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

Alessandro Tarozzi
Ricardo Maertens
Kazi Matin Ahmed
Alexander van Geen*

June 2019

Abstract

Millions of villagers in Bangladesh are chronically exposed to arsenic by drinking contaminated water from private wells. Free testing for arsenic has been shown to encourage households with unsafe wells to switch to safer sources that are often within walking distance. We describe results from a cluster randomized controlled trial conducted in 112 villages in Bangladesh to evaluate the effectiveness of different schemes to sell information on well-water quality at inducing households to stop using arsenic-contaminated water for drinking. We study whether either informal inter-household agreements to share water from wells that are found to be safe, or visual reminders of well status in the form of metal placards mounted on the well pump, can increase risk-mitigating behavior relative to simple individual sales of privately provided information. At a price of about USD0.60, only one in four households purchased a test and sales were not increased by risk-sharing agreements or visual reminders. However, switching away from an unsafe wells almost doubled in response to agreements or placards relative to the one in three proportion of households who switched away from an unsafe well with simple individual sales.

JEL: I12, I15, I18, Q53
Key words: Arsenic, Bangladesh, Environmental Health Risk

*We are very grateful to Prabhat Barnwal for conversations about risk-sharing that started this project. We acknowledge partial support from the Earth Clinic at the Earth Institute, Columbia University and from NIEHS grant P42 ES010349. 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. We also thank 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, anonymous referees, as well as many colleagues at Lund University, 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 University of New South Wales, Sydney. The study was approved by the Columbia University IRB (Protocol AAAN9900) and the Government of Bangladesh NGO Affairs Bureau. Tarozzi (corresponding author), Department of Economics and Business, Universitat Pompeu Fabra, Barcelona GSE and CRES, alessandro.tarozzi@upf.edu; Maertens: Department of Economics, Harvard University; Ahmed, Department of Geology, Faculty of Earth and Environmental Sciences University of Dhaka, Dhaka 1000, Bangladesh; van Geen, Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY10964, USA. All errors are our own.

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, ranging from the poor availability and high cost of good quality health care, to the insufficient investment in prevention, and to the frequent reliance on ineffective and sometimes unnecessarily expensive treatments, see Dupas (2012), Dupas and Miguel (2016), and Tarozzi (2016) for recent reviews. Information campaigns on health risks are sometimes seen as an appealing tool in environmental and health policy, 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. In addition, some health conditions are in principle easily preventable if appropriate behavior is adopted to avoid environmental exposures. 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.

In this paper, we describe the results of a randomized controlled trial (RCT) carried out in Sonargaon sub-district, Bangladesh, to evaluate the effectiveness of different approaches to sell information on well-water quality at inducing risk-mitigating behavior. Despite much progress in numerous health indicators (Chowdhury et al. 2013), Bangladesh remains in the midst of an extremely severe health crisis due to the widespread presence of low-dose, naturally occurring arsenic (As) in shallow aquifers, see Ahmed et al. (2006), Johnston et al. (2014), and 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, un-regulated shallow tube wells for drinking and cooking. Using nationwide data from 2009, Flanagan et al. (2012) estimated that, in a country of more than 150 million people, about 20 million were likely exposed to arsenic levels above the official Bangladesh standard of 50 ppb (parts per billion, or micro-grams per liter), while almost one third of the population was likely exposed to levels above the significantly lower threshold of 10ppb adopted by the World Health Organization (WHO).

The most visible health consequences of chronic exposure to arsenic from drinking tubewell 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 et al. 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 et al. 2016). These health effects are accompanied by significant economic impacts: exposure to arsenic has been estimated to reduce household labor supply by 8% (Carson et al. 2011) and household income by 9% per every earner exposed (Pitt et al. 2015), while Flanagan et al. (2012) calculated that a predicted arsenic-related mortality rate of 1 in every 18 adult deaths represents an additional economic burden of USD13 billion in lost productivity alone over the next 20 years.

Piped water from regulated and monitored supplies would likely 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 now is that household-level water treatment, dug wells, and rain-water harvesting are not viable alternatives for lowering arsenic exposure, also 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, tubewells may offer an effective way of providing safe drinking water to the rural population of Bangladesh in the short to medium term. With the exception of the most severely affected areas of Bangladesh, 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 sub-district close to our study area, Jamil et al. (2019) estimate that community-wide 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 community wells.

The potential of well-switching as a mechanism to reduce arsenic risk also emerged during a well-testing campaign coordinated between 1999 and 2005 by the Bangladesh Department of Public Health Engineering (DPHE), with support from the World Bank, DANIDA, UNICEF and a number of non-governmental organizations (NGOs). The initiative, coordinated through the Bangladesh Arsenic Mitigation and Water Supply Program (BAMWSP) tested with a field kit close to 5 million wells, and identified them as 'safe' or 'unsafe'—according to the Bangladesh standard of 50ppb—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 typically coming from trials that provided information campaigns on arsenic health risks, and repeat visits, in some cases with health measurements taken (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 et al. 2014; Pfaff et al. 2017). Despite these partial successes, a substantial fraction of households continue using unsafe wells, 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, also because to date testing campaigns such as BAMSWP 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.[1] 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.

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

The primary objective of our study was thus 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 in Sonargaon sub-district, we offered field tests at a (subsidized) price of BDT45 (about USD0.60 at current nominal exchange rates, close to the price of one kg of rice in Dhaka), an amount we had estimated would be just enough to cover for the salary of the surveyors hired for the project.[2] In an additional subset of 48 villages, our surveyors were instead incentivised to offer tests—at the same price of BDT45—to *groups* of buyers, where 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. Immediately after conducting the tests, the result of all tests were communicated to all group members. The agreement was not binding legally, but our 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.[3]

We also examine 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 us to include 15 villages in this experimental arm, reducing statistical power. In these villages, individuals who purchased a test at the same price of BDT45 were also given a metal placard of a color depending on the arsenic level: blue for arsenic below 10ppb, green if between 10 and 50, and red if 'unsafe', that is, above the government threshold of 50ppb. 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 for instance during the BAMWSP testing campaign—of applying to the well spout red or green paint that would often become invisible within a year (Pfaff et al. 2017). Such visible indicators of safety can act both as a reminder about the safety of the well water, and as a means to facilitate the spread of information about which wells are safe within a 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 BDT80) is high enough to increase significantly the total cost of testing campaigns. It was thus important to determine whether they made any difference relative to the alternative solution (adopted in the two experimental arms described earlier) of informing the household via a simple and inexpensive laminated card to be kept in the house, with the indication of

[2]Throughout the paper, when we convert Bangladesh Takas (BDT) into United States Dollars (USD) we use a nominal exchange rate of 80BDT/USD, and a PPP exchange rate of 23.145, as indicated in World Bank (2015, Table 2.1).

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

the test result.

Despite the low and subsidized sale price, and widespread awareness about the arsenic problem coupled with little information about the safety status of individual wells, we found that only about one in four households purchased a test, regardless of the offer type. 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 (Cohen and Dupas 2010, Dupas 2014, Tarozzi et al. 2009, Tarozzi et al. 2014), to de-worming drugs (Kremer and Miguel 2007) and water-disinfectant (Ashraf et al. 2010). However, while the offer type barely affected demand, it did matter for how households responded to the information. Our Intent-to-Treat estimates (*not* conditional on purchase) show that while standard individual sales led 3.7% of households to switch water source, the fraction was 4.3% with group sales (14% higher, 95% C.I. $[-0.0137, 0.0339]$, p-value 0.404) and 6.4% (75% higher, 95% C.I. $[0.003, 0.072]$, p-value 0.031) when metal plates were attached to the well spout in case of purchase, although the difference is only statistically significant in the latter case. In addition, we find very substantial differences between arms in the response of households who receive 'bad news' about the safety of they well water. Among these households, switching rates almost doubled from 30 to 56% with group sales relative to standard individual sales (95% confidence interval of the difference adjusted for baseline covariates is 0.011-0.362), and it more than doubled when metal plates were supplied (from 30 to 72%, 95% C.I. of adjusted difference 0.188-0.609).

Our 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 devise risk mitigation strategies. We focus 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 we study demand for information on environmental factors, earlier work has looked at demand for information on health status. Cohen et al. (2015) study how subsidies for rapid-diagnostic tests (RDTs) for malaria affect both the demand for the tests as well as demand for anti-malarial drugs. They find that while heavy subsidies increase considerably demand for RDTs, about half of individuals who test negative for malaria still decide to purchase anti-malarial drugs. Bai et al. (2017) study demand for commitment contracts to schedule preventive doctor visits by hypertensive patients in rural Punjab, India. They find that a large share of patients purchase these contracts but do not follow through with the commitment, leading to monetary losses. Thornton (2008) show that monetary incentives increased substantially demand for HIV testing in rural Malawi, but she also finds that behavioral responses, in the form of increased demand for condoms, were muted. Using data from Kenya and Tanzania, Gong (2015) finds that individuals surprised by an HIV-positive test increased risky sexual behaviour. 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 achieving desirable policy objectives.

Our paper relates to that of Barnwal et al. (2017), who estimate a demand curve for arsenic tests in Bihar, India, another location with a groundwater arsenic problem. They find that test uptake fell from 69% to 22% of households when the price increased from INR10 to INR50, where the latter was about equivalent to the daily per capita income. Further, they estimated that at INR40 (about BDT49) uptake was 25%, which is about the same as what we estimate at a very similar price of BDT45. Unlike Barnwal et al. (2017), we cannot estimate how demand changes with price, but we examine the role of non-price factors on demand and behavioral responses to information. In our setting, 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. In the next section, we provide additional background information on the extent of the arsenic problem in the study area and describe the experimental design. In Section 3, we describe the data collection protocol, present selected summary statistics, and show 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). In Section 4 we present the conceptual framework that guided the study design and that will be useful to interpret the results, which are then described and interpreted in Section 5. The cost effectiveness of the interventions is evaluated in Section 6. Finally, Section 7 offers a further discussion of the results and highlights a number of limitations of our study.

2 Study design

This study was carried out in Sonargaon, a sub-administrative unit (or *upazila*) of Narayanganj district, located approximately 25 kilometers south-east of the capital Dhaka. According to the 2011 Census of Bangladesh, Sonargaon had a population of about 400,000, and administrative records at the time of the study listed a total of 365 villages, in a 171 square kilometers territory. Sonargaon is located in a part of the country where arsenic contamination of shallow tubewell water is widespread. According to a blanket testing conducted in 1999-2000, under the supervision of the Bangladesh Arsenic Mitigation and Water Supply Program (BAMWSP), in about 80 percent of villages 40 percent or more of tubewell water had arsenic levels above the Bangladesh standard of 50ppb, while the median proportion of unsafe wells was a high 86% (Chowdhury et al., 2000).[4]

For this paper, we first selected all 128 villages in Sonargaon with more than 10 wells and with a share of unsafe wells between 40 and 90 percent according to data from the BAMWSP blanket testing campaign 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.[5]

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

[5] According to BAMWSP data, in 42% of villages in the sub-district more than 90% of wells were unsafe according to Government of Bangladesh standards.

In each study village, surveyors would conduct home visits to households to identify all wells, regardless of whether they had been tested before. Privately owned wells were linked to household who owned it, while the very few public wells were linked to the main caretaker or user. In each household linked to a well they would then explain the risk of consuming arsenic-contaminated tubewell water to an adult—typically the most senior woman—and offer to test the well for a fee. Additionally, surveyors recorded the geographic coordinates of all wells using the GPS receiver of their smartphone, regardless of whether the owner bought a test or not. When a test was purchased, tubewell water was tested using the Arsenic Econo-Quick (EQ) test kit, which has been shown to be reliable when used in the field, and can deliver results within ten minutes, see George et al. (2012) for details. 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 USD0.30 for volume purchases, although the total cost per test was estimated to be about USD2.4 per test, for a testing campaign that also covered the costs of trained personnel and metal placards to be attached to the tubewell spouts (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 warnings: (i) that arsenicosis is not a communicable disease, (ii) that arsenic cannot be removed by boiling water, (iii) that testing tubewell water for arsenic is important, and (iv) that the Bangladesh safety standard for arsenic concentration in water is 50ppb. Black cards with these messages were given to households who did not buy a test, while in case of test purchase the laminated card was blue (if arsenic was up to 10ppb), green (if it was 25 or 50ppb), or red (if the test showed an arsenic concentration above 50ppb, 'unsafe' by Government of Bangladesh standards). Owners of unsafe ('red') wells were also encouraged face-to-face to switch to a safe (blue or green) well, while owners of wells with concentrations below 10ppb were encouraged to share their well water with their neighbors. Green card holders were both encouraged to share their water and to switch to a safer (blue) well, if possible.

Our experimental variation comes from differences in selling schemes for arsenic tubewell water tests across villages. In a first group of villages, which we term group A, surveyors offered to test tubewell water for a fee of BDT45 (about USD0.60). This fee was expected to cover the salary of the testers and their supervisor. Of the BDT45 charged per test, testers kept BDT30 to cover their transportation expenses and salary, and handed over the remaining BDT15 to their supervisor. The price was determined assuming that a field worker would test about 15 wells/day for 20 days/month, leading to a monthly salary of BDT9,000 (USD112.5/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, the supervisor of 10-15 workers would earn BDT45,000-67,500 (USD563-844), a range that spans what he earned while supervising the testing in Araihazar in 2012-2013. Across all experimental arms, the cost of the field kits (USD0.30/test) was covered by the project. In practice, demand (and thus testers' compensation) was lower than expected, see Section 7 for discussion.

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 three and up to ten households, while individual sales were not allowed.[6] Surveyors would help group formation, for instance by proposing a sale to all household within the same compound (or *bari*), and then coordinating the inclusion of additional buyers via mobile phones. Each household would be responsible only for the payment of BDT45 to have its well tested, after an informal well-sharing agreement between all group members. 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 of the same group were informed of the test results for all within-group wells.

A limitation of our data is that surveyors did not keep accurate records of who belonged to which group. In other words, while our data indicate for which wells a test was purchased, we 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 comparison of demand for and responses to tests between arms A and B was the primary objective of the study. Assuming that the testing campaign would uncover five unsafe wells per village, that 30% of households with unsafe wells would switch in the control group (a rate at the lower end of what observed in previous studies), and an intra-village correlation of 12%, we determined that 50 villages per arm would ensure 80% power for a two-sided test of equality at the 5% significance level, and with a 15 percentage points difference in switching rates between arms.[7] The trial was not registered, and we did not prepare a pre-analysis plan.

Given that the available funding allowed the inclusion of a larger number of villages, we opted for the inclusion of additional experimental arms, although for these additional arms sample size was dictated by budget constraints rather than power calculations. In a third group of 16 villages (C), households were again assigned to receive individual test offers at BDT45 (as in group A). However, 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 color whether the arsenic concentration was below 10ppb (blue), between 10ppb and 50ppb (green), or above 50ppb (red). Accordingly, as shown in Figure 1, they 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.

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 further gave a bonus of BDT12 per sale to testers, to compensate for the additional effort necessary to coordinate group formation.[8]

[6] In practice, most groups included 7-10 buyers, although some had as few as three and some had 11.

[7] In this scenario the effect size is $0.15/\sqrt{0.3 \times 0.7} = 0.33$.

[8] The experimental design also included two exploratory arms, with only six villages each, where tests were sold individually either at a village-level price of BDT45 or BDT90, but with payment required only in case of 'good news', that is, in case of arsenic level no higher than 50ppb. The inclusion of these proof-of-concept conditional sales were motivated by the observation made during focus group that a number of respondents were averse to the idea of 'paying

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).[9] 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 re-assignment of treatments was thus due to a data-entry error, and not to imperfect implementation of the protocol in the field. In addition, the checks for balance in covariates are very similar if we use assigned or actual treatment (see below). For this reason in the analysis we define treatment status as *actual* treatment. Second, in four cases, surveyors were unable to differentiate a village from the one adjacent to it. While we have data from households in these four villages and the ones adjacent to them, we can only distinguish pairs, and both villages in each pair received the same treatment. For this reason, in the statistical analysis we have effectively 112 clusters divided into experimental arms A (49 clusters), B (48), and C (15). For simplicity, in the rest of the paper we will refer to these clusters as 'villages'.

The spatial distribution of villages in our intervention is displayed in Figure 2. As expected, the randomization—with the caveats described above—led to large spatial variation in treatments.

3 Data

Between January and June 2016, surveyors completed a census of all wells in the 112 study villages, and for each well they identified the household who owned it or, for a small number of community-owned wells, the primary user. Almost all wells (98.6%) were privately owned, and for simplicity in the rest of the paper we will somewhat loosely use the term 'owner' to refer to the household who owned the well, or to the household who was the primary user of community wells.

Well owners were then offered an arsenic test under one of the three selling schemes. Concomitantly, enumerators administered a short 'baseline' household survey and recorded information on sales and, in case of purchase, the result of the test. The result was immediately communicated to the buyer. Additionally, surveyors recorded GPS coordinates of all wells and whether, at the time of the visit, there were already any visible labels attached to the well indicating the arsenic-safety status. A total of 12,606 wells were listed at baseline, and the baseline questionnaire was completed for all but three of the corresponding owners. It is important to note that, while our data give us a complete census of *wells* at the time of the baseline, we do not have a census of *households*. An implication of this is that

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, we do not discuss the results in detail although these are available upon request from the authors.

[9] Unions are the third smallest administrative unit, and are formed by several *mouzas*, which in turn are composed of two to three villages. The 128 study villages belong to nine unions.

we only have information on households who owned a well, and while these were the majority, we make no claim that well owners are a representative sample of the study population. The choice to survey only owners was due to budgetary constraints, but of course an implication of it is that we cannot study whether the choice of the primary source of drinking water changed also among non-owners.

The endline survey was completed between August 2016 and January 2017.[10] The average time elapsed between baseline and endline surveys was 7.7 months, and 86% of households had their follow-up interview between seven and nine months after the baseline interview.

At the time of the endline survey, interviewers were instructed to return to the wells identified during the test sales, and to verify whether the corresponding household was still using the well as primary source of water for cooking and drinking. In case of a negative answer, the surveyor would ask the respondent to accompany him/her to the actual source, and then would record the new GPS location, verify the presence of visible indicators of arsenic safety (for instance the presence of one of the metal placards distributed during our intervention), and then would ask the respondent about perceived safety of the new source as well as about the primary reason for switching to a different source.

3.1 Summary Statistics at Baseline

We present selected summary statistics measured at baseline in Table 1. Throughout this paper, unless otherwise noted, we restrict our analysis to the large majority of households (91%) that used their own well at baseline, as this is the sample for which we can determine baseline water source and post-intervention switching.[11] All summary statistics except those on the first line of Table 1 are thus calculated for households who used the well for drinking and cooking.

The head was male in 85% of households, while 29% of the heads were wage-workers and 42% were self-employed, with the remaining largely occupied in domestic activities. Household heads had low levels of educational attainment on average, with the majority having only primary schooling or less. Most households were poor, with only 17% of the houses having a concrete roof (an indicator of wealth), while the rest had tin or (in rare cases) mud roofs. Further, most households were small in size, with an average of 3.6 members in total, of which 1.5 where children.

The average well owner in our study area lived in a village where 75% 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%) did not know whether their well was safe or unsafe with respect to arsenic. In contrast, only 7% of them thought that their well was unsafe, while the remaining 17% reported having a safe well. It is also important to note that, despite about one

[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.

[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%) were believed to be safe by the owner. For these households, during the follow-up visits, we can only determine the new source if the respondent stated that they now used the well for drinking and cooking, see Section 5.3.2.

quarter of respondents saying that they knew the status of their well, more than 99% of wells had in fact no visible sign of safety status, such as the spout painted red or green, or a metal placard attached to it. 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 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 clearly reflected in our data. Virtually all respondents replied 'yes' to the question "[h]ave you ever heard about arsenic in tubewell water?" Similarly, all but a handful of respondents replied yes when asked "[a]re you aware of the health risks of drinking tubewell water containing arsenic?"

Almost all wells (98.6%) were privately owned and on average relatively shallow (179 feet, or 55 meters) and about nine years old—which again suggests that a significant proportion of wells were installed after the BAMWSP blanket testing. The average reported installation cost of wells in our sample was BDT7,560, or about USD100 (USD323 using the PPP exchange rate from World Bank 2015). Well depth is a key predictor of installation costs: in our data, the elasticity of cost with respect to depth is 0.72 (s.e. 0.04). The BDT45 price charged for the test in our study thus represented slightly more than one half of a one percent of the installation cost.

It is interesting to note that well-sharing was already common in our study area: while the average household had fewer than four 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]

Column 7 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% level. The differences among arms was just due to chance and recall that, because baseline data were collected at the time of the arm-specific sales, we had very limited ability to enforce balance through stratification. The figures show that in a few cases there were substantive differences among arms along likely important characteristics such as the education of the household head, or knowledge about the safety status of one's well with respect to arsenic. For instance, while on average 19% of household heads in our study area had no schooling, this number drops to about 5% in treatment C and is close to 26% in arm A. Or, the fraction of respondents who did not know the status of their well ranged from 68% in arm A to 90% in arm C, while the fraction with a well described as safe ranged from 4% in C to 22% in A. Both the group-specific means and the tests of significance are very similar if we repeat the estimation using treatment as initially randomized rather that actual treatment (recall that there were some discrepancies due to clerical errors in programming the smartphones used for the survey).[13]

[12] Recall that we did not survey households who did not own a well.

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

In order to more systematically gauge the overlap in the distribution of covariates between arms, we also adopt a number of approaches described in Imbens and Rubin (2015, Ch. 14). First, for each covariate X in Table 1 and for each pair of experimental arms $a, a', \in \{A, B, C\}$, $a \neq a'$, in columns 8-10 we show *normalized differences* calculated as

$$\hat{\Delta}(X)_{a,a'} = \frac{\bar{X}_a - \bar{X}_{a'}}{\sqrt{\frac{s_a^2 + s_{a'}^2}{2}}} \tag{1}$$

where $\bar{X}_a$ and s_a^2 are the sample mean and sample variance of the variable X in arm a, respectively. While the usual t-statistics used to construct the tests for balance have a numerator 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. We also report, for each pair of arms a and a', a 'multivariate difference' estimated with a Mahalanobis distance calculated as

$$\hat{\Delta}_{a,a'} = \sqrt{(\bar{\mathbf{X}}_a - \bar{\mathbf{X}}_{a'})' \left(\frac{\hat{\mathbf{\Sigma}}_a + \hat{\mathbf{\Sigma}}_{a'}}{2} \right)^{-1} (\bar{\mathbf{X}}_a - \bar{\mathbf{X}}_{a'})}, \tag{2}$$

where $\bar{\mathbf{X}}_a$ and $\hat{\mathbf{\Sigma}}_a$ are the vector of means and the covariance matrix for all variables in arm a.[14] 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 to be likely 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 results in column 8 suggest 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 calculated as in (2) is 0.604. In contrast, the figures in columns 9 and 10 suggest that lack of balance is more problematic when we compare either arm A or 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 if also relatively large and equal to 1.1. The comparisons between B and C also show that four of 22 standardized differences are larger than 0.3, with a multivariate difference equal to 0.720.

Because we find lack of balance in characteristics (such as beliefs or schooling) that may affect behavior, we will also show results that control for observed covariates. We will show that the estimates are qualitatively robust to such inclusion, although the point estimates are in some cases affected,

[14] In calculating these multivariate differences, we exclude—because there is barely any variation in the data—the dummy for awareness of arsenic in well water, the dummy for awareness of associated heath risks, and the dummy for the well being privately owned.

and the standardized differences described above suggest that some caution should be exercised in particular when making comparisons that involve group C.

At the bottom of Table 1 we look at attrition. Overall, 8.8% of the households drinking from the index well could not be matched to the endline data, either because of true attrition (6 percentage points) or because errors in inputting the 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 our attrition measures.

4 Conceptual Framework

Before discussing the results, it is useful to think about the main factors likely to influence purchase choices and, conditional on test results, risk-mitigating behavior. In doing so we will not use a formal model but rather offer a simple conceptual framework that should help interpreting the results, also in light of ex-ante predictions about the impact of factors that were experimentally varied.

Willingness to pay for a test likely required 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 case of 'bad news' there will be safe (or at least *safer*) mitigation strategies available. We argue that all these conditions were present in our empirical context.

First, we have shown that a large majority (76%) of respondents did not know whether their well water was safe or unsafe to drink (see Table 1), and even among the rest there may have been some uncertainty. In addition, we have also seen that almost all wells had no visible sign of safety status such as paint or metal placards on the tubewell spout, so some may have valued the possibility to demonstrate water quality to others by displaying test results.[15] 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 et al. 2015, Bennett et al. 2017, Alsan and Wanamaker 2017, Martinez-Bravo and Stegmann 2017). Although we do not have direct measures of trust, we have explained that many wells had already been tested in the past in Sonargaon, so the local population was likely used to water samples being tested. In addition, earlier work carried out in the neighboring Araihazar sub-district found that self-reported switching behavior among households learning that their well was unsafe was validated by reductions in arsenic concentrations measured via biomarkers (see for instance Chen et al. 2007, Opar et al. 2007).

[15] In principle, such value needs not 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 more likely to be removed by households than those indicating low levels of arsenic two years after installation, 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 and Bennear et al. 2013).

Moving now to the second factor affecting demand (the relevance of the information), we have seen that virtually all respondents knew—at least in a general sense—about the presence and health risks of arsenic in tubewell water. Data on risk perceptions collected in the neighboring Araihazar sub-district in 2008 were also consistent with this, and a majority of respondents actually showed to be aware not only of the serious nature of arsenic risk, but also that the risk becomes more severe with prolonged exposure, see Tarozzi (2016, Figure 4) for details.

The third key factor affecting decisions is the perceived availability of alternative sources of drinking water. Well-sharing was already practiced in the area (see Section 3), so to some extent several households were already aware of the possibility of using neighbours' well water for drinking. We have also discussed how a testing campaign had been conducted by BAMWSP in the area about ten years earlier, and evidence from the neighboring Araihazar sub-district has shown that such campaign led to a substantial degree of well sharing. However, in Table 1 we have seen that about three quarter of wells had been found to be unsafe in our study area, and this may have reduced the perceived chance to have safe options nearby in case of bad news, although the average household had more than 10 other wells within a short 50m radius and about 30 within 100 meters.

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). Still, conditional on test results, we expected the decision to change source to depend primarily on the information made available by test results. Hence, regardless of the experimental arm, we predicted very little switching 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. Our 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, we expected 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 higher proportion of households to stop drinking from the tested well, relative to the individual sales in group A. In contrast, we did not have clear priors about the switching rates in B relative to C. Our predictions for *demand* were not as sharp: while we expected factors leading to higher willingness to react to information to also lead to higher willingness to pay for it, key factors such as health risks and availability of alternative sources were likely to become more salient after the realization of the test result.

5 Results

In this section, we first estimate the effect of our selling schemes on demand for testing. Next, we describe the information on arsenic levels that was revealed by the tests, and finally we discuss to what extent such information changed household behavior in terms of choice of water source for drinking.

In describing the results we mostly focus on behavior among households who were using the well as primary water source for drinking and cooking at baseline, given that for those who were not our baseline records do not indicate what the main source was.

5.1 Demand

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

$$buy_{svh} = \beta^B B_v + \beta^C C_v + \gamma X_{svh} + \delta_s + \epsilon_{svh}, \tag{3}$$

where buy_{svh} is equal to one if household h in village v and stratum s bought a test at baseline, and zero otherwise, B_v and C_v are village-specific indicator variables for the respective treatments, X_{svh} is a set of predetermined household and tubewell characteristics, and ϵ_{svh} is an error term. To account for the stratified design, we further include strata fixed effects (δ_s). Recall that we stratified treatment by the prevalence of unsafe wells based on BAMWSP data and by union. All standard errors and statistical inference are robust to the presence of intra-village correlation of residuals.

In Figure 3 we show 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 three experimental arms. As in many earlier studies looking at demand for health-related preventive products, even a relatively small fee, for potentially vital health-related information, led to low demand among potential beneficiaries.

We show the regression results in Table 2, where recall that consistent with equation (3) we adopt arm A as the reference group, so that the arm-specific coefficients represent the differences relative to the mean in A. Not surprisingly, the small differences in demand between arms A, B and C are not statistically significant. A comparison of the results in columns 1 and 2 shows that the inclusion of the strata fixed effects barely changes the point estimates, although the estimates become substantially more precise.

In column 3 we show that the results are quite robust to the inclusion of controls. This is important, because we have seen that despite the randomization there were some potentially important differences in means between arms. Because missing values in one or more of the controls lead to the loss of about 20% of observations, in column 4 we re-estimate the model without controls but including only the observations with complete observations used in column 3. 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, if only at the 10% 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 3, we see that low schooling likely decreased demand, while the higher prevalence of wells of unknown safety likely increased it, conditional on other observed characteristics. Both these factors suggest that the omission of controls may have biased demand upwards in arm C, although substantively the point estimates remain very close.

Although the coefficient estimates for the controls in column 3 cannot be interpreted causally, we highlight some interesting findings. First, demand was higher among households with better socio-economic status, as proxied by schooling of the head, better quality roofing, and more expensive wells. Finally, 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).

5.2 Test Results

Although the purchase rate was far from 100%, our 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 first describe such information. The test results are summarized in Table 3, where we also include the detailed summary statistics about switching behavior that we will describe later, and so we focus on the 10,412 households (91.3% of the total) that could be tracked in the endline survey. Of these, 2,417 where tested during our intervention.

Overall, 19% (455/2417) of the tested wells which had been used for drinking at baseline had 'unsafe' arsenic levels based on the Government of Bangladesh standards. Notably, the fraction was much lower than what observed at the time of the BAMWSP testing campaign, about 10 years earlier: in fact, recall that we included in our study only villages where BAMWSP estimated a fraction of unsafe wells in the 40 to 90% range. The reduction in the fraction of unsafe wells over time is consistent with a degree of learning about local arsenic risk, but also with economic development leading to an increasing number of households able to afford deeper wells, which are on average safer but are also more expensive to drill. Our data are broadly consistent with both hypotheses. First, we find that more recent wells, dug no more than 10 years earlier, were 22% deeper and 25 percentage points less likely to be unsafe relative to older ones.[16] Second, and although the majority of households were not sure about the safety of their wells, we also find that their beliefs about safety were strong predictors of actual safety status, suggesting a degree of sophistication or at least learning from previous testing (such as during BAMWSP), although of course we can only gauge the relationship for households who purchased the test. If we regress a dummy equal to one for unsafe wells on dummies for whether the respondent thought that the well is safe, or unsafe, we find that the belief of drinking from a safe well

[16] The average depth for older wells was 154 feet, while more recent ones were on average 34 feet deeper. The fraction of unsafe wells was 23 and 17% among the older and more recent wells, respectively. The test of equality is rejected at the 1% level for depth and at the 5% level for safety.

decreases the predicted probability of the well being unsafe by 15 percentage points while the belief of the well being unsafe *increases* it by 41 percentage points (both coefficients are significant at any standard level). A degree of consistency between beliefs about safety and actual safety status was also found in van Geen et al. (2014, Figure 4).

There was also some variation in the test results across different treatments. Recall that we are looking at results *conditional on demand* so that the randomization across treatments was in no way a guarantee of similar distributions across arms, even in large samples. The distribution of arsenic was overall similar among arms A and B, the two largest arms. Arm C has more unsafe wells (27%, versus 19 and 16% 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). Again consistent with the existence of a degree of awareness about arsenic risk, group C was by far the one with the smallest fraction of respondents thinking that their well was safe, although the fraction believing the well being 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, we estimate that at the time of the endline survey, of all the wells found to be unsafe, 30% had wells identified as safe within 25 meters, 57% had at least one within 50 meters, and 78% had at least one within 100 meters. This confirms that, in principle, switching to a nearby safe well was indeed 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, we find that 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 meters, such frequencies ranged across arms A, B and C, from 27 to 33%, from 55 to 60% and from 73 to 85%, respectively, and the null of equality in the frequencies is never rejected at standard levels. Note also that these figures surely 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 nearby whose status was unknown because the owner did not purchase the test.

5.3 Responses to test results

Next, we gauge to what extent households responded to information by using data collected at endline, about eight months after the test sales. At this time, interviewers verified whether the corresponding household was using the well as primary source of water for cooking and drinking. In case of a negative answer, the surveyor would ask the respondent to accompany him/her to the actual source, whose GPS location would then be recorded, together with the perceived safety of the new source and the primary reason for switching. Switching behavior was thus self-reported, but earlier work in the neighboring Araihazar sub-district found that switching behavior recorded in a way similar to our study was actually consistent with urinary arsenic concentrations, an objective biomarker of exposure (Chen et al. 2007).

Despite this, we cannot conclusively rule out that self-reported information on switching behavior may have suffered from courtesy bias, possibly differentially between experimental arms. For instance, safety information provided publicly (as in arm C) or to a group (as in B) may have increased this type of bias. Our data do not include biomarkers, and so we cannot validate self-reported behavioral responses against reliable indicators of exposure to arsenic. However, the requirement to accompany the enumerator to the new source likely reduced courtesy bias. In addition, 'switchers' in arms B and (especially) C reported walking longer distances to the new well relative to what we recorded in A, suggesting that the different sale strategies really affected the search for arsenic-safe options. If the higher switching rates in B and C relative to A had been driven by courtesy bias, we would have expected to observe shorter distances, given that the respondent had to accompany the interviewer to the new source of drinking water, and a well nearby would have likely been an easier and faster choice if the only purpose was to back up cheap talk.

In Figure 4, we show the raw switching rates observed in each experimental arm, without any control or strata fixed effects, and as usual with 95% confidence intervals for each arm-specific estimate. Regression results are shown in Table 4, where we always include strata fixed effects, and inference is robust to intra-cluster correlation of residuals. Our prior was that little or no behavioral change would be observed among households who did not buy a test, and among those who found that the well they used for drinking was safe. Indeed the figures in Table 3 show that barely anyone moved from a well tested as safe (12/1,950), while less than 3% of untested wells (224/7,995) stopped being used for drinking. The two bottom bar charts in Figure 4 show that for these two groups the switching rates were similarly very small in all arms. The main focus of our study was the behavioral response of users of unsafe wells, but because such responses are conditional on the choice to purchase a test and on the test result, we first describe the intent-to-treat (ITT) estimates, which measure unconditional switching rates.

In Figure 4.1 (top left) we show the raw ITT: while standard individual sales (A) led 3.7% of households to switch water source, the fraction was 4.3% with group sales (16% higher) and 6.4% (73% higher) when metal plates were attached to the well spout in case of purchase. We show the regression results in columns 1-2 of Table 4, where we estimate models as in equation (3). The difference in switching rates between B and A is 0.01 (95% C.I. -0.012-0.03), while the difference between C and A is 0.03 (95% C.I. 0.001-0.066). Consistent with Figure 4.1 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% level—for arm C. Overall, the unconditional switching rates are small in each arm, also reducing the statistical power when making between-arm comparisons, but this is in large part due to two factors. First, about three quarter of households did not purchase a test and, second, the large majority of wells were found to be safe. We thus turn to analyzing switching rates among households who purchased a test and found their well water to have unsafe arsenic levels.

In arm A, 30% of households (60/200) switched from unsafe wells. Although far from negligible, this figure is at the lower end of the range of switching rates observed in earlier studies, some of which documented rates above 2/3 (see Opar et al. 2007, Chen et al. 2007, Madajewicz et al. 2007, Bennear et al. 2013, Balasubramanya et al. 2014, George et al. 2012 and Inauen et al. 2014). On the one hand, this may appear surprising, given that in earlier studies the information had been provided for free, and so there was no self-selection into purchase of households that may have been expected to be relatively more responsive to information. On the other hand, free testing campaigns could have also led to larger switching rates by revealing a larger number of safe wells in the vicinity of unsafe wells. Moreover, in a number of such earlier studies tests were conducted in the context of intensive research efforts that may have contributed to a stronger response (Chen et al. 2007, Madajewicz et al. 2007).

Looking now at the switching rates observed with group sales (arm B) or with metal placards posted on the pump head (arm C), we found that rates in both were much higher relative to A. In B, 56% of households switched (89/160), while in C the rate was even higher, at 72% (68/95). In column 3 of Table 4 we show the corresponding regression results. Both estimated differences are large and significant at the 5% level or below, with $\hat{\beta}^B = 0.27$ (95% C.I. 0.061-0.484) and $\hat{\beta}^C = 0.46$ (95% C.I. 0.231-0.696). The difference in switching rates between B and 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 we include baseline controls (column 4). For both arms B and C the coefficients are almost identical (column 5) when we estimate the model without controls but including only the complete observations used in a adjusted regression in column 4. This suggests that the impacts are not substantively biased by differences in the level of observed confounders.

It is worth emphasizing how certain controls predict switching, although the results should not be interpreted causally, given that the covariates may be correlated with unobserved factors that also matter for the choice of water source. Most coefficients are small and not significant at standard level. Conditional on test purchase, households with a better educated head were *not* more likely to change the source of drinking water, a finding that contrasts with earlier work that evaluated switching behavior following arsenic testing offered at no cost, see Chen et al. (2007), Madajewicz et al. (2007), Pfaff et al. (2017). A notable—and perhaps disheartening—result is that prior beliefs about the well being safe reduced predicted switching by 11 percentage points, although the coefficient is not significant at standard levels. Beliefs about water safety may have thus been rather persistent for some households, despite the evidence offered by the tests. Recall that in our study the minimum arsenic level communicated to owners of 'unsafe' wells was 100ppb, so this finding was not due to respondents who thought that the well was unsafe but then learned that it was instead 'barely unsafe'.

Another concerning result is that, again conditional on the well being unsafe, higher levels of arsenic do *not* predict more switching. To the contrary, and using 100 as the omitted category, dummies for the arsenic level being equal to 200, 300 or 500/1000ppb are *negative* and in some case 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] In Figure 5 we show indeed that, with the exception of arm

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

B, switching rates as a function of the test result were well approximated by a step function jumping from about 0 for arsenic levels up to 50ppb to a larger and rather constant level for 'unsafe' arsenic level of 100 or above. Note also that very high arsenic levels are not rare. In our sample, less than 30% of tested wells were unsafe, but among those more than half had arsenic levels above 100 ppb and were therefore very unsafe.

In column 6, we also include as regressor a dummy for the presence of a safe well within 50 meters, where we define a neighboring well as safe when it was identified as such by our 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 we lose some observations due to errors in the geo-location of the wells. We also control for the total number of wells in a 50-meter radius, and we interact the dummy for safe wells with the treatment indicators. Among owners of unsafe wells in arm A we find that, as expected, having a safe alternative nearby increases switching. The coefficient is large (27 percentage points) and significant at the 1% 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 group signing or verbal commitment leading some households to share wells with other group members, with less concern of 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 our data do now allow us to determine with certainty if 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.

A limitation of our data is that we cannot gauge to what extent switching was associated to a reduction in arsenic risk, because we do not have records for the arsenic level of the new source of drinking water. Surveyors were asked to record the GPS location of the new source, but the smartphone's GPS sensor we used—with an approximate precision of 10m at best—was not sufficiently precise to identify uniquely the well, also due to the dense network of wells within the study area. Despite this, some useful information can be gleaned from households that changed the source of drinking water. Of the 217 'switchers', almost all (214) listed safety concerns as the primary reason for their decision. However, about one third of these (79/217) had switched to a different well which was itself perceived 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 our study area a degree of arsenic exposure remained even among a sizeable fraction of households who reacted to the new information by switching to a different water source for drinking and cooking.

5.3.1 Interpretation

These results confirm our prediction that group signing or metal placards lead to more switching relative to privately provided information. In this section we provide some evidence to support possible mechanisms behind the results, although we acknowledge that our results are tentative, and we cannot conclusively separate 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). In particular, our data face two limitations. First, we do not know which precise beliefs each household had about safe and unsafe wells in the village. We can thus only use data on beliefs about the wells actually used, and on the actual arsenic levels of all tested wells, but the latter information was not necessarily known to households. Second, in arm B, where tests were only sold to groups of buyers, our data only indicates whether a household purchased a test, but they do not indicate who formed a group with whom. We thus can 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 this limitation, our data at least 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 who did not want the information to be known to others, could have removed the placards, although detaching the metal wire holding them to the pumphead would have required some effort or a tool. However, we find that this behavior was rare. At the time of the return visits, the vast majority of the 348 placards installed on the well spout at the time of the test 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 our study.

Our data also indicates that the placards were more effective than result cards only at reminding users of the arsenic status of their well water, again suggesting that salience likely played a relevant role in the relatively high responses to information in arm C. At the time of the endline survey, we recorded again the beliefs about the safety of the well water. Although in some cases the respondent may have changed between baseline and endline survey, this allows to gauge the degree of learning that took place after being informed about the test result. Overall, we find that almost 90% of respondents correctly reported whether the water was found to be unsafe, but we also find that 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 them as such was 83% in arm A (135/162), 88% in B (122/139), and a remarkable 98% in arm C (83/85).

We also find some 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 of the group. Earlier we have explained that of the 217 households who stopped drinking from their unsafe well, only 88 (41%) had switched to a well that they believed to be safe. When we disaggregate by arm, we find that the fraction was 47, 27, and 53% in arms A, B and C, respectively. It is possible that also switchers from unsafe wells in B started drinking water from wells saf*er* that the original one (unfortunately we cannot check if this was the case), but that in C 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 our 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% 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 we look at the 124 observations with likely correct records we find that, while distances from the new source were almost identical in arms A and B (on average 68 and 80 meters, respectively, with the p-value of the difference = 0.518), the distance was substantially larger in C (190 meters, with the p-value of the difference with respect to A < 0.01).

In sum, we argue that the data suggest that placards (C) likely made households more aware of the risks associated to drinking from their own unsafe wells and allowed better choices, sometimes at the cost of longer distances traveled to fetch drinking water. In contrast, there is less that we can say about the mechanisms that made group sales (B) relatively successful, although the key factors delineated in our 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 who used their well for drinking and cooking at baseline. Perhaps not surprisingly, demand was significantly lower among 'non-users' (12%, vs. 25% among users), and for these households we did not record the main source of water for drinking. Switching behavior is thus harder to analyze, also because the sample is small. However, our records allow us to determine that many of these households reacted to 'good news' by switching to the well they were not initially using. In our sample, 139 non-users purchased a test, and of these 126 (91%) were re-interviewed at endline. Of these, exactly half (63) found out that their well was safe (As$\leq$ 50), and *all but 3* reported that they *were* using the well for drinking and cooking at the time of the endline.[18] In contrast, only nine of the 63 with unsafe wells reported that they were using the well. However, for them we do not know if the well they were using at baseline was found to have an arsenic level even higher than their own.

[18] 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, we do not analyze in detail the results by arm.

6 Cost Effectiveness

In this section we evaluate the cost-effectiveness of the different sale strategies. Because our 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), we gauge the merit of free provision as compared to our sales strategy by assuming a range of switching rates consistent with earlier studies.[19] In addition, recall that we cannot estimate reliably the change in arsenic contamination for 'switchers', given that we only observe the arsenic level of the initial source.

Using figures from our project, we assume that each test costs USD0.30 (or BDT24 using an exchange rate of 1USD/80BDT), and that personnel is paid BDT45 per test delivered, with an additional bonus of BDT12 for group sales. We also add BDT80 per test in arm C, to account for the cost of the placards. Consistent with our experimental results, we assume a take up rate of 25% in arms A, B and C, while consistent with results from earlier blanket testings we assume a 100% testing rate when tests are offered for free. Again using estimates from our RCT, we assume a 30% switching rate among users of unsafe wells in arm A, while for arms B and C we use the estimates adjusted for controls and strata fixed effects in column 4 of Table 4. That is, we assume that switching rates are 0.49 in arm B and 0.70 in arm C. In the case of free provision, we vary switching rates from 0.3 to 0.75, consistent with what found in earlier work that evaluated switching after free provision. In our study area, the fraction of unsafe wells varied in the 16-27% range, while the earlier BAMWSP figures in these same villages varied from 40 to 90%. To cover a wide range of possibilities we thus provide calculations using a fraction of unsafe wells that is either low (20%), or medium (40%) or high (80%). We summarize the results in Table 5, assuming that a policy maker is deciding how best to allocate a total and fixed budget of USD10,000.

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. 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

[19] Jamil et al. (2019), using data from blanket free testing in Araihazar, estimates a total cost of <USD1 per person whose exposure was reduced.

observed in arm A. Note that in the table above 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 when the fraction of unsafe wells increases, because this leads to an increase in the number of households who may benefit from switching, while the fraction that does is not affected under our assumption.

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 USD1.15 (group sales with high prevalence of unsafe wells) to USD14.4 (free provision with low switching rates and low prevalence of unsafe wells). However, Pitt et al. (2015) estimated a present discounted value of per-household gains from switching to safe water sources over twenty years ranging from USD1400 to USD1000 for discount rates of 3% to 8%. 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 10ppb. In addition, Keskin et al. (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.[20] 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 quarter of households not to learn about the safety status of their well. For instance, even under the scenario of only 20% 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 who are isolated, either geographically or socially.

7 Discussion and Conclusions

Information on household-specific environmental health risks can be a relatively inexpensive policy tool, but the design of information campaigns often has to contend with low demand and 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

[20] There are few estimates of the value of a statistical life (VLS) for Bangladesh, and their range is very broad. For instance Mahmud et al. (2019), using data on WTP 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 ten times larger.

may severely limit the effectiveness of such campaigns, especially if targeted information is only supplied for a fee. These considerations are salient in Bangladesh, a country where millions of people use water from shallow tubewells for drinking and cooking, and where a large fraction of such water is estimated to be contaminated by naturally occurring arsenic in concentrations high enough to have extremely deleterious health consequences in case of long-term exposure. This is generating one of the most severe public health crisis 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 to allow households to undertake risk-avoiding behavior, by simply changing their primary source of drinking water.

In this paper we have described the results from a randomized field experiment where we study the effect of different arsenic test selling schemes on test uptake and well switching. Despite the fees our team of surveyors managed to sell more than 2,800 tests for a total of about 11,400 wells. This allowed to uncover the presence of hundreds of wells with arsenic levels above the threshold adopted by the Government of Bangladesh, and overall about half of the users decided to switch to a different source of drinking water. We show that relatively subtle differences in the way information was sold and provided, while barely affecting demand, led to very substantial gaps in behavioral responses: both group sales that leveraged informal local solidarity networks, and the addition of metal placards posted on the wells more than doubled the fraction of users of unsafe wells that reported having switched to a different water source at the time of our return visits, relative to simple, individual sales where test results were provided privately to the buyers. These findings should be useful for the design of information campaigns that aim at providing measures of risk exposure that vary at the household level. In our context, information was supplied for a fee only to household who chose to purchase a test, but we conjecture that similar considerations will likely be relevant also 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 be however emphasized. First, data limitations do not allow us to conclusively disentangle the mechanisms underlying the results, 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 tubewell spouts.

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

Third, our trial did not include an arm where tests were provided free of charge, and so we need to resort to a number of assumptions when making cost-effectiveness comparisons between sales and free provision. With this caveat, our results suggest that cost sharing under some scenarios could allow to 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.

Fourth, to the extent that our results can be extrapolated to the rest of the country, we have shown that tests-for-fee campaigns can only provide a partial solution to the public health crisis due to arsenic in shallow aquifers. In our study area, about three quarters of wells remained untested, despite the fact that in a large majority of cases the users had no idea about the safety of the water they routinely used for drinking and cooking. Unfortunately, given that we know the arsenic status only for tested wells, we do not observe the total number of unsafe wells in each arm, which would be a good approximation of the total number of well owners needing mitigation of arsenic risk. However, recall that treatment assignment was stratified based on geographical area and fraction of unsafe wells as estimated years earlier by the BAMWSP blanket testing campaign, so we would expect the fractions of unsafe wells to be approximately the same among treatment arms. Indeed, the fraction of wells that turned out to be unsafe upon testing was broadly similar across arms, ranging from 16% in B to 27% in C. Given that the large majority of households did not know the safety of their well, it is probably safe to assume that the actual prevalence of safe wells was similar to these figures in all experimental arms, and therefore in the 15-30% range. Taken together, and given that the unconditional ITT estimates of switching rates were in the 3-6 percentage point range, these back-of-the-envelope estimates suggest that, despite the many tests sold, switching rates achieved by our test sales program remained well below the likely fraction of unsafe wells.

Fifth, the low demand also raises concerns on the sustainability of a test-for-fee selling scheme at this price. On an average work day, we calculate that surveyors visited 25 well owners to offer As tests (with surveyor-specific averages ranging from 12 to 36 visits), but only sold six 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 by selling 15 tests a day. The low demand thus implies that the actual wage fell short of the expected one.

Sixth, 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 sub-district, where Pfaff et al. (2017) document that following the BAMWSP blanket testing campaign, about 30% of households whose well water was found to be unsafe had switched to other wells identified as unsafe or of unknown status.

Seventh, our study is silent as to whether demand was limited by our 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. We also cannot exclude that 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).

Although our study did not include an experimental arm where *all* wells were tested, the relatively low behavioral responses were likely at least in part due to the fact that fees, by causing many wells to remain untested, substantially reduced the set of safe alternatives available for many households relative to what could have been achieved with blanket testing. In other words, although exposure to arsenic is not an infectious disease, there are clear positive externalities in the decision to test a well, and given the low demand observed even at very low prices this may be another case where free provision may be the optimal policy strategy (Cohen and Dupas 2010), at least if budget limitations allow it.

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. Our results 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. However, and although this cannot be gauged directly from our analysis, we also conjecture that these strategies may be best adopted while providing tests for free and disseminating widely information on safe sources. Given the magnitude of the public health problem in Bangladesh, this would require significant investments from the government or from donors, but free provision would also avoid screening out individuals with low ability to pay, and it would possibly facilitate switching decisions by increasing the number of viable safe options.

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*, 1687–1988.

Alsan, M. and M. Wanamaker (2017). Tuskegee and the health of black men. *Quarterly Journal of Economics 133*(1), 407–455.

Argos, M., T. Kalra, P. J. Rathouz, Y. Chen, B. Pierce, F. Parvez, T. Islam, A. Ahmed, M. Rakibuz-Zaman, R. Hasan, G. Sarwar, V. Slavkovich, A. van Geen, J. Graziano, and H. Ahsan (2010). Arsenic exposure from drinking water, and all-cause and chronic-disease mortalities in Bangladesh (HEALS): a prospective cohort study. *Lancet 376*, 252–258.

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), 2383–2413.

Bai, L., B. Handel, E. Miguel, and G. Rao (2017, August). Self-control and demand for preventive health: Evidence from hypertension in India. Working Paper 23727, National Bureau of Economic Research.

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), 631–647.

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*, 225–240.

Bennett, D., S. A. A. Naqvi, and W.-P. Schmidt (2017). Learning, hygiene, and traditional medicine. *The Economic Journal Forthcoming.*

Carson, R., P. Koundouri, and C. Nauges (2011). Arsenic mitigation in Bangladesh: a household labor market approach. *American Journal of Agricultural Economics 93*(2), 407–414.

Chen, Y., J. H. Graziano, F. Parvez, M. Liu, V. Slavkovich, T. Kalra, M. Argos, T. Islam, A. Ahmed, M. Rakibuz-Zaman, R. Hasan, D. Levy, A. van Geen, and H. Ahsan (2011). Arsenic exposure from drinking water and cardiovascular disease mortality: A prospective cohort study in Bangladesh. *British Medical Journal 342*, 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), 917–923.

Chowdhury, A., M. Hossain, R. Nickson, M. Rahman, M. Jakariya, and M. Shamimuddin (2000). Combating a deadly menace: Early experiences with a community-based arsenic mitigation project in Bangladesh. *BRAC Research and Evaluation Division, Research Monograph Series* (16).

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. *The Lancet 382*(9906), 1734–1745.

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), 609–645.

Dupas, P. (2012). Health behavior in developing countries. *Annual Review of Economics 3*, 425–449.

Dupas, P. (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 (2016). Impacts and determinants of health levels in low-income countries. NBER Working Paper 22235.

Fafchamps, M. (2011). *Risk Sharing Between Households*, Volume 1A of *Handbook of Social Economics*, Chapter 24, pp. 1255–1280. 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, K. M. Ahmed, J. Moon-Howard, A. Tarozzi, X. Liu, P. Factor-Litvak, and J. Graziano (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), 11213–11219.

Goldberg, J., M. Macis, and P. Chintagunta (2018). Incentivized peer referrals for tuberculosis screening: Evidence from India. NBER Working Paper No. 25279.

Gong, E. (2015). HIV testing and risky sexual behaviour. *The 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), 346–355.

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), 701–719.

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), 5596–5604.

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*, 477–483.

Keskin, P., G. K. Shastry, and H. Willis (2017). Water quality awareness and breastfeeding: Evidence of health behavior change in Bangladesh. *The Review of Economics and Statistics 99*(2), 265–280.

Kremer, M. and E. Miguel (2007). The illusion of sustainability. *Quarterly Journal of Economics 122*(3), 1007–1065.

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), 731–754.

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 (2017). 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), 4171–4185.

Miller, G. and A. M. Mobarak (2013). Intra-household externalities and low demand for a new technology: experimental evidence on improved cookstoves. National Bureau of Economic Research Working Paper no. 18964.

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), 164–172.

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), 331–339.

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. Working Paper 21741, National Bureau of Economic Research.

Pop-Eleches, C., H. Thirumurthy, J. P. Habyarimana, J. G. Zivin, M. P. Goldstein, D. De Walque, L. Mackeen, J. Haberer, S. Kimaiyo, J. Sidle, 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*, 131–137.

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), 1093–1103.

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), 1909–1941.

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), 231–235.

Thornton, R. L. (2008). The demand for, and impact of, learning HIV status. *American Economic Review 98*(5), 1829–1863.

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 25^2km area of Bangladesh. *Science of the Total Environment 488*, 484–492.

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), 632–638.

van Geen, A., H. Ahsan, A. Horneman, R. Dhar, Y. Zheng, I. Hussain, K. Ahmed, A. Gelman, M. Stute, H. Simpson, S. Wallace, C. Small, F. Parvez, V. Slavkovich, N. Lolocano, M. Becker, Z. Cheng, H. Momotaj, M. Shahnewaz, A. Seddique, and J. Graziano (2002). Promotion of well-switching to mitigate the current arsenic crisis in Bangladesh. *Bulletin of the World Health Organization 80*(9), 732–737.

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), 1729–1740.

Viscusi, W. K. and C. J. Masterman (2017). Income elasticities and global values of a statistical life. *Journal of Benefit-Cost Analysis 8*(2), 226–250.

Wasserman, G. A., X. Liu, F. Parvez, H. Ahsan, P. Factor-Litvak, J. Kline, A. van Geen, V. Slavkovich, N. J. LoIacono, D. Levy, Z. Cheng, , and J. H. Graziano (2007). Water arsenic exposure and intellectual function in 6-year-old children in Araihazar, Bangladesh. *Environmental Health Perspectives 115*(2), 285–289.

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.

Figure 1: Metal Placards

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

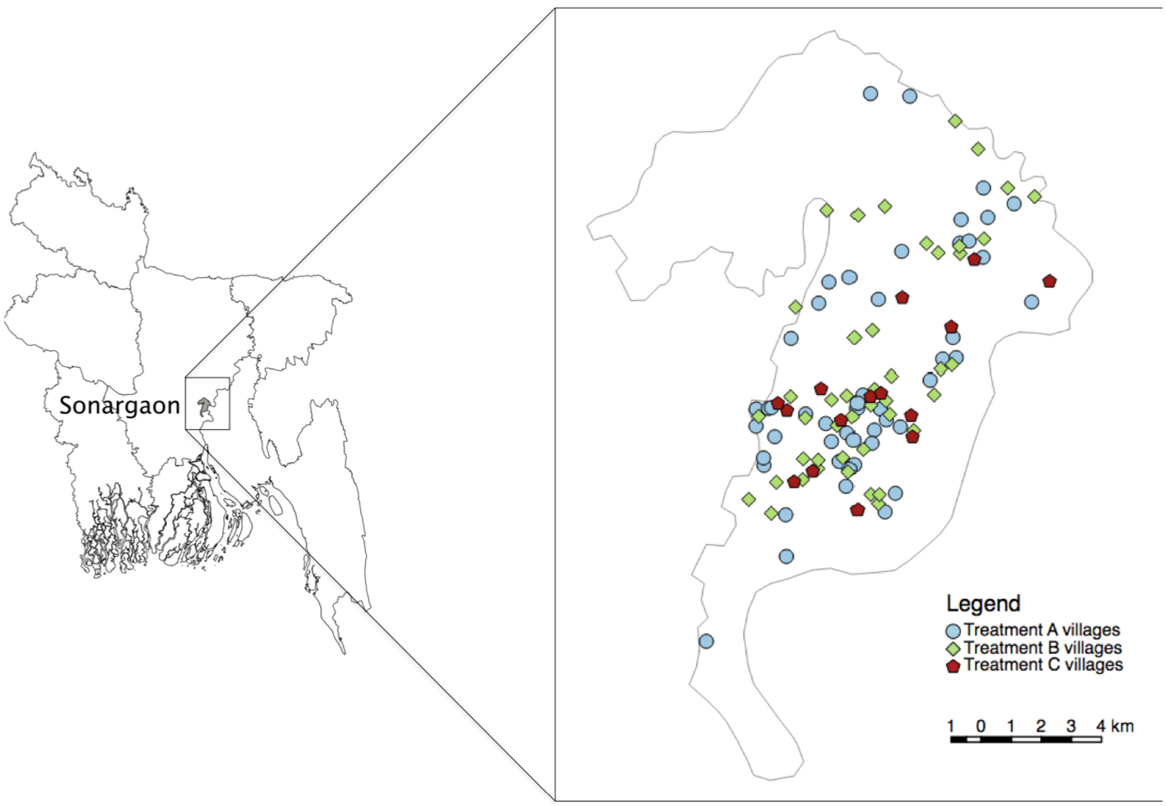


Figure 2: Study area and treatment assignment

Notes: Author's illustrations from the geo-location of study villages recorded at baseline. Each village is placed at the mean latitude and longitude of all well-owners interviewed at baseline in the village.

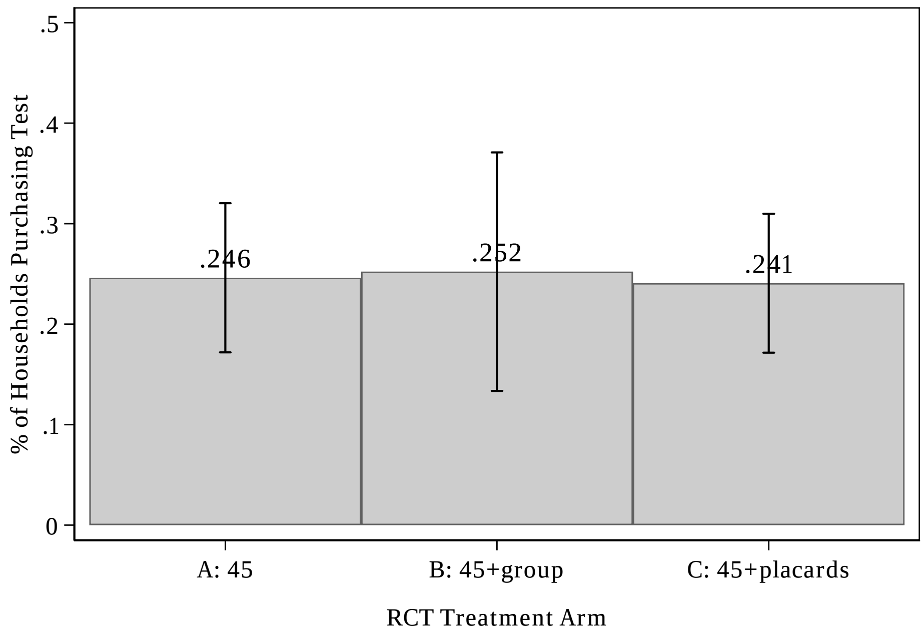
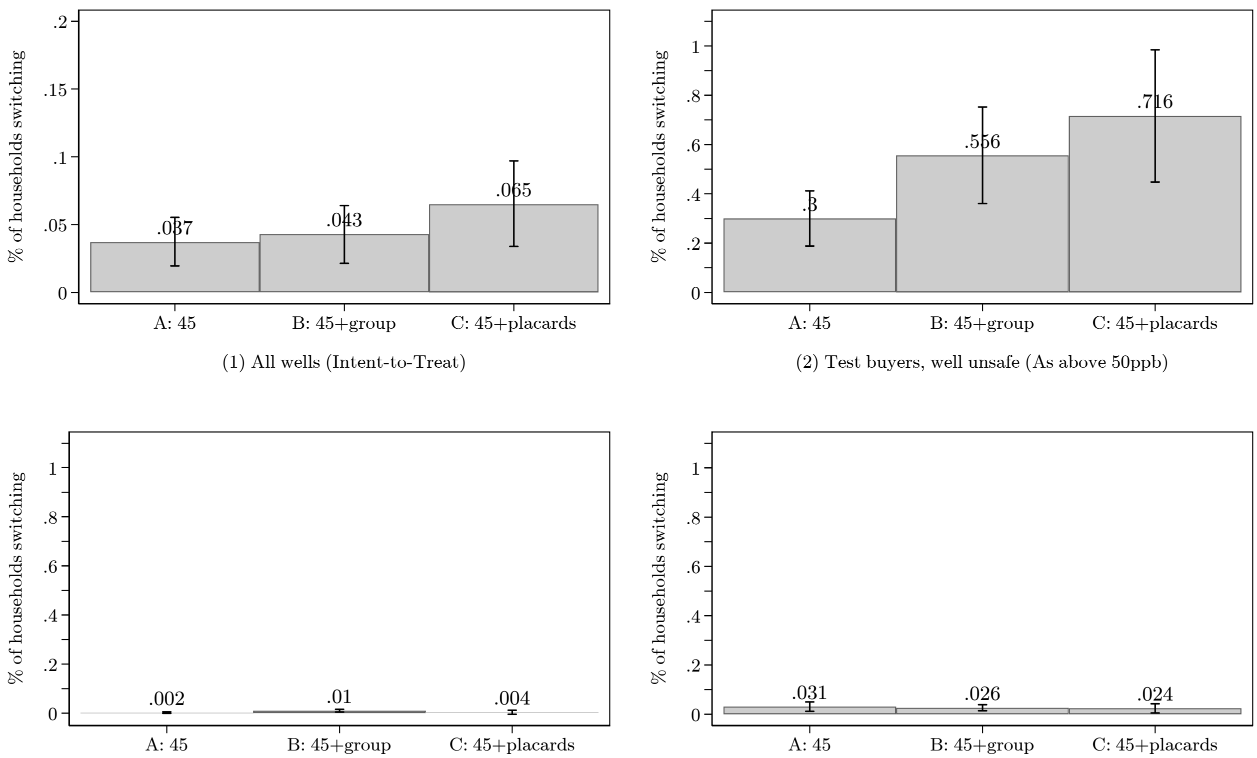


Figure 3: Demand for Tests of Arsenic Concentration in Well Water

Source: Authors' estimations from baseline data (January to June 2016). Each bar is labeled with the arm-specific purchase rate. The vertical intervals represent 95% 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).



(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

Figure 4: Switching rates

Source: Authors' estimations from endline data (August 2016 to January 2017). Each figure shows the fraction of households who 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% 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: $n_A = 3639$, $n_B = 3252$, $n_C = 1104$.

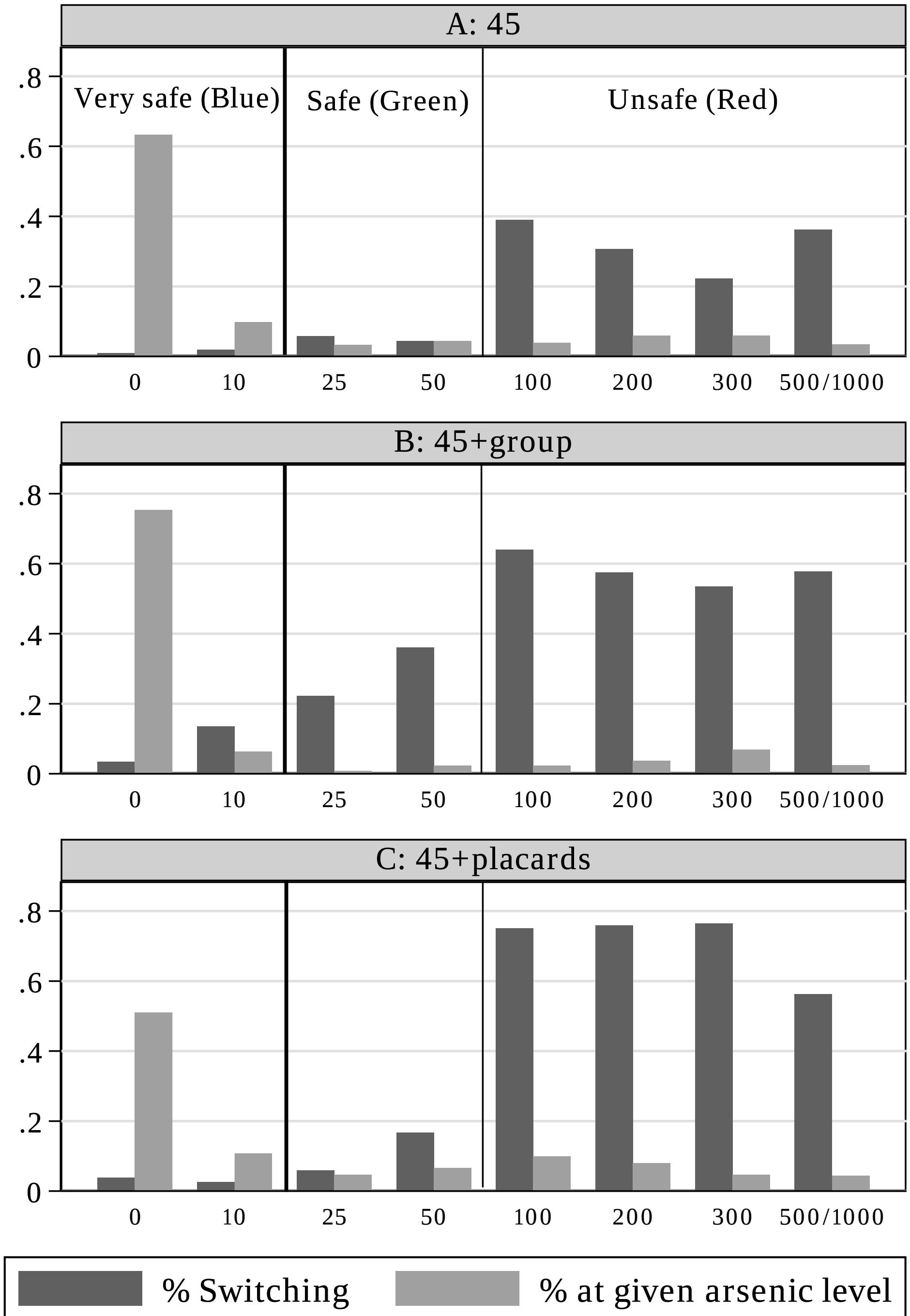


Figure 5: Switching rates from tested wells, by arsenic level and experimental arm

Source: Authors' estimations from baseline (January to June 2016) and endline (August 2016 to January 2017) data. The figures show, for each experimental arm and for tested wells, the prevalence of each arsenic level as identified by the test (light grey bars, in ppb, or micrograms per litre) and the fraction of households who were no longer using the well at endline (dark grey bars). The field tests identified the arsenic level as a value in the set $As \in \{0, 10, 25, 50, 100, 200, 300, 500, 1000\}$. The values on the horizontal line are not drawn at scale. Results of As= 1000 were rare and hence we pooled 500 and 1000 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.

Table 1: Baseline Summary Statistics and Balance across Treatment Arms

	(1)	(2) Overall	(3)	(4)	(5) Means by experimental arm	(6)	(7) Tests of equality (p-values)	(8)	(9) Standardized differences	(10)
	Obs.	Mean	St.Dev.	A	B	C	H_0 A=B=C	A/B	A/C	B/C
Drink from well at baseline	12603	0.905	0.293	0.930	0.884	0.891	0.2520	0.161	0.139	-0.022
Household head is male	11410	0.848	0.359	0.843	0.850	0.860	0.9210	-0.018	-0.047	-0.028
Household head wage worker	10890	0.285	0.452	0.247	0.298	0.378	0.4260	-0.115	-0.287	-0.171
Household head self-employed	10890	0.419	0.493	0.440	0.438	0.285	0.3390	0.004	0.327	0.323
Household head no schooling	10890	0.193	0.395	0.262	0.163	0.054	<0.001***	0.245	0.596	0.355
Household head primary school	10890	0.328	0.470	0.317	0.343	0.319	0.7900	-0.055	-0.005	0.05
Heard about As in well water	11410	0.996	0.062	0.996	0.996	0.997	0.7430	-0.007	-0.029	-0.022
Aware of health risks of As	11410	0.998	0.044	0.997	0.998	0.999	0.2950	-0.018	-0.047	-0.031
House has concrete roof	11252	0.173	0.378	0.175	0.184	0.133	0.4060	-0.024	0.117	0.142
Household members	11045	3.600	1.300	3.630	3.570	3.590	0.8270	0.048	0.037	-0.014
Number of Children	11045	1.460	1.040	1.480	1.430	1.470	0.6920	0.054	0.016	-0.041
Well As status unknown (belief)	10515	0.758	0.428	0.684	0.789	0.903	0.0005***	-0.242	-0.562	-0.319
Well As status unsafe (belief)	10515	0.069	0.253	0.097	0.042	0.057	0.2590	0.214	0.151	-0.066
Well As status safe (belief)	10515	0.173	0.379	0.220	0.168	0.041	<0.001***	0.13	0.553	0.427
Well labeled safe	10515	0.002	0.040	0.003	0.001	0.001	0.1020	0.049	0.05	0.002
Wells within 50m	10260	12.300	15.600	10.400	14.500	12.100	0.2530	-0.239	-0.222	0.14
Wells within 50m labeled safe	10260	0.015	0.128	0.028	0.003	0.007	0.0302**	0.19	0.157	-0.046
Share unsafe wells (BAMSWP)	11410	0.746	0.132	0.759	0.720	0.782	0.4120	0.295	-0.193	-0.46
Well is privately owned	11410	0.986	0.117	0.980	0.991	0.992	0.1690	-0.087	-0.103	-0.018
Well depth ($\times$100 feet)	11410	1.790	1.080	1.830	1.770	1.730	0.8800	0.054	0.096	0.045
Well age (years)	11410	9.130	7.570	8.860	9.490	9.000	0.0105**	-0.084	-0.019	0.062
Well cost ($\times$10000 BDT)	11410	0.756	0.642	0.763	0.764	0.707	0.7340	-0.001	0.119	0.095
Persons drinking from well	11343	8.840	11.100	8.880	8.490	9.750	0.5710	0.036	-0.071	-0.101
Attrition	11410	0.088	0.283	0.094	0.089	0.063	0.3150			
Lost after baseline	11410	0.055	0.228	0.058	0.055	0.045	0.5990			
Duplicate I.D. at baseline	11410	0.032	0.177	0.036	0.033	0.018	0.4170			
							Multivariate standardized differences	0.604	1.103	0.720

Notes: Author's calculations from baseline data (January to June 2016). The unit of observation is the primary household attached to a specific well. The number of clusters (villages) in the five 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 household who used the specific well for cooking and drinking at baseline. Differences in the number of observations across these variables are explained by missing entries during the data collection. The p-values in column 7 are for tests of the null of equal means across treatment arms (robust to intra-village correlation). Asterisks denote test significance: *** p<0.01, ** p<0.05, * p<0.1. The normalized differences in columns 8-10 are calculated as in equation (1), while the multivariate standardized differences in the last row are calculated as in equation (2), see Section 3.1 for details.

Table 2: Demand for tests

	(1)	(2)	(3)	(4)
	Dependent variable:			
	Indicator = 1 if household purchased test			
B: 45+group	0.006	-0.004	-0.012	-0.001
	(0.071)	(0.037)	(0.018)	(0.018)
C: 45+placards	-0.005	0.034	0.030	0.059*
	(0.051)	(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 50m			-0.003***	
			(0.001)	
No. of visibly safe wells within 50m			0.006	
			(0.037)	
Fraction unsafe wells in village (BAMWSP)			0.093	
			(0.086)	
Well depth ('00 feet)			0.033***	
			(0.009)	
Well age (years)			-0.001	
			(0.001)	
Well cost ('0000 BDT)			0.027*	
			(0.014)	
Believes well is safe			-0.115***	
			(0.018)	
Believes well is unsafe			-0.049***	
			(0.015)	
Observations	11,410	11,410	8,892	8,892
R-squared	0.000	0.130	0.111	0.040
Controls	No	No	Yes	No
Strata FE	No	Yes	Yes	Yes
Mean in A	0.246	0.246	0.246	0.246
Clusters	112	112	102	102

Source: Authors' estimations from baseline data (January to June 2016). The dependent variable is binary and is = 1 if the household purchased the test at baseline. All regressions are estimated with OLS. Regressions with strata fixed effects include union fixed effects and a dummy = 1 in villages where the % of unsafe wells in the village (estimated by BAMWSP) was below the median. Standard errors are clustered at the village level. The smaller sample size in column 3 relative to columns 1-2 is due to missing values in one or more controls, while in column 4 we do not include controls but we only use observations with complete observations. Significance: *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Number of wells by safety status and switching decision

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		Tested						Not tested	
				Safe		Unsafe			
				Switched		Switched			
	Total wells used for drinking and cooking	Total	Unsafe (proportion)	Yes	No	Yes	No	Total	Switched Yes
A: BDT45	4,679	1,040	0.192	2	838	60	140	3,639	113
B: BDT45+Group	4,281	1,029	0.155	9	860	89	71	3,252	85
C: BDT45+Placards	1,452	348	0.273	1	252	68	27	1,104	26
Total	10,412	2,417	0.188	12	1,950	217	238	7,995	224
Tests of equality (p-value)									
H_0 : A=B=C			0.3174						

Notes: Authors' calculations using information from a total of 10,412 wells that were used at baseline for drinking and cooking purposes. We exclude from the analysis 768 wells used by households that could not be re-contacted at endline, and 406 wells with a duplicate ID at baseline which can thus not be matched to endline data on switching decisions.

Table 4: Choice of water source at endline

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: Indicator = 1 if switcched (no longer uses same well used at baseline)					
	All wells (ITT)		Conditional on purchase and unsafe well (As> 50ppb)			
B: BDT45+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: BDT45+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 ('00 feet)		-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 ('0000 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 = 200ppb				-0.039		0.012
				(0.062)		(0.066)
As = 300ppb				-0.151**		-0.157**
				(0.061)		(0.066)
As = 500 or 1000ppb				-0.148*		-0.114
				(0.077)		(0.084)
Number of wells within 50m						-0.002
						(0.005)
There is at least one safe well within 50m						0.270***
						(0.085)
B × at least one safe well within 50m						-0.188
						(0.123)
C × at least one safe well within 50m						-0.076
						(0.121)
Observations	10,412	9,385	455	407	407	355
R-squared	0.012	0.028	0.169	0.261	0.176	0.298
Controls	No	Yes	No	Yes	No	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes
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

Notes: Authors' estimations from baseline and endline data. The Intent-to-Treat results in columns 1 and 2 show switching rates *not* conditional on purchase or test result, including all households who used the well at baseline and who could be matched between baseline and endline surveys. 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). In column 2 (relative to column 1), and in column 4 (relative to 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 location was 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. All regressions are estimated using a linear probability model where the dependent variables is a dummy equal to one if the well was no longer used for cooking and drinking at endline. Standard error are clustered at the village level. Asterisks denote statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Cost Effectiveness of Different Sale Strategies

		Demand (%)	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 sales+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

The estimate show the responses to different sale strategies, assuming a total budget of USD10,000. Each test is assumed to cost USD0.30 (BDT24), while testers are assumed to be paid BDT45 per test delivered, with an additional bonus of BDT12 for group sales. BDT80 per test are added in arm C, to account for the cost of the metal placards. We assume a take up rate of 25% in arms A, B and C, and 100% when tests are offered for free. Switching rates among users of unsafe wells are assumed to be 30%, 49% and 70% in arms A, B, and C, respectively. In the case of free provision, we vary switching rates from 0.3 to 0.75, consistent earlier studies in neighboring areas.