or unsafe to drink (table 1). In addition, very few wells had visible signs of safety status such as paint, so even households with strong priors about the safety of their well water may have valued the possibility to demonstrate water quality to others by displaying test results.[14] That the tests could provide new information also required trust, as there is growing evidence that lack of trust in health-related information may hinder the adoption of behavior that could reduce health risks (Cohen, Dupas, and Schaner 2015; Bennett, Naqvi, and Schmidt 2018; Alsan and Wanamaker 2017; Martinez-Bravo and Stegmann 2019). Although the data do not include measures of trust, water tests were not a novelty because many wells had been tested in the past in Sonargaon. In addition, earlier work carried out in the neighboring Araihazar subdistrict found switching rates from unsafe to safe wells after testing of between one-third and three-quarters (see the Introduction for references), consistent with a high degree of trust in the results.

The second condition (relevance of the information) is also satisfied because virtually all respondents knew in a general sense about the presence and health risks of arsenic in tube-well water. Data on risk perceptions collected in neighboring Araihazar in 2008 indicated that a majority of respondents were aware not only of the serious nature of arsenic risk, but also that the risk becomes more severe with prolonged exposure; see Tarozzi (2016, fig. 4) for details.

Finally, the third condition (perceived availability of alternative sources of drinking water) was also likely to be satisfied given that well sharing was already practiced in the area (see the Data section). To some extent, households were therefore already aware of the possibility of using neighbors' well water for drinking. Data from neighboring Araihazar show that the BAMWSP testing campaign conducted about 10 years earlier led to a substantial degree of well sharing (Balasubramanya et al. 2014). On the one hand, the figures in table 1 show that about three-quarters of wells had been found to be unsafe by BAMWSP in the study area, and this may have reduced the perceived chance of having safe options nearby in the case of bad news. On the other hand, the average household had more than 10 other wells within a short 50 m radius and about 30 within 100 m, and such a dense network of wells is likely to have increased the perceived likelihood of having safe wells close by.

Overall, differences in switching rates between any two experimental arms could have emerged either from different selection into purchase or from the way information was provided (in which case even identical buyers may have reacted differently). The decision to change source was still likely to depend primarily on any new information made available by the test results. Hence, regardless of the experimental

---

[14] In principle, such value does not need to be positive, for instance if knowledge of a high-arsenic well is perceived as lowering land value, or signaling poor health among household members, or is more generally stigmatized. In Bihar, India, Barnwal et al. (2017) find that placards indicating unsafe arsenic levels were, two years after installation, more likely to have been removed by households than those indicating low levels of arsenic, although such behavior may have also been justified by the desire not to be reminded constantly about the health risks. However, earlier research has shown that households very rarely refuse testing when this is offered for free, even when the results are posted on the wells (see for instance Madajewicz et al. 2007; Opar et al. 2007; Bennear et al. 2013).

arm, very little switching was expected to take place from untested wells (driven perhaps by "free riders" who moved to wells nearby found to be safe), and even less from wells that were found to be safe. The prior was also that the difficulty in predicting arsenic contamination without a test would mean that the likelihood of having an unsafe well, even if conditional on purchase, would be uncorrelated with the mode of sale and thus similar between groups. Finally, conditional on finding out that one's water is unsafe, the expectation was that the soft commitment and the easier access to within-group information on safe alternatives (in group B), and the salience and visibility of the tags posted on wells (in group C), would lead a larger proportion of households to stop drinking from the tested well relative to the individual sales in group A. In contrast, priors about the switching rates in B relative to C were not as clear. However, the effectiveness of C at inducing switching from unsafe wells rested at least in part on households not removing the metal plates from the pump heads, thereby maintaining the ability of the plate to make safety status visible and to discourage drinking from unsafe wells.

Regardless of the sale method, switching from unsafe wells should be more likely when safer alternatives are available nearby. If households recognize that health risks increase with the arsenic concentration in the water, then switching from unsafe wells should increase with the arsenic level, a desirable pattern that has been observed elsewhere; see Madajewicz et al. (2007). The method of information delivery may also affect the choice of the new source conditional on switching. In particular, the labeling of wells in C could make safe wells easier to identify for the whole village relative to B, possibly leading to the choice of safer wells. Finally, it is possible that buyers that were *not* using their well for this purpose at the time of the sales started doing so if they learned that the water was safe.

Predictions for differences in *demand* between arms were not as sharp: while factors leading to higher willingness to react to information were expected to also lead to higher willingness to pay for it, key factors such as perceived health risks and availability of alternative sources were likely to become more salient after the realization of the test result. However, consistent with the conceptual framework described above, households that reported knowing the safety of their well water were expected to be less likely to purchase a test.

5. Results

This section first describes the estimated effect of the selling schemes on the demand for testing. Next, it describe the information on arsenic levels that was revealed by the tests, and finally it discusses to what extent such information changed household behavior in terms of choice of water source for cooking and drinking. In describing the results the focus is primarily on households that used the well as primary water source for cooking and drinking at baseline, given that for those who were not the baseline records do not indicate what the main source was.

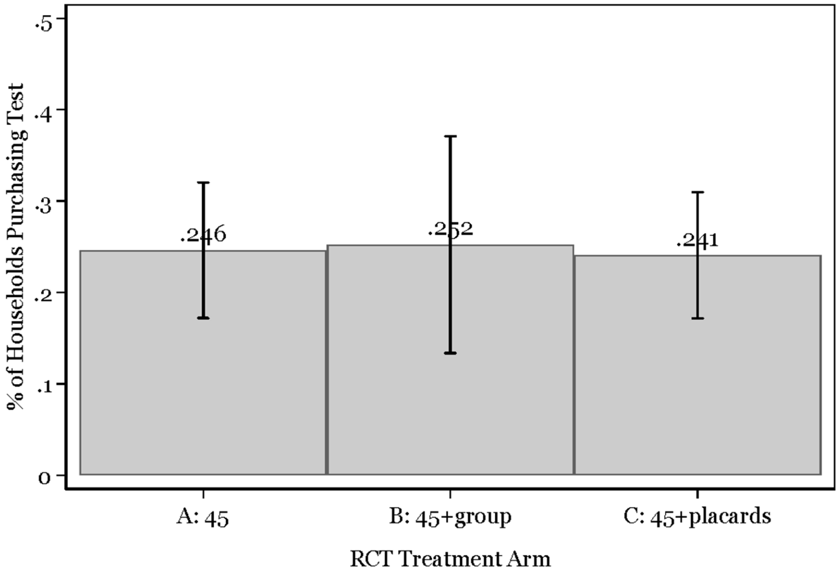
5.1. Demand

Of the 11,410 households that used their own well for cooking and drinking at baseline and who were offered an arsenic test, 2,829 (25 percent) bought a test under one of the selling schemes. To estimate the average treatment effect of selling schemes B and C relative to A, the following equation is estimated using a linear probability model:

$$\mathrm{buy}_{svh} = \beta^B B_v + \beta^C C_v + \gamma X_{svh} + \delta_s + \epsilon_{svh}, \tag{1}$$

where buy_{svh} is equal to 1 if household h in village v and stratum s bought a test at baseline and 0 otherwise, B_v and C_v are village-specific indicator variables for the respective treatments, X_{svh} is a set of predetermined household and tube-well characteristics, and ϵ_{svh} is an error term. To account for the stratified design, regressions also include strata fixed effects (δ_s). Recall that treatment was stratified by the prevalence of unsafe wells based on BAMWSP data and by union. The coefficients of interest are β^B and

Figure 2. Demand for Tests of Arsenic Concentration in Well Water



Source: Authors' estimations from baseline data (January to June 2016).

Note: Each bar is labeled with the arm-specific purchase rate. The vertical intervals represent 95 percent confidence intervals, estimated allowing for intra-village correlation of residuals. The number of observations by arm are, from left to right, $n = 5{,}164$ (A), 4,697 (B), and 1,549 (C).

β^C, which capture the causal impact on demand of selling schemes B and C, relative to A, respectively. All standard errors and statistical inference are robust to the presence of intra-village correlation of residuals.

Figure 2 shows graphically the simple comparison of take-up rates across arms without the inclusion of controls or strata fixed effects. A first clear result is that neither the group sales nor the addition of the metal placard made any appreciable difference for demand. A second finding is that demand was overall quite low, with about one-quarter of households purchasing the test in each of the three experimental arms. As in many earlier studies looking at demand for health-related preventive products, even a small fee led to low demand, despite the potentially vital information provided by the tests.[15]

Table 2 displays the corresponding regression results. Column (1) shows that, as expected, the small differences in demand between arms A, B, and C are not statistically significant. Column (2) shows that the results are quite robust to the inclusion of controls. Because missing values in one or more of the controls lead to the loss of about 20 percent of observations, in column (3) the model is estimated without controls but including only the observations with complete observations used in column (2). In this case, $\hat{\beta}^B$ barely changes, consistent with the overall good balance between arms A and B suggested by the Imbens and Rubin approach. In contrast, $\hat{\beta}^C$ doubles in magnitude from 3 to 6 percent (s.e. 0.034) and it becomes significant, although only at the 10 percent level. Recall that arm C appears to be different mostly because, on average and relative to the other arms, (a) household heads had better education and (b) the fraction of wells whose safety was unknown at the time of the test sales was higher and the fraction believed to be safe was lower. In column (2) it can be seen that low schooling predicts lower demand, while the higher prevalence of wells of unknown safety is positively associated with it, conditional on other observed

[15] The low demand also raises concerns on the sustainability of a test-for-fee selling scheme at this price. On an average work day, surveyors visited 25 well owners to offer As tests (with surveyor-specific averages ranging from 12 to 36 visits), but sold only 6 tests (with surveyor-specific averages ranging from 3.5 to 10.8 tests). Recall that the test price was chosen so that surveyors would earn a wage similar to that earned in the neighboring district of Araihazar for similar work, selling 15 tests a day. The low demand thus implies that the actual wage fell short of the expected one.

Table 2. Demand for Tests

	Dependent variable:[a] Indicator = 1 if household purchased test		
	(1)	(2)	(3)
B: 45+group	−0.004	−0.012	−0.001
	(0.037)	(0.018)	(0.018)
C: 45+placards	0.034	0.030	0.059^{*}
	(0.032)	(0.028)	(0.034)
Household head is male		-0.070^{***}	
		(0.018)	
Household head works for wage		0.005	
		(0.016)	
Household head self-employed		0.037^{**}	
		(0.015)	
Household head has no schooling		-0.126^{***}	
		(0.014)	
Household head has primary only		-0.066^{***}	
		(0.012)	
Concrete house roof		0.073^{***}	
		(0.015)	
No. household members besides children		0.039^{***}	
		(0.009)	
No. of children of head in household		0.052^{***}	
		(0.006)	
No. of wells within 50 m		-0.003^{***}	
		(0.001)	
No. of visibly safe wells within 50 m		0.006	
		(0.037)	
Fraction unsafe wells in village (BAMWSP)		0.093	
		(0.086)	
Well depth (×100 ft.)		0.033^{***}	
		(0.009)	
Well age (years)		−0.001	
		(0.001)	
Well cost (×10,000 BDT)		0.027^{*}	
		(0.014)	
Believes well is safe		-0.115^{***}	
		(0.018)	
Believes well is unsafe		-0.049^{***}	
		(0.015)	
Observations[b]	11,410	8,892	8,892
R-squared	0.130	0.111	0.040
Controls	No	Yes	No
Strata FE[c]	Yes	Yes	Yes
Mean in A	0.246	0.246	0.246
Clusters	112	102	102

Source: Authors' estimations from baseline data (January to June 2016).

Note: a) The dependent variable is binary and =1 if the household purchased the test at baseline. All regressions are estimated with OLS, and group A is the omitted category, as in model (1). Standard errors are clustered at the village level. Significance: *** $p < .01$, ** $p < .05$, * $p < .1$. b)The smaller sample size in column (2) relative to column (1) is due to missing values in one or more controls, while in column (3) controls are not included, but only observations with complete observations are used. c)Strata fixed effects include union fixed effects and a dummy = 1 in villages where the fraction of unsafe wells in the village (estimated by BAMWSP) was below the median.

Table 3. Number of Wells by Safety Status and Switching Decision

		Tested						Not tested	
				Safe		Unsafe			
	Total wells used for drinking and cooking	Total	Unsafe (proportion)	Switched		Switched		Switched	
				Yes	No	Yes	No	Yes	No
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A: BDT 45	4,679	1,040	0.192	2	838	60	140	113	3,526
B: BDT 45+group	4,281	1,029	0.155	9	860	89	71	85	3,167
C: BDT 45+placards	1,452	348	0.273	1	252	68	27	26	1,078
Total	10,412	2,417	0.188	12	1,950	217	238	224	7,771
Tests of equality (p-value)									
H_0: A=B=C			0.3174						

Source: Authors' calculations using information from a total of 10,412 wells that were used at baseline for drinking and cooking purposes.

Note: Excluded from the analysis are 768 wells used by households that could not be recontacted at endline, and 406 wells with a duplicate ID at baseline which can thus not be matched to endline data on switching decisions.

characteristics. Both these factors suggest that the omission of controls may have biased demand upwards in arm C, although the point estimates remain very close.

The coefficient estimates for the controls in column (2) cannot be interpreted causally, but it can be noted that beliefs about the safety status of the well strongly predict demand: well owners thinking that their well is safe had little to gain from buying a test, and indeed they were 12 percentage points less likely to purchase the test (p-value < 0.01). The belief that the water was *unsafe* also decreased the probability of purchase, although by less than half as much ($\hat{\beta} = -0.049$, p-value < 0.01). Overall this is consistent with the conceptual framework, where it was highlighted that a key factor for willingness to pay for the test is that its result will provide new information.

5.2. Test Results

Although the purchase rate was low, the intervention generated a large increase in the number of tested wells in Sonargaon. Before looking at the responses to the information made available by the tests, it is useful to describe such information. The test results are summarized in table 3, which also includes the detailed figures on switching behavior that will be described later, and so the statistics are calculated for the 10,412 households (91.3 percent of the total) that could be tracked in the endline survey. Of these, 2,417 purchased tests during the intervention.

Overall, 19 percent (455/2,417) of the tested wells that had been used for cooking and drinking at baseline had unsafe arsenic levels relative to the national standard of 50 ppb.[16] Recall that these results are *conditional on demand* so that the randomization across treatments does not guarantee similar distributions across arms, even in large samples. However, the distribution of arsenic was overall similar between arms A and B, the two largest arms. Arm C had more unsafe wells (27 percent, versus 19 and

---

[16] The prevalence of unsafe wells was much lower than the 40–90 percent observed at the time of the BAMWSP testing campaign, about 10 years earlier. This is consistent with a degree of learning over time about local arsenic risk and how to avoid it, in particular by digging deeper wells, which is more expensive but perhaps made more affordable by economic development. Indeed, the data indicate that a majority of wells were of recent construction and were deeper than the older ones. The data also show that the beliefs about the safety of their well water among the minority of respondents who reported to know it were strongly correlated with actual test results. Detailed results for these findings are available upon request from the authors.

16 percent in arms A and B, respectively), although the null of equality among these three arms cannot be rejected at standard levels (p-value = 0.32). Consistent with the existence of a degree of awareness about arsenic risk, at baseline group C was by far the one with the smallest fraction of respondents thinking that their well was safe, although the fraction believing that the well was unsafe was fairly similar between groups; see table 1. The larger share of unsafe wells in arm C may thus have been the result of lack of balance at baseline arising by chance, possibly due to the small number of clusters (15) in this treatment arm.

Overall, at the time of the endline survey, of all the wells found to be unsafe, 30 percent had at least one well identified as safe within 25 m after the testing, 57 percent had at least one within 50 m, and 78 percent had at least one within 100 m. This confirms that, in principle, switching to a nearby safe well was a feasible strategy to mitigate arsenic risk for the large majority of households. In addition, and consistent with the similarity across arms in the prevalence of purchases and unsafe results, the different testing strategies produced very similar frequencies of safe alternatives in the vicinity of high-arsenic wells. At distances of 25, 50, and 100 m, such frequencies ranged across arms from 27 to 33 percent, from 55 to 60 percent, and from 73 to 85 percent, respectively, and the null of equality is never rejected at standard levels. Note also that these figures may underestimate substantially the *potential* role of switching to reduce arsenic risk, given that they do not take into account the likely presence of safe wells that were not tested.

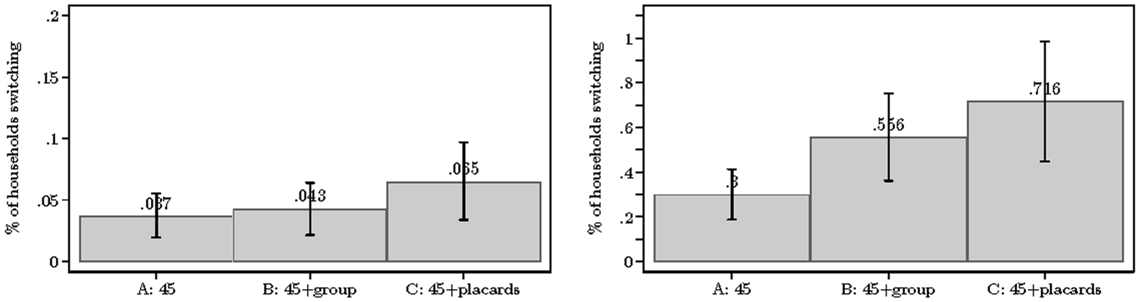
5.3. Responses to Test Results

Figure 3 shows the raw switching rates—and corresponding confidence intervals—observed in each experimental arm, without any control or strata fixed effects. Corresponding regression results are shown in table 4. Consistent with the conceptual framework, little well switching was found among households that did not buy a test, and among those who found that the well they used for drinking was safe. The estimates in table 3 show that barely anyone moved away from a well identified as safe (12/1,962), while less than 3 percent of untested wells (224/7,995) stopped being used for drinking. The two bottom bar charts in fig. 3 show that for these two groups the switching rates were similarly very small in all arms.

The primary outcome of interest in this study was the response of users of unsafe wells, but because such responses are conditional on the choice to purchase a test and on the test result, the ITT estimates are presented first, to describe unconditional switching rates. In fig. 3(1) it can be seen that while standard individual sales (A) led 3.7 percent of households to switch water source, the fraction was 4.3 percent with group sales (16 percent higher) and 6.4 percent (73 percent higher) when metal plates were attached to the well spout in the case of purchase. The regression results are displayed in columns (1)–(2) of table 4, where the estimated models are as in equation (1). When controls are included (column 2), the difference in switching rates between B and A is 0.01 (95 percent C.I. [−0.012, 0.03]), while the difference between C and A is 0.03 (95 percent C.I. [0.003, 0.065]). Consistent with fig. 3(1), both estimates suggest that group sales and especially placards increased switching rates relative to individual sales, but the point estimates are small, and the null of equality is only rejected—at the 5 percent level—for arm C. Overall, the unconditional switching rates were small in each arm, reducing the statistical power when making between-arm comparisons. This is in large part because about three-quarters of households did not purchase a test, and a large majority of wells were found to be safe.

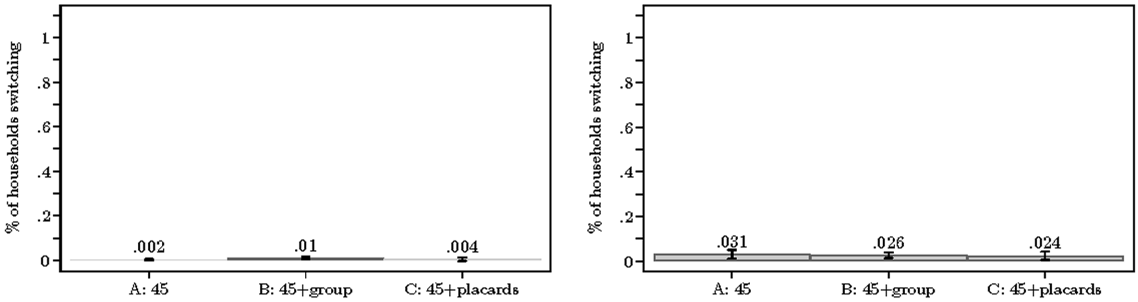
Turning next to the switching rates among households that purchased a test and found their well water to have unsafe arsenic levels, individual sales (arm A) led to a 30 percent switching rate (60/200), while both group sales (arm B) and metal placards posted on the pump head (arm C) led to substantially higher responses. In B, 56 percent of households switched (89/160), while in C the rate was even higher, at 72 percent (68/95). Column (3) of table 4 shows the corresponding regression results. Both estimated differences are large and significant at the 5 percent level or below, with $\hat{\beta}^B = 0.27$ (95 percent C.I. [0.061, 0.484]) and $\hat{\beta}^C = 0.46$ (95 percent C.I. [0.231, 0.696]). The difference in switching rates between B and

Figure 3. Switching Rates



(1) All wells (Intent-to-Treat)

(2) Test buyers, well unsafe (As above 50ppb)



(3) Test buyers, well safe (As 50ppb or below)

(4) Untested wells

Source: Authors' estimations from endline data (August 2016 to January 2017).

Note: Each figure shows the fraction of households that stopped using the baseline well for drinking and cooking and switched to a different water source, by experimental arm. Switching rates are shown separately for all wells (graph 1, top left), those that tested unsafe (2, top right), safe (3, bottom left), and for wells that were not tested because the test had not been purchased (4, bottom right). The vertical intervals within each bar are 95 percent confidence intervals robust to intra-village correlation. The number of wells n_T, $T \in \{A, B, C\}$ used in each bar are as follows. All wells: $n_A = 4{,}679$, $n_B = 4{,}281$, $n_C = 1{,}425$; unsafe wells: $n_A = 200$, $n_B = 160$, $n_C = 95$; safe wells: $n_A = 838$, $n_B = 869$, $n_C = 253$; untested wells: $n_A = 3{,}639$, $n_B = 3{,}252$, $n_C = 1{,}104$.

C is substantively important (19 percentage points) but is not estimated precisely (p-value=0.163). The estimated differences become smaller but remain large and significant when baseline controls are included (column 4). For both arms B and C the coefficients are almost identical (column 5) when the model does not include controls but it uses only the complete observations used in the regression with controls in column (4). This suggests that the impacts are not substantively biased by differences in the level of observed confounders.

Contrary to expectations, unsafe higher levels of arsenic do *not* predict more switching. Using 100 as the omitted category, dummies for the arsenic level being equal to 200, 300, or 500/1,000 ppb are actually *negative* and in some cases very large and statistically significant. This finding is consistent with most households gauging safety primarily in a binary way, an unfortunate possibility given that in reality arsenic health risk is to first order proportional to arsenic concentration.[17] Figure 4 shows indeed that overall switching rates were well approximated by a step function with a jump at 100 ppb. Note also that very high arsenic levels were not rare in the sample: although less than 30 percent of tested wells were unsafe, more than half of these had arsenic levels above 100 ppb.

The estimates in column (6) also include as regressor a dummy for the (endogenous) presence of a safe well within 50 m, where a neighboring well is defined as safe when it has been identified as such by the

---

[17] In an RCT carried out in 2008 in the neighboring Araihazar subdistrict, Bennear et al. (2013) showed that attempts to highlight the existence of such a gradient did not increase switching, with some evidence that it actually *decreased* it.

Table 4. Choice of Water Source at Endline

	Dependent variable:[a] Indicator = 1 if switched (no longer uses same well used at baseline)					
	All wells (ITT)[b]		Conditional on purchase and unsafe well (As > 50 ppb)[c]			
	(1)	(2)	(3)	(4)	(5)	(6)
B: BDT 45+group	0.010	0.009	0.273**	0.186**	0.191*	0.344***
	(0.012)	(0.010)	(0.106)	(0.088)	(0.103)	(0.097)
C: BDT 45+placards	0.038**	0.034**	0.463***	0.398***	0.390***	0.389***
	(0.017)	(0.016)	(0.117)	(0.106)	(0.121)	(0.121)
Household head is male		−0.020**		−0.063		−0.068
		(0.009)		(0.078)		(0.091)
Household head works for wage		0.002		−0.033		0.012
		(0.006)		(0.056)		(0.058)
Household head self-employed		0.029***		0.233***		0.208***
		(0.010)		(0.070)		(0.057)
Household head has no schooling		−0.007		0.076		0.151
		(0.009)		(0.097)		(0.118)
Household head has primary only		−0.009		0.031		0.027
		(0.007)		(0.059)		(0.062)
Concrete house roof		0.005		0.028		0.079
		(0.006)		(0.079)		(0.078)
No. household members besides children		−0.002		−0.052		−0.055
		(0.002)		(0.032)		(0.045)
No. of children of head in household		0.004**		0.013		0.013
		(0.002)		(0.028)		(0.030)
Well depth (×100 ft.)		−0.015**		0.011		−0.034
		(0.006)		(0.032)		(0.041)
Well age (years)		0.001		−0.004		−0.002
		(0.000)		(0.004)		(0.004)
Well cost (×10,000 BDT)		−0.016***		−0.090***		−0.078**
		(0.006)		(0.034)		(0.038)
Believes well is unsafe		0.029		0.060		0.028
		(0.020)		(0.081)		(0.076)
Believes well is safe		−0.022***		−0.113		−0.123
		(0.007)		(0.108)		(0.123)
As = 200 ppb				−0.039		0.012
				(0.062)		(0.066)
As = 300 ppb				−0.151**		−0.157**
				(0.061)		(0.066)
As = 500 or 1000 ppb				−0.148*		−0.114
				(0.077)		(0.084)
Number of wells within 50 m						−0.002
						(0.005)
There is at least one safe well within 50 m						0.270***
						(0.085)
B × at least one safe well within 50 m						−0.188
						(0.123)
C × at least one safe well within 50 m						−0.076
						(0.121)

Table 4. Continued

	Dependent variable:[a] Indicator = 1 if switched (no longer uses same well used at baseline)					
	All wells (ITT)[b]		Conditional on purchase and unsafe well (As > 50 ppb)[c]			
	(1)	(2)	(3)	(4)	(5)	(6)
Observations[d]	10,412	9,385	455	407	407	355
R-squared	0.012	0.028	0.169	0.261	0.176	0.298
Mean in A	0.0374	0.0374	0.300	0.300	0.300	0.300
Test of equality B = C, *p*-value	0.124	0.136	0.163	0.0676	0.140	0.728
Clusters	112	105	76	71	71	66

Source: Authors' estimations from baseline and endline data.

Note: [a)] All regressions are estimated using a linear probability model where the dependent variable is a dummy equal to 1 if the well was no longer used for cooking and drinking at endline. Standard errors are clustered at the village level. All regressions include strata fixed effects. Asterisks denote statistical significance: *** $p < .01$, ** $p < .05$, * $p < .1$. [b)] The intent-to-treat results in columns (1) and (2) show switching rates *not* conditional on purchase or test result, including all households that used the well at baseline and who could be matched between baseline and endline surveys. [c)] All regressions in columns (3)–(6) include only observations for which the well was used for cooking and drinking purposes at baseline, a test was purchased, and the test indicated unsafe levels of arsenic in the water (As > 50 ppb). [d)] In column (2) (relative to column 1), and in column (4) (relative to column 3) the decrease in the number of observations is due to missing values in controls, and in column (6) some additional observations are lost because the GPS locations were not recorded correctly. The model in column (5) is the same as in column (3) but uses only observations with complete data used in column (4).

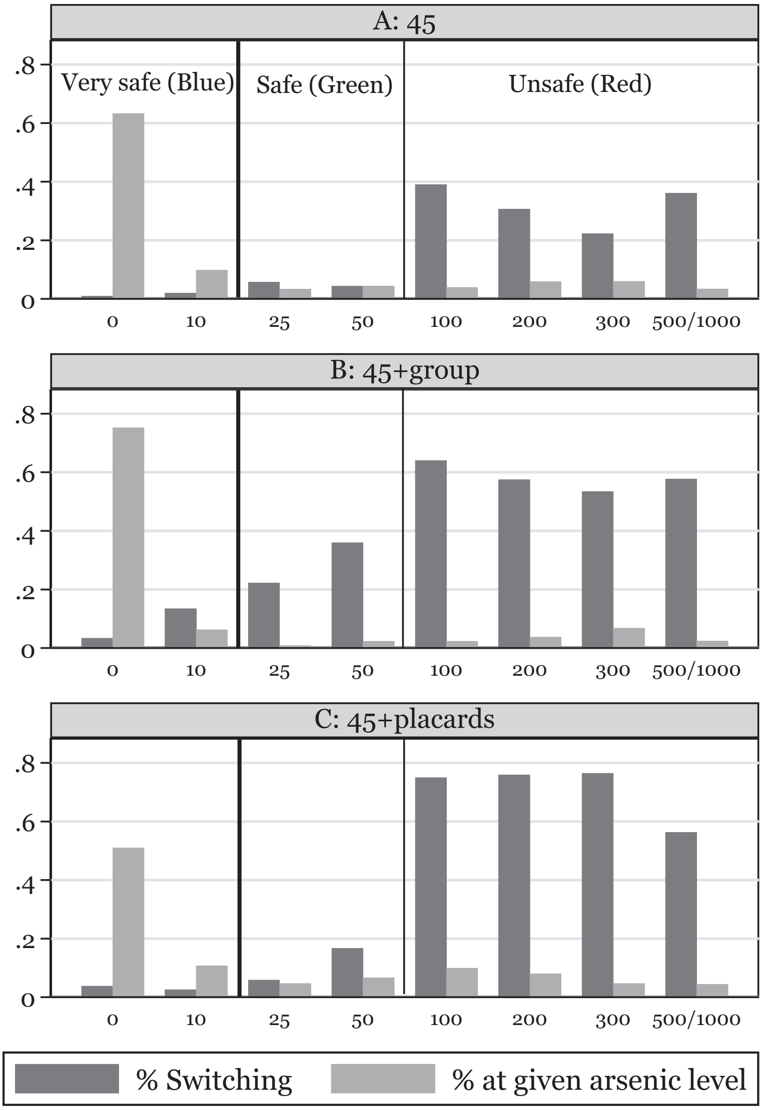
research team. Recall that, at baseline, very few wells could be identified as safe by visible signs such as placards or paint on the well spout. In this model some observations are lost due to implausible entries for the geo-location of some wells. Controls are also included for the total number of wells in a 50 m radius, and the dummy for safe wells is interacted with the treatment indicators. Among owners of unsafe wells in arm A, as expected, having a safe alternative nearby increases switching. The coefficient is large (27 percentage points) and significant at the 1 percent level. In group B, this association is weaker, given that the interaction (=−0.19) is *negative* and its magnitude is about two-thirds of that observed in arm A, although it is estimated imprecisely and is thus not significant at standard levels. This is consistent with informal group agreements leading some households to share wells with other members, with less concern for geographical distance, something which may have happened if geographical proximity was a poor proxy for sorting into the same risk-sharing group. This remains, however, a conjecture, given that the data do now allow determining with certainty whether the well being used at endline belonged to a group member. The interaction between distance to a safe well and the treatment C dummy is again negative (=−0.08) but smaller and not significant at standard levels.

Of the 217 "switchers," almost all (214) listed safety concerns as the primary reason for their decision. Although the arsenic level in the new source of drinking water for switchers cannot be determined, about one-third of these (79/217) had switched to a different well that was itself perceived by the respondent as being unsafe, while 88 had switched to a well reported as being safe, and the remaining 50 households did not know the status of the well. In principle, even a switch to an unsafe well, if the new well is *safer*, can reduce exposure to arsenic, but this finding suggests that in the study area a degree of arsenic exposure remained even among a sizeable fraction of households that reacted to the new information by switching to a different water source for drinking and cooking.

5.3.1. Mechanisms

These results confirm the conceptual framework, according to which group signing or metal placards would lead to more switching relative to privately provided information. This section provides evidence to support possible mechanisms behind the findings, although the arguments are tentative as it is not

Figure 4. Switching Rates from Tested Wells, by Arsenic Level and Experimental Arm



Source: Authors' estimations from baseline (January to June 2016) and endline (August 2016 to January 2017) data.

Note: The figures show, for each experimental arm and for tested wells, the prevalence of each arsenic level as identified by the test (light gray bars, in ppb, or micrograms per liter) and the fraction of households that were no longer using the well at endline (dark gray bars). The field tests identified the arsenic level as a value in the set {0, 10, 25, 50, 100, 200, 300, 500, 1000}. The values on the horizontal line are not drawn at scale. Results of As = 1000 were rare and hence 500 and 1000 were pooled together. Wells with arsenic below the thick vertical line were the safest, while those with arsenic above the second and thin vertical line were labeled unsafe. A household was described as having switched if, at the time of the endline survey, the respondent stated that the main source of water used for drinking and cooking was no longer the well used at baseline.

possible to separate conclusively the relative role of the increase in the information about alternatives versus the soft commitment (in arm B) or the added salience of the placards (in arm C). There are two key limitations. First, the data include respondents' beliefs about the well used by their household, and also include the arsenic levels of all tested wells, but the latter was not necessarily known to households. As a consequence, one cannot measure how each intervention changed the whole information set for each household. Second, in arm B the data do not include group composition. It is thus possible to examine neither the nature of the specific groups, nor whether households whose water turned out to be unsafe were being allowed to drink water from wells belonging to other members of the same group.

Despite these limitations, the data suggest that the added salience provided by the placards in arm C played a role in explaining the higher switching rates. In principle, owners who did not want to be reminded of their well water being unsafe, or did not want the information to be known to others, could have removed the placards, although detaching the metal wire holding them to the pump head would have required some effort. However, this behavior was rare. At the time of the return visits, the vast majority of the 348 placards installed on the well spouts at the time of the tests were still in place, regardless of their color. Of the 95 red placards installed on unsafe wells, 90 were still visible, while no placard was visible in two wells and a black placard (perhaps a data entry error) was found on the remaining three. Almost all blue and green placards remained similarly in place during the study period. This suggests that the testing campaign led to a persistent increase in the salience and visibility of information in villages included in arm C. This result stands in contrast with Barnwal et al. (2017), who found that placards indicating unsafe arsenic levels in Bihar, India, were significantly more likely to be removed by households, although such actions were observed two years after installation, a much longer time interval relative to the average of eight months in this study.

The data also indicate that the placards were more effective than result cards only (as in arm A) at reminding users of the arsenic status of their well water, again suggesting increased salience. At the time of the endline survey, almost 90 percent of buyers correctly reported whether the water was found to be unsafe, but while learning about unsafe water was similar in arms A and B, it appeared to be better in C, consistent with the role of placards as reminders. Among respondents whose well water was found to be unsafe, the fraction who correctly identified it as such was 83 percent in arm A (135/162), 88 percent in B (122/139), and a remarkable 98 percent in arm C (83/85).

There is also evidence that the placard allowed switchers to make better choices, while group sales, despite inducing more switching relative to individual sales, may have induced households to share wells within the group despite the existence of better options outside the group. Of the 217 households that stopped drinking from their unsafe well, only 88 (41 percent) had switched to a well that they believed to be safe. Looking at this by arm, the fraction was 47, 27, and 53 percent in arms A, B, and C, respectively. It is possible that switchers from unsafe wells in B started drinking water from wells *safer* than the original one (unfortunately it cannot be checked whether this was the case), but in C, the fact that switchers were almost twice as likely as in B to change to a safe well suggests that the placards played a role in allowing better choices. This is also consistent with data on the distance to the new well. Recall that when a respondent reported a change in the main source of water for drinking since baseline, the surveyor would ask to be accompanied to the new source, whose GPS location would then be recorded. Unfortunately such GPS records were clearly incorrect for about 40 percent of the 217 switchers (this was evident because the new source was located too far from the household residence, usually even outside the village borders), but when one looks at the 124 observations with records likely to be correct, while distances from the new source were almost identical in arms A and B (on average 68 and 80 m, respectively, with the p-value of the difference = 0.518), the distance was substantially larger in C (190 m, with the p-value of the difference with respect to A <0.01).[18]

---

[18] This finding also suggests that the results are not driven by courtesy bias in reporting switching behavior. In principle, safety information provided publicly (as in arm C) or to a group (as in B) may have induced some respondents to over-report switching behavior if this was perceived as socially desirable. If the higher switching rates in B and C relative to A had been driven by courtesy bias, one would have expected to observe shorter distances in the two former arms relative to A (the opposite of what the data show), given that the respondent had to accompany the interviewer to the new source of drinking water, and a well nearby is likely to have been an easier and faster choice if the only purpose was to back up cheap talk.

In sum, the data suggest that placards (C) are likely to have made households more aware of the risks associated with drinking from their own unsafe wells and to have allowed better choices, sometimes at the cost of longer distances traveled to fetch drinking water. In contrast, there is less that can be said about the mechanisms that made group sales (B) relatively successful, although the key factors delineated in the conceptual framework are consistent with the results.

5.3.2. Responses among Non-users

The results discussed so far are related to the large majority of households that used their well for cooking and drinking at baseline. Perhaps not surprisingly, demand was significantly lower among "non-users" (12 percent, vs. 25 percent among users), and for these households there is no record of their main source of water for drinking. Switching behavior is thus harder to analyze, also because the sample is small. However, our data allow us to determine that many of these households reacted to "good news" by switching to the well they were not initially using. In the sample, 139 non-users purchased a test, and of these 126 (91 percent) were re-interviewed at endline. Of these, exactly half (63) found out that their well was safe (As ≤ 50), and *all but* 3 reported that they *were* using the well for drinking and cooking at the time of the endline.[19] In contrast, only 9 of the 63 with unsafe wells reported that they were using the well. However, for them it is not known whether the well they were using at baseline was found to have an arsenic level even higher than their own.

6. Cost Effectiveness

This section evaluates the cost-effectiveness of the different sale strategies. Because the RCT did not include an arm with free provision (unlike earlier studies that *only* included free provision; see for instance Madajewicz et al. 2007 or Bennear et al. 2013), the merit of free provision as compared to our sales strategy is gauged by assuming a range of switching rates consistent with earlier studies.[20] In addition, recall that the change in arsenic contamination for "switchers" cannot be estimated reliably, given that the data include the arsenic level of the initial source only.

Consistent with figures from this study, the calculations assume that tests costs USD 0.30 (or BDT 24 using an exchange rate of BDT 80/USD 1), and that personnel are paid BDT 45 per test delivered, with an additional bonus of BDT 12 for group sales. An amount of BDT 80 per test is added in arm C to account for the cost of the placards. Consistent with the experimental results, a take-up rate of 25 percent is set in arms A, B, and C, while consistent with results from earlier blanket testing a 100 percent testing rate is used when tests are offered for free. Again using estimates from the RCT, the calculations use a 30 percent switching rate among users of unsafe wells in arm A, while for arms B and C they use the estimates adjusted for controls and strata fixed effects in column (4) of table 4. That is, switching rates are assumed to be 0.49 in arm B and 0.70 in arm C. In the case of free provision, switching rates are varied from 0.3 to 0.75, consistent with findings from earlier work that evaluated switching after free provision. In the study area, the fraction of unsafe wells varied in the 16–27 percent range, while the earlier BAMWSP figures in these same villages varied from 40 to 90 percent. To cover a wide range of possibilities, calculations are shown using a fraction of unsafe wells that is either low (20 percent), medium (40 percent), or high (80 percent). The results are summarized in table 5, assuming that a policy maker is deciding how best to allocate a total and fixed budget of USD 10,000.

While free provision maximizes switching opportunities within a given locality, charging a fee allows for a larger coverage—at the cost of reducing uptake among those with low willingness or ability to pay.

[19] Of these 126 households, 11 were in arm A, 39 in arm B, and 13 in arm C. Because of the very small numbers involved, the results by arm are not analyzed in detail.

[20] Jamil et al. (2019), using data from blanket free testing in Araihazar, estimates a total cost of <USD 1 per person whose exposure was reduced.

Table 5. Cost Effectiveness of Different Sale Strategies

		Demand (percent)	Labor cost per test (BDT)	Unit cost for placard (BDT)	Total cost per test (BDT)	Wells tested given budget	Wells found to be unsafe	Total number unsafe wells in area	Switch rate	Wells no longer used	Cost per switch (USD)	Wells not tested
Fraction with unsafe wells	0.20											
Free provision		1	45	0	69	11,594	2,319	2,319	0.30	696	14.38	0
Free provision		1	45	0	69	11,594	2,319	2,319	0.75	1,739	5.75	0
A: Individual sales		0.25	0	0	24	33,333	6,667	26,667	0.30	2,000	5.00	20,000
B: Group sales		0.25	12	0	36	22,222	4,444	17,778	0.49	2,178	4.59	13,333
C: Individual sales+placards		0.25	0	80	104	7,692	1,538	6,154	0.70	1,077	9.29	4,615
Fraction with unsafe wells	0.40											
Free provision		1	45	0	69	11,594	4,638	4,638	0.30	1,391	7.19	0
Free provision		1	45	0	69	11,594	4,638	4,638	0.75	3,478	2.88	0
A: Individual sales		0.25	0	0	24	33,333	13,333	53,333	0.30	4,000	2.50	40,000
B: Group sales		0.25	12	0	36	22,222	8,889	35,556	0.49	4,356	2.30	26,667
C: Individual sale+placards		0.25	0	80	104	7,692	3,077	12,308	0.70	2,154	4.64	9,231
Fraction with unsafe wells	0.80											
Free provision		1	45	0	69	11,594	9,275	9,275	0.30	2,783	3.59	0
Free provision		1	45	0	69	11,594	9,275	9,275	0.75	6,957	1.44	0
A: Individual sales		0.25	0	0	24	33,333	26,667	106,667	0.30	8,000	1.25	80,000
B: Group sales		0.25	12	0	36	22,222	17,778	71,111	0.49	8,711	1.15	53,333
C: Individual sales+placards		0.25	0	80	104	7,692	6,154	24,615	0.70	4,308	2.32	18,462

Source: Authors' simulations.

Note: The estimates show the responses to different sale strategies, assuming a total budget of USD 10,000. Each test is assumed to cost USD 0.30 (BDT 24), while testers are assumed to be paid BDT 45 per test delivered, with an additional bonus of BDT 12 for group sales. BDT 80 per test are added in arm C to account for the cost of the metal placards. The take-up rate is assumed to be 25 percent in arms A, B, and C, and 100 percent when tests are offered for free. Switching rates among users of unsafe wells are assumed to be 30 percent, 49 percent, and 70 percent in arms A, B, and C, respectively. In the case of free provision, switching rates are varied from 0.3 to 0.75, consistent with earlier studies in neighboring areas.

Given the fixed budget, the total number of tests ranges from a maximum of 33,333 with individual sales, to a minimum of 7,692 with sales of tests supplied with a metal placard, so that the placards make arm C even more expensive (per test) than free provision without placards. Under the simplifying assumption that the probability of uncovering an unsafe well does not depend on the mode of supply, these figures imply that individual sales (A) would be the strategy that maximizes the number of unsafe wells uncovered, followed by group sales (B) and free provision, while sales with placard (C) would be the worst in this respect. However, the relative performance of the strategies changes once the different switching rates are taken into account. In particular, given the high switching rates observed in B and C, and the relatively low cost of group sales (B), it is group sales that maximize the number of unsafe wells that cease to be used for drinking. Individual sales (A) are second best, followed by either free provision (under the high-switching scenario) or sales with placards (C), while free provision is the worst under the assumption of switching rates as low as those observed in arm A. Note that the relative performance of the different strategies does not depend on the prevalence of unsafe wells. In contrast, the average cost "per switch" from an unsafe well decreases with the fraction of unsafe wells, as this leads to an increase in the number of households that may benefit from switching, while the fraction that does is not affected under our assumptions.

These estimates also show that even the strategies with the highest average cost per unsafe well averted are highly cost effective. The cost ranges from USD 1.15 (group sales with high prevalence of unsafe wells) to USD 14.4 (free provision with low switching rates and low prevalence of unsafe wells). However, Pitt, Rosenzweig, and Hassan (2015) estimated a present discounted value of per-household gains from switching to safe water sources over 20 years ranging from USD 1,400 to USD 1,000 for discount rates of 3 percent to 8 percent. Such estimates only take into account income gains that result from avoiding productivity losses due to consumption of arsenic-contaminated water, while they ignore the additional utility gains from better health and reduced mortality. Argos et al. (2010) estimate substantial declines in all-cause mortality over a 10-year period associated with high arsenic content of drinking water, with hazard rates ranging from 1.09 to 1.68 relative to "safe" wells with arsenic below 10 ppb. In addition, Keskin, Shastry, and Willis (2017) find that testing campaigns also reduced mortality among young children because arsenic risk induced mothers to breastfeed longer. Such reductions in mortality would make testing even more cost effective.[21] Note also that testing would remain cost effective even if only a fraction of switchers actually moved to a safe source.

Last but not least, it must be highlighted that while group sales with cost sharing appears to be the most cost-effective strategy for a given budget, it comes at a high cost in terms of equity. This is of course a by-product of low demand, which leads three-quarters of households not to learn about the safety status of their well. For instance, even under the scenario of only 20 percent of unsafe wells, while 4,444 unsafe wells would be identified, and 2,178 of them would no longer be used for drinking, we also find that in the same communities where sales took place there would be an additional 13,333 unsafe wells that would not be tested. Equity concerns, if individuals sales are not allowed, may be particularly serious for households that are isolated, either geographically or socially.

7. Conclusions

Information on household-specific environmental health risks can be a relatively inexpensive policy tool to mitigate those risks. However, the design of information campaigns often has to contend with resistance to behavioral change even when the presence of such risks has been revealed to target households. This may be especially true in developing countries, where poverty, low literacy, and other constraints may severely limit the effectiveness of such campaigns, especially if targeted information is supplied only for a fee. These

---

[21] There are few estimates of the value of a statistical life (VLS) for Bangladesh, and their range is very broad. For instance Mahmud, Sawada, and Yamada (2019), using data on willingness to pay to reduce mortality risk from air pollution in urban Bangladesh, estimates a value of 17,500–22,500 in PPP terms, or about 10–12 times per capita GDP, while Viscusi and Masterman (2017) use a figure about 10 times larger.

considerations are salient in Bangladesh, a country where millions of people use water from shallow tube wells for cooking and drinking, and where a large fraction of such water is estimated to be contaminated by naturally occurring arsenic in concentrations high enough to have serious health consequences in the case of long-term exposure. This is generating one of the most severe public health crises worldwide (Ahmed et al. 2006). Given that wells with unsafe water are often located at walking distance from safe wells, the provision of information on well-specific arsenic levels represents a potentially life-saving tool, allowing households to mitigate arsenic risk by simply changing their primary source of drinking water.

This article has described the results from a randomized field experiment to study the effect of different arsenic test selling schemes on test uptake and well switching. Despite the fees, a team of surveyors managed to test 2,800 of a total of about 11,400 wells. This allowed the uncovering of the presence of hundreds of wells with arsenic levels above the threshold adopted by the Government of Bangladesh. Overall, about half of the users of contaminated wells decided to switch to a different source of drinking water. Relatively subtle differences in the way information was sold and provided, while barely affecting demand, led to very substantial gaps in behavioral responses. Relative to simple, individual sales where test results were provided privately, both (a) group sales that leveraged informal local solidarity networks and (b) the addition of metal placards posted on the wells more than doubled switching rates among users of unsafe wells. These findings should be useful for the design of information campaigns that aim to provide measures of risk exposure that vary at the household level. In the context of this study, information was supplied only to households that chose to purchase a test, but similar considerations may also be relevant when information is provided for free, for instance through blanket testing campaigns such as the one conducted now more than 10 years ago by BAMWSP.

A number of caveats and limitations should, however, be emphasized. First, data limitations do not allow the mechanisms underlying the results to be conclusively disentangled, although the observed patterns are consistent with a conceptual framework where the adoption of health-protecting behavior is increased by pre-commitment to share drinking water (despite the absence of enforcing mechanisms), by the ease of access to information on safe sources, and by "reminders" on water safety provided by placards affixed to the tube-well spouts.

Second, the data do not include objective measures of exposure to arsenic, and so (unlike some earlier studies) it is not possible to determine whether the self-reported changes in the main source of drinking water were reflected in actual reductions in exposure. This also limits the ability to evaluate the cost-effectiveness of the interventions.

Third, the trial did not include an arm where tests were provided free of charge, and so it is necessary to resort to a number of assumptions when making cost-effectiveness comparisons between sales and free provision. With this caveat, the results suggest that cost sharing under some scenarios could achieve a significantly larger number of unsafe wells no longer being used for drinking for a given budget, but this would come at the expense of equity, as many wells would remain untested due to low demand. In addition, free, blanket testing campaigns may also increase switching rates among owners of unsafe wells by increasing the number of known safer alternatives available nearby.

Fourth, to the extent that the results can be extrapolated to the rest of the country, tests-for-fee campaigns can only provide a partial solution to the public health crisis due to arsenic in shallow aquifers. In the study area, about three-quarters of wells remained untested. That the vast majority of well owners did not know the safety status of their well suggests that the share of unsafe wells among these untested wells was similar to that among tested ones and, therefore, in the 15–30 percent range. These rates are much larger than the unconditional ITT estimates of the impact of the selling schemes on switching rates, which are in the 3–6 percentage point range. This large discrepancy suggests that, despite the many tests sold, switching rates achieved by the test sales program remained well below the likely fraction of unsafe wells.

Fifth, despite the likely selection into purchase of households more responsive to arsenic-related information, about half of users of unsafe wells were still using the same source at the time of the return visit.

Further, among those who switched to a different source, many switched to a well that was either still unsafe (although possibly *safer*) or with unknown contamination levels. That more guidance is needed to facilitate switching to safe water sources is also consistent with findings from the neighboring Araihazar subdistrict, where Pfaff et al. (2017) document that following the BAMWSP blanket testing campaign, about 30 percent of households whose well water was found to be unsafe had switched to other wells identified as unsafe or of unknown status.

Sixth, this study is silent as to whether demand was limited by the strategy of approaching women (usually the most senior woman in the household) to offer the tests. Miller and Mobarak (2013) show that women in Bangladesh were less likely than men to purchase improved cookstoves to reduce indoor pollution, despite their stronger preference for the new technology justified by them bearing much of the health costs of traditional stoves. These factors may have mediated the differences in switching rates between arms. For instance, it is possible that the strength of women's bargaining power in the choice to buy the tests, or in the choice of water source for drinking, was affected by group dynamics (especially in arm B) or by the public nature of the test results (especially in arm C).

In sum, and until game-changers such as regulated piped water become widely available, much remains to be learned about the optimal design of campaigns for the provision of information on environmental health risks. The findings discussed in this article suggest that facilitating the spread of information on safe options, reminders, and mechanisms that leverage the presence of peer groups may represent promising ways to maximize the adoption of risk-avoiding behavior.

References

Ahmed, M. F., S. Ahuja, M. Alauddin, S. J. Hug, J. R. Lloyd, A. Pfaff, T. Pichler, C. Saltikov, M. Stute, and A. van Geen. 2006. "Ensuring Safe Drinking Water in Bangladesh." *Science* 314(5806): 1687–8.

Alsan, M., and M. Wanamaker. 2017. "Tuskegee and the Health of Black Men." *Quarterly Journal of Economics* 133(1): 407–55.

Argos, M., T. Kalra, P. J. Rathouz, Y. Chen, B. Pierce, F. Parvez, and T. Islam et al. 2010. "Arsenic Exposure from Drinking Water, and All-Cause and Chronic-Disease Mortalities in Bangladesh (HEALS): A Prospective Cohort Study." *Lancet* 376(9737): 252–8.

Ashraf, N., J. Berry, and J. Shapiro. 2010. "Can Higher Prices Stimulate Product Use? Evidence from a Field Experiment in Zambia." *American Economic Review* 100(5): 2383–413.

Bai, L., B. Handel, E. Miguel, and G. Rao. 2017, August. "Self-Control and Demand for Preventive Health: Evidence from Hypertension in India." Forthcoming, Review of Economics and Statistics.

Balasubramanya, S., A. Pfaff, L. Bennear, A. Tarozzi, K. M. Ahmed, A. Schoenfeld, and A. van Geen. 2014. "Evolution of Households' Responses to the Groundwater Arsenic Crisis in Bangladesh: Information on Environmental Health Risks Can Have Increasing Behavioral Impact over Time." *Environment and Development Economics* 19(5): 631–47.

Barnwal, P., A. van Geen, J. von der Goltz, and C. K. Singh. 2017. "Cost-Sharing in Environmental Health Products: Evidence from Arsenic Testing of Drinking-Water Wells in Bihar, India." *Journal of Environmental Economics and Management* Forthcoming.

Bennear, L., A. Tarozzi, H. B. Soumya, A. K. Matin, A. Pfaff, and L. 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(2): 225–40.

Bennett, D., S. A. A. Naqvi, and W.-P. Schmidt. 2018. "Learning, Hygiene, and Traditional Medicine." *The Economic Journal* 128(612): F545–F574.

Carson, R., P. Koundouri, and C. Nauges. 2011. "Arsenic Mitigation in Bangladesh: A Household Labor Market Approach." *American Journal of Agricultural Economics* 93(2): 407–14.

Chen, Y., J. H. Graziano, F. Parvez, M. Liu, V. Slavkovich, T. Kalra, and M. Argos et al. 2011. "Arsenic Exposure from Drinking Water and Cardiovascular Disease Mortality: A Prospective Cohort Study in Bangladesh." *British Medical Journal* 342(7806): d2431.

Chen, Y., A. van Geen, J. H. Graziano, A. Pfaff, M. Madajewicz, F. Parvez, A. I. Hussain, V. Slavkovich, T. Islam, and H. Ahsan. 2007. “Reduction in Urinary Arsenic Levels in Response to Arsenic Mitigation Efforts in Araihazar, Bangladesh.” *Environmental Health Perspectives* 115(6): 917–23.

Chowdhury, A. M. R., A. Bhuiya, M. E. Chowdhury, S. Rasheed, Z. Hussain, and L. C. Chen. 2013. “The Bangladesh Paradox: Exceptional Health Achievement despite Economic Poverty.” *Lancet* 382(9906): 1734–45.

Chowdhury, S., A. Krause, and K. F. Zimmermann. 2016. “Arsenic Contamination of Drinking Water and Mental Health.” Industrial Relations Section Working Paper No. 607, Princeton University.

Cohen, J., and P. Dupas. 2010. “Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment.” *Quarterly Journal of Economics* 125(1): 1–45.

Cohen, J., P. Dupas, and S. G. Schaner. 2015. “Price Subsidies, Diagnostic Tests, and Targeting of Malaria Treatment: Evidence from a Randomized Controlled Trial.” *American Economic Review* 105(2): 609–45.

Dupas, P. 2012. “Health Behavior in Developing Countries.” *Annual Review of Economics* 3(1): 425–49.

———. 2014. “Short-Run Subsidies and Long-Run Adoption of New Health Products: Evidence from a Field Experiment.” *Econometrica* 82(1): 197–228.

Dupas, P., and E. Miguel. 2017. “Impacts and Determinants of Health Levels in Low-Income Countries.” In *Handbook of Economic Field Experiments*, edited by A. Banerjee and E. Duflo, 3–93. Amsterdam: North-Holland.

Fafchamps, M. 2011. “Risk Sharing between Households.” In Volume 1A of *Handbook of Social Economics*, Chapter 24, pp. 1255–80. San Diego and Amsterdam: North-Holland.

Flanagan, S. V., R. B. Johnston, and Y. Zheng. 2012. “Arsenic in Tube Well Water in Bangladesh: Health and Economic Impacts and Implications for Arsenic Mitigation.” *Bulletin of the World Health Organization* 90(11): 839–846.

George, C. M., A. van Geen, V. Slavkovich, A. Singha, D. Levy, T. Islam, and K. M. Ahmed et al. 2012. “A Cluster-Based Randomized Controlled Trial Promoting Community Participation in Arsenic Mitigation Efforts in Bangladesh.” *Environmental Health* 11(1): 41.

George, C. M., Y. Zheng, J. H. Graziano, S. B. Rasul, Z. Hossain, J. L. Mey, and A. Van Geen. 2012. “Evaluation of an Arsenic Test Kit for Rapid Well Screening in Bangladesh.” *Environmental Science & Technology* 46(20): 11213–9.

Goldberg, J., M. Macis, and P. Chintagunta. 2018. “Incentivized Peer Referrals for Tuberculosis Screening: Evidence from India.” NBER Working Paper No. 25279, National Bureau of Economic Research, Cambridge, MA.

Gong, E. 2015. “HIV Testing and Risky Sexual Behaviour.” *Economic Journal* 125(582): 32–60.

Howard, G., M. Ahmed, A. Shamsuddin, S. Mahmud, and D. Deere. 2006. “Risk Assessment of Arsenic Mitigation Options in Bangladesh.” *Journal of Health, Population and Nutrition* 24(3): 346–55.

Imbens, G., and D. Rubin. 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction.* New York: Cambridge University Press.

Inauen, J., R. Tobias, and H.-J. Mosler. 2014. “The Role of Commitment Strength in Enhancing Safe Water Consumption: Mediation Analysis of a Cluster-Randomized Trial.” *British Journal of Health Psychology* 19(4): 701–19.

Jamil, N. B., H. Feng, K. M. Ahmed, I. Choudhury, P. Barnwal, and A. van Geen. 2019. “Effectiveness of Different Approaches to Arsenic Mitigation over 18 Years in Araihazar, Bangladesh: Implications for National Policy.” *Environmental Science & Technology* 53(10): 5596–604.

Johnston, R., S. J. Hug, J. Inauen, N. I. Khan, H.-J. Mosler, and H. Yang. 2014. “Enhancing Arsenic Mitigation in Bangladesh: Findings from Institutional, Psychological, and Technical Investigations.” *Science of the Total Environment* 488: 477–83.

Keskin, P., G. K. Shastry, and H. Willis. 2017. “Water Quality Awareness and Breastfeeding: Evidence of Health Behavior Change in Bangladesh.” *Review of Economics and Statistics* 99(2): 265–80.

Kremer, M., and E. Miguel. 2007. “The Illusion of Sustainability.” *Quarterly Journal of Economics* 122(3): 1007–65.

Krupoff, M., A. M. Mobarak, and A. van Geen. 2020. “Evaluating Strategies to Reduce Arsenic Poisoning in South Asia: A View from the Social Sciences.” *Asian Development Review* (Forthcoming).

Madajewicz, M., A. Pfaff, A. van Geen, J. Graziano, I. Hussein, H. Momotaj, R. Sylvi, and H. Ahsan. 2007. “Can Information Alone Change Behavior? Response to Arsenic Contamination of Groundwater in Bangladesh.” *Journal of Development Economics* 84(2): 731–54.

Mahmud, M., Y. Sawada, and E. Yamada. 2019. “Willingness to Pay for Mortality Risk Reduction from Air Quality Improvement: Evidence from Urban Bangladesh.” JICA Research Institute Working Paper No. 190.

Martinez-Bravo, M., and A. Stegmann. 2019. “In Vaccines We Trust? The Effect of Anti-vaccine Propaganda on Immunization: Evidence from Pakistan.” CEMFI Working Paper No. 1713.

McArthur, J., D. Banerjee, S. Sengupta, P. Ravenscroft, S. Klump, A. Sarkar, B. Disch, and R. Kipfer. 2010. "Migration of As, and 3 H/3 He Ages, in Groundwater from West Bengal: Implications for Monitoring." *Water Research* 44(14): 4171–85.

Miller, G., and A. M. Mobarak. 2013. "Intra-household Externalities and Low Demand for a New Technology: Experimental Evidence on Improved Cookstoves." NBER Working Paper No. 18964, National Bureau of Economic Research, Cambridge, MA.

Opar, A., A. Pfaff, A. A. Seddique, K. Ahmedc, J. H. Graziano, and A. van Geen. 2007. "Responses of 6500 Households to Arsenic Mitigation in Araihazar, Bangladesh." *Health and Place* 13(1): 164–72.

Pfaff, A., A. Schoenfeld, K. Ahmed, and A. van Geen. 2017. "Reduction in Exposure to Arsenic from Drinking Well-Water in Bangladesh Limited by Insufficient Testing and Awareness." *Journal of Water Sanitation and Hygiene for Development* 7(2): 331–9.

Pitt, M. M., M. R. Rosenzweig, and N. Hassan. 2015. "Identifying the Cost of a Public Health Success: Arsenic Well Water Contamination and Productivity in Bangladesh." NBER Working Paper No. 21741, National Bureau of Economic Research, Cambridge, MA.

Pop-Eleches, C., H. Thirumurthy, J. P. Habyarimana, J. G. Zivin, M. P. Goldstein, D. De Walque, and L. Mackeen et al. 2011. "Mobile Phone Technologies Improve Adherence to Antiretroviral Treatment in a Resource-Limited Setting: A Randomized Controlled Trial of Text Message Reminders." *AIDS (London, England)* 25(6): 825.

Rahman, A., L.-Å. Persson, B. Nermell, S. El Arifeen, E.-C. Ekström, A. H. Smith, and M. Vahter. 2010. "Arsenic Exposure and Risk of Spontaneous Abortion, Stillbirth, and Infant Mortality." *Epidemiology* 21(6): 797–804.

Raifman, J., H. Lanthorn, S. Rokicki, and G. Fink. 2014. "The Impact of Text Message Reminders on Adherence to Antimalarial Treatment in Northern Ghana: A Randomized Trial." *PLoS One* 9(10): e109032.

Sanchez, T. R., D. Levy, M. H. Shahriar, M. N. Uddin, A. B. Siddique, J. H. Graziano, A. Lomax-Luu, A. van Geen, and M. V. Gamble. 2016. "Provision of Well-Water Treatment Units to 600 Households in Bangladesh: A Longitudinal Analysis of Urinary Arsenic Indicates Fading Utility." *Science of the Total Environment* 563: 131–7.

Smith, A., E. Lingas, and M. Rahman. 2000. "Contamination of Drinking-Water by Arsenic in Bangladesh: A Public Health Emergency." *Bulletin of the World Health Organization* 78(9): 1093–103.

Tarozzi, A. 2016. "Health Interventions in Low Income Countries: A (Not So) Low Hanging Fruit?" Els Opuscles del CREI, No. 42.

Tarozzi, A., A. Mahajan, B. Blackburn, D. Kopf, L. Krishnan, and J. Yoong. 2014. "Micro-loans, Bednets and Malaria: Evidence from a Randomized Controlled Trial in Orissa (India)." *American Economic Review* 104(7): 1909–41.

Tarozzi, A., A. Mahajan, J. Yoong, and B. Blackburn. 2009. "Commitment Mechanisms and Compliance with Health-Protecting Behavior: Preliminary Evidence from Orissa, India." *American Economic Review Papers and Proceedings* 99(2): 231–5.

———. 2008. "The Demand for, and Impact of, Learning HIV Status." *American Economic Review* 98(5): 1829–63.

van Geen, A., E. B. Ahmed, L. Pitcher, J. L. Mey, H. Ahsan, J. H. Graziano, and K. M. Ahmed. 2014. "Comparison of Two Blanket Surveys of Arsenic in Tubewells Conducted 12 Years Apart in a 25 km^2 Area of Bangladesh." *Science of the Total Environment* 488: 484–92.

van Geen, A., K. Ahmed, A. Seddique, and M. Shamsudduha. 2003. "Community Wells to Mitigate the Arsenic Crisis in Bangladesh." *Bulletin of the World Health Organization* 81(9): 632–8.

van Geen, A., H. Ahsan, A. Horneman, R. Dhar, Y. Zheng, I. Hussain, and K. Ahmed et al. 2002. "Promotion of Well-Switching to Mitigate the Current Arsenic Crisis in Bangladesh." *Bulletin of the World Health Organization* 80(9): 732–7.

van Geen, A., Z. Cheng, Q. Jia, A. A. Seddique, M. W. Rahman, M. M. Rahman, and K. M. Ahmed. 2007. "Monitoring 51 Community Wells in Araihazar, Bangladesh, for up to 5 Years: Implications for Arsenic Mitigation." *Journal of Environmental Science and Health Part A* 42(12): 1729–40.

Viscusi, W. K., and C. J. Masterman. 2017. "Income Elasticities and Global Values of a Statistical Life." *Journal of Benefit-Cost Analysis* 8(2): 226–50.

Wasserman, G. A., X. Liu, F. Parvez, H. Ahsan, P. Factor-Litvak, J. Kline, and A. van Geen et al. 2007. "Water Arsenic Exposure and Intellectual Function in 6-Year-Old Children in Araihazar, Bangladesh." *Environmental Health Perspectives* 115(2): 285–9.

World Bank. 2015. "Purchasing Power Parities and the Real Size of World Economies: A Comprehensive Report of the 2011 International Comparison Program." Technical Report, World Bank. doi:10.1596/978-1-4648-0329-1.