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

Ricardo Maertens
Alessandro Tarozzi
Kazi Matin Ahmed
Alexander van Geen*

February 2018

Abstract

Lack of access to reliable information on environmental exposure limits opportunities for risk-avoiding behavior, particularly in developing countries. Private markets could potentially play a role in providing such information if households cannot rely on the public sector. We describe results from a randomized controlled trial conducted in Bangladesh to determine, first, to what extent charging a modest fee for an environmental test limits demand and, second, how different modes of information delivery can affect demand for and behavioral responses to such information. In rural Bangladesh, millions of individuals are chronically exposed to arsenic by drinking contaminated water from their private well. Well tests for arsenic have previously been shown to encourage a sizeable fraction of households owning an unsafe well to switch to a safer source that, typically, is within walking distance. However, the safety status of millions of tube wells remains unknown and there is no well-established market for well tests. In our study, tests were sold under different conditions during field trials conducted in 128 villages of Sonargaon sub-district, Bangladesh. At a price of BDT45 (about USD0.60), only one in four households purchased a test, despite widespread awareness about the risk of exposure to arsenic but little knowledge about the safety status of wells. Sales were not increased by "nudges" in the form of visible metal placards indicating safety status or by offers to promote the sharing of safe wells through informal agreements. However, conditional on learning about the unsafe status of one's tube well water, both informal agreements and visible placards nearly doubled the fraction of households that switched away from their contaminated well. Large increases in demand and behavioral responses were also observed with sales that required payment only in case of "good news", although this contract variation was only adopted in a subset of only 12 villages. We conclude that households should not be charged for well tests but considerably more attention should be paid to the way tests results are delivered to encourage the sharing of safe wells.

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

---

*We gratefully acknowledge financial 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 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 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, Venezia Cà Foscari, CEPREMAP India-China Conference (PSE, Paris), the VI Navarra Center for International Development (Fundación Ramón Areces, Madrid), Oxford (CSAE), Stanford, 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.
Maertens: Department of Economics, Harvard University; Tarozzi (corresponding author), Department of Economics and Business, Universitat Pompeu Fabra, Barcelona GSE and CRES, alessandro.tarozzi@upf.edu; 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 features 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 other health policy, because they can be relatively inexpensive to run when compared to other options. In addition, some health conditions are in principle easily preventable if appropriate behavior is adopted to avoid environmental exposures. On the other hand, 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 anyway 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 study, first, demand for information on the quality of drinking water and, second, the adoption of risk-mitigating behavior conditional on the information received. 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 As 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 As from drinking tube well water in South Asia, such as cancerous skin lesions and loss of limb, were recognized in the state of West Bengal, India in the mid-1980s (Smith et al. 2000). It has since been shown on the basis of long-term studies in neighboring Bangladesh that As exposure increases mortality due to cardiovascular disease and several forms of cancer, 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 As 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 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 As exposure because of the cost and logistics of maintaining such systems in rural South Asia (Ahmed et al. 2006; Howard et al. 2006; Sanchez et al. 2016). In contrast, and despite being the main source of As exposure, tube wells may also offer an effective way of providing safe drinking water to the rural population of Bangladesh. With the exception of the most severely affected areas of Bangladesh, the spatial distribution of high- and low-As wells is highly mixed, even over small distances. At the same time, whether a well is contaminated with As or not rarely changes over time (van Geen et al. 2007; McArthur et al. 2010). Therefore, exposure among users of As-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). Previous studies in Bangladesh have documented switching rates from an unsafe to a safe well after testing of between one-third and three-quarters, depending on how much effort was put into persuading users to do so (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).

Well-sharing as an effective risk mitigation strategy, however, relies on knowledge about the safety of both one's own water source and that of neighbors. Between 1999 and 2005, the World Bank, DANIDA and UNICEF, in collaboration with other NGOs and the Bangladesh Department of Public Health Engineering (DPHE), conducted a blanket testing campaign that 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. Unfortunately, such a testing campaign, coordinated through the Bangladesh Arsenic Mitigation and Water Supply Program (BAMWSP), has not been repeated. Meanwhile, millions of new wells have sprouted in the country, and in most cases users do not know the As level of the water. There are a few commercial laboratories in Dhaka with the capability to test wells for As, but few rural households are aware of these services. 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, and the willingness to pay for them is not known.

The first objective of this paper is to make progress in understanding the sustainability of a market for As field tests, by offering them at a price of BDT45 (about USD0.60) in 49 randomly selected villages in Sonargaon sub-district. Such amount was close to the price of one kg of rice in Dhaka, and was deemed to be high enough to cover for the salary of the surveyors hired for the project. The results contribute to a growing literature that studies demand for health-protecting technologies in developing countries. Prior research in different countries has documented very low uptake of a variety of such products, ranging from insecticide-treated nets (Cohen and Dupas 2010, Dupas 2014b, Tarozzi et al. 2014), to de-worming drugs (Kremer and Miguel 2007) and water-disinfectant (Ashraf et al. 2010). Our work complements this literature by looking at demand for health-related *information* that can be exploited by households to devise risk mitigation strategies. In Bihar, India, another location with a severe As problem, Barnwal et al. (2017) estimated very high elasticity of demand, with uptake of tests falling from 69% to 22% of households when the price increased from INR10 to INR50, where the latter was about equivalent to daily per capita income. They also found that sales repeated two years later were substantially higher, suggesting that experience with the tests were an important factor. In our study area, and despite low prices and widespread awareness about the As problem coupled with little information about the safety status of individual wells, we found that only about one in four households purchased the test. In this arm, conditional on learning about the unsafe status of one's drinking water, we estimate that 30% of households stated that they switched to a different source at the time of our return visit.

The second and more novel objective of the paper is to determine whether the existence of informal, within-village solidarity networks could be leveraged to increase demand for testing and especially well-sharing. A large literature documents the importance of village networks to cope with shocks, including health shocks, see Fafchamps (2011) for a review. In our case, and in an additional subset of 48 villages, we study demand for As tests—at the same price of BDT45—to self-formed *groups* of up to ten individuals, where group members were asked to sign an informal agreement according to which those with safe wells would share their well with others 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 increasing the salience of safe options within the village. Although we find that demand for tests remained about the same as with sales that did not involve informal water-sharing agreements, we estimate that switching rates relative to individual sales almost doubled from 30 to 56%, while the 95% confidence interval of the difference-in-differences (DD) adjusted for baseline covariates is $0.012 - 0.392$. Although the result is striking, data limitations do not allow us to probe conclusively whether the increase in switching was driven by sharing within groups of buyers.

In an additional subset of 15 villages, individuals who purchased a test at the usual price were also given a metal placard of a color depending on the As level: blue for As below 10ppb, green if between 10 and 50, and red if 'unsafe', that is, above the Government threshold of 50ppb. The placard was then attached to the well spout.

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. 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 the test result. While we find that demand for testing was barely affected by the concurrent offer of a metal placard, the switching rates more than doubled (from 30 to 72%, 95% C.I. of adjusted DD 0.165-0.553) relative to those recorded when test results were communicated individually via a simple laminated card.

Additionally, we describe the findings from two other experimental arms, although the small sample size (six villages per group) means that estimates are imprecise. Conversations during focus groups with locals indicated that respondents disliked the idea of paying for tests that may have delivered 'bad news', especially when these are not accompanied by the offer of clear and effective mitigation strategies. Such resistance to receiving a potentially negative update about one's health circumstances is also consistent with behavior observed in other contexts. For instance, in Malawi Thornton (2008) found that the fraction of individuals tested for HIV, at a time when anti-retroviral therapies were not yet widely available, testing prevalence remained low if incentives were not provided. Refusal rates for HIV testing remain a potential problem for the correct estimation of prevalence in population surveys (Janssens et al. 2014). To probe these issues further, in six villages we offered the As tests at the same price of BDT45 but with the agreement that payment would be requested only in case of 'good news', i.e. a low As test result. This conditional payment mechanism also had the additional potential benefit of providing stronger incentives to test for individuals who thought their well was likely unsafe. This may have helped targeting more at-risk individuals (as long as beliefs were correlated with actual safety status). In this regard, the results contribute to a literature that studies mechanisms to reduce wasteful spending in public health interventions (Dupas 2014a). We find that conditional pricing more than doubled demand relative to the simple offer (58 vs. 25 percent), although of course the comparison between the two sale mechanisms is confounded by the effective reduction in the 'perceived price' induced by the 'pay-for-good-news' contract, at least for everyone except the rare individuals who felt completely sure that their water was safe. The difference in demand is so large to be significant even at the 1% level (95% C.I. of adjusted DD 0.225-0.449), despite the small number of villages in this experimental arm. Even when we doubled the price at BDT90 in another group of six villages, demand remained substantially larger (95% C.I. of adjusted DD 0.031-0.200). We also find the this type of contract significantly increased the proportion of unsafe wells among those being tested.

The paper proceeds as follows. In the next section we provide some 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 overall substantively 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 in Section 5. Finally, we conclude in Section 6.

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 squared kilometers territory. Sonargaon is located in a part of the country where arsenic contamination of shallow tube well water is widespread. According to a blanket testing conducted between June 1999 and June 2000, under the supervision of the Bangladesh Arsenic Mitigation and Water Supply Program (BAMWSP), in about 90 percent of villages 40 percent or more of tube well 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).[1]

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 the BAMWSP blanket testing campaign. In each study village, surveyors would walk across the whole area identifying all wells, regardless of their ownership status or of evidence of having been tested before. For privately owned wells, the surveyors would identify which household owned the well, while for public wells the surveyors would identify which household was the main caretaker or user. Surveyors then conducted a home visit, where they would explain the risk of consuming arsenic-contaminated tube well water to an adult—typically the most senior woman—and then 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, tube well water was tested using Arsenic Econo-Quick (EQ) test kits, which have 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 result (in ppb), is rounded to the nearest integer in the following sequence $\{0, 10, 25, 50, 100, 200, 300, 500, 1000\}$. 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 tube well spouts (van Geen et al. 2014).

Surveyors would also administer a short household questionnaire and distribute 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 tube well water for arsenic is important, and (iv) that the Bangladesh safety standard for arsenic concentration in water is 50ppb. Black cards were given to households who did not buy a test, while in case of test purchase the laminated card was blue (if As was 10ppb or below), green (if it was 25 or 50ppb), or red (if the test showed an As 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.

[1] Blanket testing in Sonargaon was carried out by BRAC, a partner NGO of BAMWSP. A total of 25,048 tube wells were tested for arsenic, although the test kits utilized did not allow to measure precisely As in the 0-100ppb range.

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 tube well water tests across villages. In a first group of villages, which we term group A, surveyors offered to test tube well water for a fee of BDT45 (about USD0.60). This fee was expected to cover the salary of the testers and their supervisor. In particular, 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 (USD115/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 (USD578-867), 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 (USD0.35/test) was covered by the project.

In a second group of villages (B), groups of at most 10 neighbors were gathered and asked (i) if they wanted to test their water for a fee of BDT45 each and (ii) if they wanted to sign an agreement to share any safe water among the signatories—before conducting any testing. The agreement had no legal standing, and was meant to serve as a soft commitment device. Anecdotal information from the field indicated that buyers were sometime uncomfortable about signing a document, in which case a verbal agreement took place instead. All neighbors that committed to sharing their tube well water were allowed to see the test results from all other neighbors in the group. Both well owners that committed to share their well (formally or verbally) and those who did not could purchase the test.

In a third group of villages (C), households were offered to test tube well water for BDT45. But, for households that bought the test, 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). Further, as shown in Figure 1, they displayed two hands holding drinking cups, one hand holding a drinking cup, and a large cross over a hand holding a drinking cup, depending on As 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 to testers per every household that signed the well-sharing agreement or that verbally committed to sharing their well within the group of buyers.

In addition, our experimental design included two arms, with only six villages each, where payment for a test was only required if the water tested safe. This 'conditional fee' was set at BDT45 (arm D) or BDT90 (arm E). Further, when a sale was made and the well tested safe, the tester would keep BDT30 and BDT60, respectively, and hand over the difference to the supervisor. The intervention's protocol stipulated that sale offers would be made in villages assigned to group E first, in order to avoid having to increase the price in some villages later in the project, which could cause discontent.

All 128 villages considered in this study were stratified by high or low arsenic contamination, determined by whether the share of unsafe wells in the BAMWSP testing campaign carried out years earlier was below or above the median. Villages belonging to arms A, B and C were also stratified by union (a larger administrative unit).[2] Then, using a pseudo-random number generator, 50 villages were assigned to treatment arms A and B each, 16 villages were assigned to treatment arm C while the remaining 12 villages were randomly assigned to treatment arms D or E. 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 due to a data-entry error and was thus, in principle, unrelated to village characteristics. 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 five 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 124 clusters divided into experimental arms A (49 clusters), B (48), C (15), D and E (6 each): in the rest of the paper we will refer to them as 'villages', although as we have explained this is not always correct.

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 December 2015 and June 2016, surveyors completed a census of all wells in the 124 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 (98.5%) 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.

All households were then offered an arsenic test under one of the various selling schemes. Concomitantly, enumerators administered a short household ('baseline') 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 13,989 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 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

[2]Unions are the third smallest administrative unit, after *mouzas* (groups of 2 to 3 villages) Our 128 villages come from 9 unions. Villages in arms D and E were only stratified by high vs. low As because each group only included six villages.

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.[3] The average time elapsed between baseline and endline surveys was about eight months, and 86.4% of the households had their follow-up interview between seven and nine months after the baseline interview. Both the baseline and the follow-up surveys recorded whether a household used the well for cooking and drinking.

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 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. This is because we did not record the choice of water source for households who did not use the well for drinking, and so for this latter subset we cannot verify switching behavior. All summary statistics except those on the first line of Table 1 are thus calculated for households who used the well for drinking.

The head was male in 85% of households, while 27% 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 3.7 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, more than 70% of wells had never been tested, and 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 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

[3] 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.

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 50m, out of an average of nearly 12 wells within that distance.

The immense public health challenge due to widespread As 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 tube well water?" Similarly, all but a handful of respondents replied yes when asked "[a]re you aware of the health risks of drinking tube well water containing arsenic?"

Almost all wells (98.6%) were privately owned and on average relatively shallow (176 feet, or 54 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 BDT7,480, or about USD320 using the PPP exchange rate from World Bank (2015).[4] This implies that the cost of BDT45 for the test charged in treatments A, B, C, and (in case of a "safe" result) D, represented slightly more than one half of a one percent of the installation cost. This amount is also close to the price of one kilogram of rice in the capital, Dhaka.

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.[5]

Column 9 of Table 1 shows the p-value for the null hypothesis of equality of means across the five treatment arms, while in column 10 we test the null of equality of means only among the three larger arms, for which we have more statistical power. In the former case, the null is rejected in 13 of 26 cases at the 10% level, while it is rejected in five cases in the latter. 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. Still, the figures show that in some cases there were substantively large 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 21% of household heads in our study area had no schooling, this number drops to about 6% in treatment C, while is close to 40% in arm E. Or, the fraction of respondents who did not know the status of their well ranged from 69% 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. Because these characteristics may affect test purchase decisions and switching behavior upon receiving news of having an unsafe well, we will also show results that control for observed covariates, and we will show that in general estimates are robust to such inclusion.

At the bottom of Table 1 we look at attrition. Overall, 8.3% of the households drinking from the

[4]Using the nominal exchange rate, the average price was about USD100. Well depth is a key predictor of installation costs: in our data, the elasticity of cost with respect to depth is about 0.5.

[5]Recall that we did not survey households who did not own a well. Had we also included them, the average number of users and the average number of household members should have been expected to be about the same.

index well could not be matched to the endline data, either because of true attrition (5 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 five arms is rejected at the 5% level for the prevalence of duplicate identifiers, but the error rates are anyway small, and the null is not rejected when we only compare the three larger experimental treatment groups.

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 conceptual framework that should help interpreting the results, especially in light of ex-ante predictions.

Thinking first about demand, and keeping everything else the same, we would expect purchase to be more likely among richer households, and less likely when prices are higher. Recall, however, that in groups A, B and C the sale price was the same for all households, so our data cannot be used to estimate an elasticity of demand with respect to price. In contrast, in arm D the sale price was nominally the same as in A, B and C, but because payment was only required in case of a 'positive' result (that is, if the test showed that As concentrations were 50ppb or below), the expected price in D was in practice always lower, except for households who were absolutely certain that their well was safe. Recall—see Table 1—that only 17% of respondents believed their well to be safe, and in practice there might have been some uncertainty even among individuals who reported such beliefs. In addition, in principle even someone with complete faith in the safety of their well may have valued the test, for instance as a signal to others about land value, or about the health of household members. Still, we know that the large majority of respondents did not know the As concentrations in their drinking water, and so in group D we expected to see lower expected prices and thus higher demand relative to A, B and C. Because the nominal price in group E was twice as large as in D we also expected—trivially—to observe lower uptake in E relative to D, although we had no clear priors about demand in E relative to A, B and C.

Another important determinant of demand is the value of the information received. Attaching value to such information requires, first, awareness about the health risks associated to continued exposure to As. In Table 1, we have already seen that virtually all respondents knew—at least in a general sense—about the presence and health risks of As in tube well water. On the one hand, such almost universal awareness confirms that households were likely to care about the information being offered, but on the other hand it implies that we have barely any variation in the data that we can exploit to predict demand. The enumerators only recorded yes/no answers to simple questions about awareness, but in practice we would expect variation in the perceived risk of As. Individual-level access to information and the ability to process such information—most likely affected by schooling—are primary drivers of heterogeneity in perceived risk. In addition, the health risk posed by As is also a function of factors such as genetic traits and nutrition that may influence As metabolism in ways that will be almost impossible to gauge for most people (Ahsan et al. 2006, Ahsan et al. 2007).

Indirect support for the existence of both variation in perceived health risks and a high degree of awareness comes from data on subjective risk perceptions collected in 2008 in the neighboring Araihazar sub-district.[6] These data, collected for a different population, cannot be used in a regression framework as a proxy for user-specific risk perceptions in our sample, but they are interesting because they come from a very similar and geographically close rural area. Respondents were asked to state the perceived probability that a child or an adult will develop either skin lesions (a frequent marker of arsenicosis) or more generic "serious health problems" as a consequence of drinking water from a well with unsafe levels of arsenic. The survey instrument described "serious health problems" as conditions that can impair normal daily activity such as working, going to school, playing or helping out with household chores. The probabilities were assessed using different lengths of exposure, ranging from one month to 20 years. In Figure 3 we display the empirical distribution of beliefs for different time horizons. The histograms show a large degree of variation in risk perceptions, but overall they also document a fairly sophisticated understanding of the risks, with progressively higher perceived risk as the time horizon lengthens. For instance, almost no one reported a probability of serious health damage above 30 percent for a 1-month exposure (the 'true' probability being most likely close to zero), while more than half of respondents thought that health damage would certainly appear after 20 years of exposure (a grim assessment that may not be too far from the truth).

Also important for demand is trust towards the test result. There is now 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. However, although we cannot completely rule out that a degree of mistrust affected demand, we do not have compelling reasons to think that any such mistrust should have been different across treatment arms A, B and C. Conceivably, in arms D and E some prospective buyers may have been suspicious of tests offered at a price that depended on the results, and this may have dampened demand.

Next, even conditional on the same perceptions about health risks, demand for information will likely depend on the utility of health and on beliefs about the economic cost of exposure to As. Both of these may depend on both demographic composition of the household and on the main type of economic activity. For instance, parents may be more willing to purchase a test if they are worried about the health of their children, or individuals engaged in manual labor may value information more if they think that their productivity could be severely affected by arsenicosis. Also important for purchase

[6] Such data were collected in the context of a RCT that evaluated the impact of providing information on As contamination, free of charge, using scripts that highlighted to a different extent the fact that arsenic risk increases with As concentration, and is not merely 'binary' (that is, 'safe' vs. 'unsafe'), as many mistakenly understand it, see Bennear et al. (2013).

decision are beliefs about the possibility of taking mitigating action in case of As contamination of one's well water. This in turn is a function both of the expected benefits from switching to a safe (or safer) water source, and of the availability of such better alternatives nearby. Earlier work carried out in the neighboring Araihazar sub-district found that distance from safer alternatives was an important predictor of switching behavior, given that fetching water from someone else's well increased significantly the time and effort devoted to the task (Madajewicz et al. 2007).

Most of the factors discussed so far (except price) are related to characteristics of the potential buyers (including beliefs) that were likely to affect demand similarly across treatments, given that in all treatment arms sales were conducted providing the same information content to prospective buyers. There are however some factors that may have generated differences in demand in ways that were, in some cases, hard to predict. First, the sale types were likely to generate differences in the availability of information on safe water sources nearby. While buyers were privately informed about the test result in treatment A (as well as D and E), information was public within groups of buyers in treatment B, and it was public in treatment C, where a metal placard indicating the test result was attached to the well spout. If households anticipated this, thus expecting easier location of safe sources in case of need, demand may have been higher in B and C versus A. On the other hand, the opposite may have happened if households disliked the prospect of sharing their well in case their result indicated low As levels, or they may have been worried about a bad result affecting negatively land value or the perceived healthiness of household members in the community. Such considerations, however, required a good deal of sophistication, and we did not expect them to lead to substantive differences in demand across arms A, B and C.

Another element that could have increased demand in D and E relative to the other groups, even conditional on the same (expected) price, is that earlier field work in the neighboring Araihazar sub-district had indicated that several people disliked the idea of 'paying for bad news', and so the contingent contract offered in D and E allowed buyers to avoid the risk of having to pay such psychological cost. However, given that we do not observe the perceived probability of the well water being safe (that is, we do not know the 'subjective price' faced by each household), we cannot disentangle the relative role of prices and psychological costs of 'paying for bad news', at least not without making strong assumptions on subjective beliefs about water safety.

Next, we discuss factors likely to matter in the decision to switch to a different water source after our interventions. Of course the strongest prediction was that switching would depend on the actual result of the test. It is less clear whether, conditional on being above or below the safety threshold, higher levels of arsenic would have led to more switching. Madajewicz et al. (2007) found strong evidence of a gradient in the relationship between arsenic and switching, but there is also evidence that many individuals mistakenly understand the problem in binary terms—safe vs. unsafe, see Bennear et al. (2013). In our analysis, we mostly focus on the decisions of households who purchased the test and received an 'unsafe' test result: in principle even other households may have reacted, but our prior (confirmed by the results, as we will show) was that responses would have been concentrated among those who cared sufficiently to purchase the test, and who received 'bad news'.

A key consideration is that any difference in switching between any two experimental arms could have emerged, in principle, either from different selection into purchase (because buyers in different arms may have been of different 'type', and hence reacted differently to the same information) or from the way information was provided (in which case even identical buyers may have reacted differently).

We expected most of the considerations relevant for demand to also affect the decision to switch, and in the same direction. So, we expected factors leading to higher willingness to pay for information to also lead to higher willingness to use such information. However, after the realization of the test result, we also expected some factors to become more salient, possibly leading to differences in switching behavior. In group B, we hoped that two key factors would lead to higher switching rates relative to A. First, the soft commitment at the time of purchase could have been strong enough to facilitate well-sharing within the group. Second, group purchases also facilitated the sharing of information within the group, thereby leading to better information about safe alternatives to one's well. Similarly, in group C the presence of the metal placard affixed to the tube well spout should have allowed more awareness in the community about safe options, given that most wells are easily observable by any passerby. In addition, the placard could also have provided a daily reminder about the health costs of drinking from an unsafe well.

An important consideration is that although these factors led us to expect higher switching rates in B and C relative to A, the differences between arms were subtle. Test results were delivered in different ways, but in principle households may have easily exchanged information in all treatment arms. Similarly, in group B the 'commitment' was soft and not enforced, and in group C well owners could have easily removed the placard from the well.

Finally, while our expectation was that differential selection into purchase would not be particularly important in groups A, B and C, the same was not true for group D and E, where payment was conditional on the result. Buyers in D were thus more likely to expect an unsafe result than in A, B and C, where the nominal price of the test was the same, and even more so in group E, where the nominal price was twice as large, and thus more 'optimistic' households were least likely to buy, everything else being the same. This generated unclear predictions about switching behavior, because more pessimistic respondents faced lower expected prices and thus were more likely to purchase, but the vast majority of households were drinking from their well at baseline *despite* their pessimistic beliefs, and so they were perhaps not the ideal households to target in a campaign aiming at increasing switching off unsafe water.

5 Results

In this section, we first estimate the effect of our intervention's test selling schemes on uptake rates. Next, we describe the information on As levels that was revealed by the testing campaign, and finally we discuss to what extent such information changed household behavior in terms of choice of water source for drinking.

5.1 Demand

Of the 12,684 households who used their own well for cooking and drinking at baseline and who were offered an arsenic test, 3,456 (27%) bought a test under one of our selling schemes. We also found that 211 of 1,302 households (16%) who were *not* using their well at baseline purchased the test, but we will not analyze their responses to the result because for them we cannot verify if they switched to a different source.

To estimate the average treatment effect of selling schemes B, C, D, and E, relative to A, we estimate the following equation using a linear probability model:

$$buy_{svh} = \beta^B B_v + \beta^C C_v + \beta^D D_v + \beta^E E_v + \gamma X_{svh} + \delta_s + \epsilon_{svh}, \tag{1}$$

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, C_v, D_v, and E_v are village-specific indicator variables for the respective treatments, X_{svh} is a set of predetermined household and tube well characteristics, and ϵ_{svh} is an error term. To account for our 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—only for treatments A, B, and C—by union. All standard errors and statistical inference are robust to the presence of intra-village correlation of residuals.

In Figure 4 we show graphically the simple comparison of take up rates across arms without the inclusion of controls. A first clear result is that neither the incentives for 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 A, B and C. 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. Third, and again consistent with our priors, the offer of a free test in case of bad news led to much higher purchase rates (58%) in group D relative to A, B and C. Fourth, doubling the price in E relative to D reduced demand by about half (35%). Again as in many other earlier studies we thus find that demand was quite elastic with respect to price $(((0.346 - 0.582)/0.582)/((90 - 45)/45) = -0.41)$, although we can only calculate it for the pay-for-good-news contract. Interestingly, despite the sharp drop in demand associated with the doubling of the price, the purchase rate in E remained about 10 percentage points higher than in A, B and C. In Figure 5, we also show, for each experimental arm, histograms of village-level purchase rates. The three histograms on the top row show that the similar purchase rates in groups A, B and C are also reflected in similarly-shaped histograms, with some purchase almost everywhere, most purchase rates in the 10 to 30% range, and few outliers with very high demand. In addition, looking now at the two histograms in the bottom row of Figure 5, we also see that the higher purchase rates observed in D and E was not driven by a few outliers but could be observed almost everywhere, with an uptake above 25% in all but one of the 12 villages included in these two arms.

We show the regression results in Table 2, where recall that consistent with equation (1) 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 also not statistically significant. In contrast, and despite the small number of clusters in arm D, the increase in demand with a pay-for-good-news contract is so large ($\hat{\beta}^D = 0.336$) that the null of equality can be rejected at the 1% level, and the 95% C.I. is 0.173-0.499, so that even the lower bound of the interval represents a substantive difference relative to A. When the conditional price was doubled (group E), the null for $\hat{\beta}^E$ $(= 0.1)$ can be rejected only at the 10% level, and the 95% C.I. is -0.014-0.214. 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, except $\hat{\beta}^E$, become substantially more precise.

In group B, where our survey team encouraged the (optional) formation of risk-sharing groups, our data indicate that such option was indeed chosen by many households: out of a total of 1,274 households purchasing the test, 545 (43%) chose this option. In addition, there are two reasons why this figure is actually a lower bound for the actual number that agreed to well-sharing at the time of purchase. First, the choice to sign a group agreement was mistakenly left unspecified for 132 wells (10%). Second, the surveyors indicated that several households agreed verbally to the informal risk-sharing arrangement but refused to sign a document: in such cases, our data thus indicate that the agreement was not *signed*, although it did exist in oral form.

In column 3 we show that the results are also very 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 among experimental arms. We thus include controls for the demographic structure of the household, the gender, education and occupation of the head, age, depth and cost of the tube well, an indicator for the quality of housing, the number of wells and of wells visibly labeled as safe within 100 meters, the fraction of unsafe wells in the village at the time of the BAMWSP blanket campaign, indicators for whether the well itself was visibly labeled as safe, and finally indicators for whether the respondent believed the well water to be safe, or believed it to be unsafe. The only substantive change is that the estimate for β^E become sufficiently precise to become significant at the 1% level. Because missing values in one or more of the controls lead to the loss of about 20% of observations, in column 4 we show that the results remain similar if we re-estimate the model without controls but including only the observations used in column 3. In this case, the main differences are that $\hat{\beta}^C$ and $\hat{\beta}^E$ become larger (0.06 and 0.16, respectively), so that even the former become significant although only at the 10% level.

The controls in Table 3 are obviously not exogenous, and so the corresponding coefficients cannot be interpreted as causal. Despite this, it is informative that the estimates are broadly consistent with the theoretical framework described earlier. Having completed only primary schooling predicts a 7 percentage points decline in demand, while the coefficient for having no schooling is almost twice as large in magnitude. This is consistent with education leading to more awareness about the arsenic problem, although it could also proxy for higher income or wealth and thus higher ability to pay. Other indicators of higher socio-economic status also predict more demand. Coefficients are positive and significant for both better quality roofing and more expensive wells. Larger households are more likely to purchase the test, but the increase in demand predicted by one more child (0.05) is only slightly larger than that observed with one more adult (0.04). Neither the number of wells nearby, nor how many of such wells were visibly labeled safe predict demand. This perhaps suggests that expectations about viable alternatives was not yet sufficiently salient at the time of the purchase, which may have contributed the explain the very similar purchase rates in groups A, B, and C. Finally, well owners thinking that their well is safe stand to gain little from buying a test, and indeed they are 13 percentage points less likely to purchase the test. Interestingly, the belief that the water is *unsafe* barely decreases the probability of purchase, and the corresponding coefficient is not significant at standard levels. These results also suggests that even respondents who stated that they knew about well status did not feel completely confident about their priors, given that a fraction of them decided to purchase the test.[7]

Because beliefs about water safety affect the expected price in groups D and E, we also re-estimate the model with interactions between beliefs and treatment indicators, and we show the results in column 5. We first note that none of the interactions with treatments B and C are significant, consistent with the hypothesis that beliefs did not substantively alter selection into purchase in these arms. However, the results are quite different in treatments D and E. In arm D, households that at baseline were drinking well water thought to be safe—that is, households who expected to pay for the test—were 42 percentage points less likely to purchase the test relative to households in the same villages who had no priors about safety ($-0.134 - 0.288$). In contrast, such beliefs only decreased demand by 13 percentage points in A, and the difference is significant at the 1% level. The estimates also show that among users of wells believed to be safe, demand was 7 percentage points higher in D relative to A ($0.354 - 0.288$). This is consistent with a degree of perceived uncertainty about safety even among households who stated that their well was safe (so that their expected price was lower than in A).

Note also that in arm A the belief of having *unsafe* water, by decreasing the value of information, decreased demand by 7 percentage points (significant at the 1% level) relative to households who did not know the safety status, with similarly negative associations in arms B and C. In contrast, in arm D the same beliefs, by reducing the expected price, *increased* demand by 12 percentage points ($0.185 - 0.066$) relative to households in the same villages who said they did not know the safety of the water. Among households reporting an unsafe well, demand was 54% higher in D relative to A ($0.354 + 0.185$), a much larger increase in demand relative to what observed for respondents thinking that their well was safe (7%), again consistent with the former households facing substantially lower expected prices relative to the latter.

When we look at arm E, where the price conditional on 'good news' was twice as high as in D, the interaction between believing that the well is unsafe and the treatment dummy is almost identical than in D, although the estimate is very noisy and not significant. The interaction with the belief of having *safe* water, while still negative in sign, is close to zero and not significant in group E. This is

[7] When we pool together all treatment arms we find that the purchase rate was 15.2% among respondents who believed their well to be safe, and 14.7% among those who believed that it was unsafe.

surprising, given that such belief implied a higher probability of payment, which in this experimental arm was twice as large as in any other.

5.2 Test Results

Although the purchase rate was far from 100%, our intervention thus generated a large increase in the number of tested wells in Sonargaon. Before looking at the responses to the tests, it is useful to analyze their results. These are summarized in Table 3, where we also include the detailed summary statistics about switching behavior that we will describe later. In addition, in Figure 6 we plot the spatial distribution of safe (diamonds), unsafe (triangles), and untested wells (black dots) across the whole study area in Sonargaon. Although only wells of buyers were tested, the map shows that there was considerable heterogeneity in water quality, even within small areas. Visual inspection of this general pattern becomes even more apparent by looking at each village at a time. As an illustration, in Figure 6 we also provide a magnified view of the spatial distribution of wells in a randomly chosen village (belonging to treatment A). In this specific example (in no way unique) it is easy to see that, first, there was a very dense network of wells and, second, all unsafe wells were not far from other wells with As levels below the Bangladesh safety standard.

Overall, 27% (808/3017) of the tested wells who had been used for drinking at baseline had 'unsafe' As levels based on the Government of Bangladesh standards. Notably, the fraction was thus 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 As 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.[8] Our data are broadly consistent with both hypothesis. First, we find that more recent wells, dug no more than 10 years earlier, were 22% deeper and 9.5 percentage points less likely to be unsafe relative to older ones.[9] 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, 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 *decreases* the predicted probability of the well being unsafe by 15 pp while the belief of the well being unsafe *increases* it by 41 pp (both coefficients are significant at any standard level). The results are also similar if we re-estimate this simple model using only wells from arms A, B and C, where the 'expected price' did not depend on beliefs.

[8]The link between depth and cost is apparent in our data, where the elasticity of cost with respect to depth is 0.51 (s.e. 0.026).

[9]The average depth for older wells was 151 feet, while more recent ones were on average 33 feet deeper. The fraction of unsafe wells was 33.5 and 24% among the older and more recent wells, respectively. The test of equality is rejected at any standard level for both depth and safety.

There was also variation in the test results across different treatments, see also Figure 7 where we show the whole arm-specific distributions of As. 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 As 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 As 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, and made possible by the relatively small number of units (15) in this treatment arm.

Also consistent with our prior, in group D, where testing was conditional on a payment to be handed over only in case of 'bad news', the prevalence of unsafe wells was enormously larger, at 65%. The prevalence was also large, but not *as* large in E (42%), where the nominal price to be paid in case of the well being safe was twice as large. On the one hand, this latter finding was unexpected, given that the higher (conditional) price led us to expect that the largest fraction of unsafe wells would be found in E. On the other hand, recall that both groups D and E only included six villages, so the prevalence of unsafe wells are imprecisely estimated, to the extent that the 95% confidence interval of the difference in the prevalence between E and D also includes positive values ($[-0.57, 0.11]$).

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, on the main source of water for cooking and drinking. Overall, we estimate that at this time, of the 808 wells found to be unsafe, 36% had wells identified as safe within 25 meters, 65% had at least one within 50 meters, and 83% had at least one within 100 meters. This confirms that switching to a nearby safe well was indeed a feasible strategy to mitigate As risk for the large majority of households.[10]

In Figure 8, we show the switching rates observed in each experimental arm, for owners of unsafe wells (the most critical target of the campaign) but also for owners of safe or untested wells. We did not expect to observe much change among the two latter categories, and indeed our data strongly confirm our prior: barely anyone moved from a well tested as safe (13/2,210), while less than 3% of untested wells (227/8,618) stopped being used for drinking. In contrast, the results related to tested wells for which the result showed As$>$ 50ppb (in the top graph) show very interesting patterns. First, 30% of households (60/200) switched from unsafe wells in arm A. Although far from negligible, such figure was at the lower end of the range of switching rates observed in earlier studies, some of which

[10]Note also that these figures underestimate the potential role of switching to reduce As 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.

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, in a number of such early studies tests were conducted in the context of intensive research efforts that may have contributed to a stronger response.

Looking now at the switching rates observed in arms B (sales with 'commitment') and C (with metal placards posted on the well spout), we found that rates in both were much higher relative to A. In B, 56% of households switched, while in C the rate was even higher, at 65%. In column 1 of Table 4 we show the corresponding regression results, confirming that the differences are not only very large but also statistically significant at the 5% level or below. The estimates remain almost identical when we include strata fixed effects (column 2) and they become smaller but remain large and significant when we include baseline controls (columns 3 and 5), suggesting (as for demand) that the impacts are not substantively biased by differences in the level of observed confounders. In all models except in column 4 (more on this below) the presence of placards was associated with more switching relative to the group signing, although the null of equality is never rejected at standard levels.

These results confirm our prediction that group signing or metal placards would lead to more switching, although we cannot separate how much of this was due to an increase in the information about alternatives, or to the (soft) commitment, in arm B, or to the added salience of the placards, in arm C. In B, we do find that having signed the agreement increases the predicted probability of switching by 20 percentage points (relative to 45% among those who did not sign). However, the decision to sign was endogenous and thus this finding cannot be interpreted as causal. In addition, if we also include controls for all confounders included in column 3, the increase in the predicted probability becomes much smaller (= 0.01) and is no longer significant. Recall also that several households did agree to well-sharing despite refusing to do so in writing, which further limits the likely predicted power of the dummy equal to one for households that signed. In arm C, 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. In particular, surveyors found that 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 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, an important consideration given that households were free to remove the placards, which would have considerably reduced the difference between the testing campaigns in arms A and C. It is also interesting to note that this result stands in contrast with Barnwal et al. (2017), who found that placards indicating unsafe As 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 our study.

In arm D, switching rates were also 20-30 percentage points higher than in A, although the lack of precision means that we can only reject the null of equality between these two arms at the 10% level, when controls are included. In contrast, switching rates in E were lower than in all other experimental arms. As we saw in Figure 8, in E switching rates were 26 percentage points lower than in A (and the difference is significant at the 1% level), but in column 2 of Table 4 we see that the difference becomes close to zero and no longer significant when we include strata fixed effects, suggesting that the small sample of this experimental arm leads to problematic inference for this outcome due to limited variation: in Table 3 we can see that in group E only 67 wells were tested, and only three households changed the main source of drinking water.[11]

The results in column 3 are also interesting because they allow us to look at predictors of switching, although once again these 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 As testing offered at no cost, see Chen et al. (2007), Madajewicz et al. (2007), Pfaff et al. (2017). This suggests that in our sampleang more years of schooling was not associated with an increase in As risk-avoiding behavior conditional on information, although recall that demand for tests was lower among households with a less educated head. A notable—and perhaps disheartening—result is that prior beliefs about the well being safe reduced predicted switching by 16 percentage points (the coefficient is significant at the 10% level). Beliefs about water safety may have thus been rather persistent for some households, despite the evidence offered by the tests (recall that we are now looking at wells that were identified as being unsafe by the field test). Recall also that in our study the minimum As 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 As did *not* predict more switching. To the contrary, and using 100 as the omitted category, dummies for the As level being equal to 200, 300 or 500/1000ppb are *negative* and in some case significant albeit only at the 10% level. This finding is consistent with most households gauging safety primarily in a binary way, an unfortunate possibility given that in reality As health risk is to first order proportional to As concentration.[12] In Figure 9 we show indeed that, with the exception of arm B, switching rates as a function of the test result were well approximated by a step function jumping from about 0 for As levels up to 50ppb to a larger and rather constant level for 'unsafe' As level of 100 or above.

In column 4 we also include as regressor a dummy for the presence of a safe well within 50 meters,

[11] Note also that such low switching rates in E cannot be explained by geographical clustering of unsafe wells that may have been more marked in these six villages because of chance. In fact, our data show that among the owners of unsafe wells in E, 30% had one safe well or more within 25 meters, while 60% had at least one within 50 meters. These figures are lower than in D, but are close to those in B and C, where switching rates were high.

[12] 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 lead to more switching, with some evidence that it actually *decreased* it.

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 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 (21 percentage points) and significant at the 1% level. Interestingly, in group B this association almost completely disappears, given that the interaction is *negative* and its magnitude is almost identical to that observed in arm A. 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 negative but small and not significant at standard levels, so in this group, as well as in A, distance appears to be an important predictor of switching. In contrast, in both D and—especially—E the interactions were large and negative, so that in D availability of safe wells barely predicts switching (and the null of no association can be rejected at standard levels), while it actually *reduces* predicted switching in E (and the null of no association is rejected at the 5% level). But recall that so few households actually switched in group E, that it is impossible to draw solid inferences as to what predicted switching in this experimental arm.

Finally, in column 5 we show that if we re-estimate the same basic model as in column 2 (with strata fixed effects), but including only the observations with non-missing covariates used in column 4, the differences in switching rates between arms, conditional on test purchase, remain substantially unchanged.

A limitation of our data is that we cannot gauge to what extent switching was associated to a reduction in arsenic risk. Surveyors were asked to record the GPS location of the new source of drinking water, but the GPS system 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. The lack of biomarkers measuring As exposure also limits our ability to evaluate the health impacts of our intervention, although earlier work in neighboring Araihazar sub-district found strong evidence of such benefits following testing campaigns, see Chen et al. (2007). Despite these limitations, some useful information can be gleaned from reports from households that changed the source of drinking water. Of the 405 'switchers', almost all (401) listed safety concerns as the primary reason for their decision. However, about half of these (201/401) had switched to a different well who was itself perceived as being unsafe, while 152 had switched to a well reported as being safe, and the remaining 48 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 As, but this finding suggests that in our study area a degree of As 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.

The results in columns 1-5 of Table 4 only use information from wells identified as unsafe, but in column 6 we also compare switching rates for all wells which had been used for drinking at baseline. In other words, these are the *unconditional* switching rates, regardless of purchase decision or test result. These results show that in group A only 3.7% of households switched well, while in B the fraction was 1 percentage point higher (not significant at standard levels) and in C it was 4 percentage points higher (significant at the 5% level). The highest switching rates were found in arm D, where almost 20% of households switched following the test sales, while the lowest was found in arm E where barely anyone switched. These results are of course not surprising and are a reflection of the findings we described earlier in terms of demand, prevalence of unsafe wells, and switching behavior conditional on safety.

Given that switching from safe wells is very rare, it would have been interesting to compare these patterns with the fraction of unsafe wells in each arm, which could be seen as the fraction of households needing mitigation of As risk. Unfortunately, given that we know the As status only for tested wells, such fractions are unobserved. On the other hand, 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. That this was the case is also suggested by the similarity across arms in terms of depth and cost of wells (see Table 1), which as we explained are strongly (negatively) correlated with As contamination. Recall also that in arms A, B and C (where the price paid did not depend on safety status) the fraction of wells that turned out to be unsafe upon testing was similar, ranging from 16% in B to 27% in C (see Table 3). 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, these back-of-the-envelope estimates suggest that, despite the many tests sold, switching rates remained well below the likely fraction of unsafe wells, perhaps with the exception of what we observed in arm D, where the perceived price was lower and purchase particularly appealing among users of wells suspected to be unsafe.

6 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 may severely limit the effectiveness of information 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 tube wells 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. 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. Unfortunately, the public sector is no longer providing free testing of tube well water, and a private market for tests still does not exist. It is thus imperative to both learn more about ways to increase demand for testing, and to identify ways to induce households informed of the unsafe nature of their water to switch to different sources.

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, conditional on learning that one's well is unsafe, their effect on well switching. Our results show that relatively subtle differences in the way information was provided, while leading to small differences in uptake, 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.

Our results are consistent with a conceptual framework where the adoption of health-protecting behavior is increased by pre-commitment to share drinking water (despite the absence of enforcing mechanisms), by the ease of access to information on safe sources, and by 'reminders' on water safety provided by placards affixed to the tube well spouts. We also show that sales where the price is demanded only when the buyer receives 'good news' can be very successful at increasing sales and possibly responses to information, although, due to limited sample size, our results are only exploratory, and we cannot disentangle the separate impact of a lower (expected) price from the possible benefits of allowing individuals to avoid the prospect of having to 'pay for bad news'.

Our results should be very useful for the design of information campaigns that aim at providing measures of risk exposure that varies 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.

On the one hand, and despite the positive fees, our team of surveyors managed to sell more than 3,000 tests for a total of about 12,000 wells. This allowed to uncover the presence of about 800 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. On the other hand, to the extent that our results can be extrapolated to the rest of the country, our results also show 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. In addition, and 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.

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. 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).

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.

Ahsan, H., Y. Chen, M. G. Kibriya, V. Slavkovich, F. Parvez, F. Jasmine, M. V. Gamble, and J. H. Graziano (2007). Arsenic metabolism, genetic susceptibility, and risk of premalignant skin lesions in Bangladesh. *Cancer Epidemiology and Prevention Biomarkers 16*(6), 1270–1278.

Ahsan, H., Y. Chen, F. Parvez, L. Zablotska, M. Argos, I. Hussain, H. Momotaj, D. Levy, Z. Cheng, V. Slavkovich, A. van Geen, G. R. Howe, and J. H. Graziano (2006). Arsenic exposure from drinking water and risk of premalignant skin lesions in Bangladesh: Baseline results from the health effects of arsenic longitudinal study. *American Journal of Epidemiology 163*(12), 1138–1148.

Alsan, M. and M. Wanamaker (2017). Tuskegee and the health of black men. *Quarterly Journal of Economics Forthcoming.*

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.

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), 407414.

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. (2014a). Getting essential health products to their end users: Subsidize, but how much? *Science 345*(6202), 1279–1281.

Dupas, P. (2014b). 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.

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.

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.

Janssens, W., J. Gaag, T. F. Rinke de Wit, and Z. Tanović (2014). Refusal bias in the estimation of HIV prevalence. *Demography 51*(3), 1131–1157.

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.

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.

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.

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*, 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.

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.

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 tube well spouts in arm C. The pictures show placards attached to, from left to right, safe (blue, $As \leq 10ppb$), marginally safe (green, $10 < As \leq 50ppb$), and unsafe (red, $As > 50ppb$) wells, respectively.

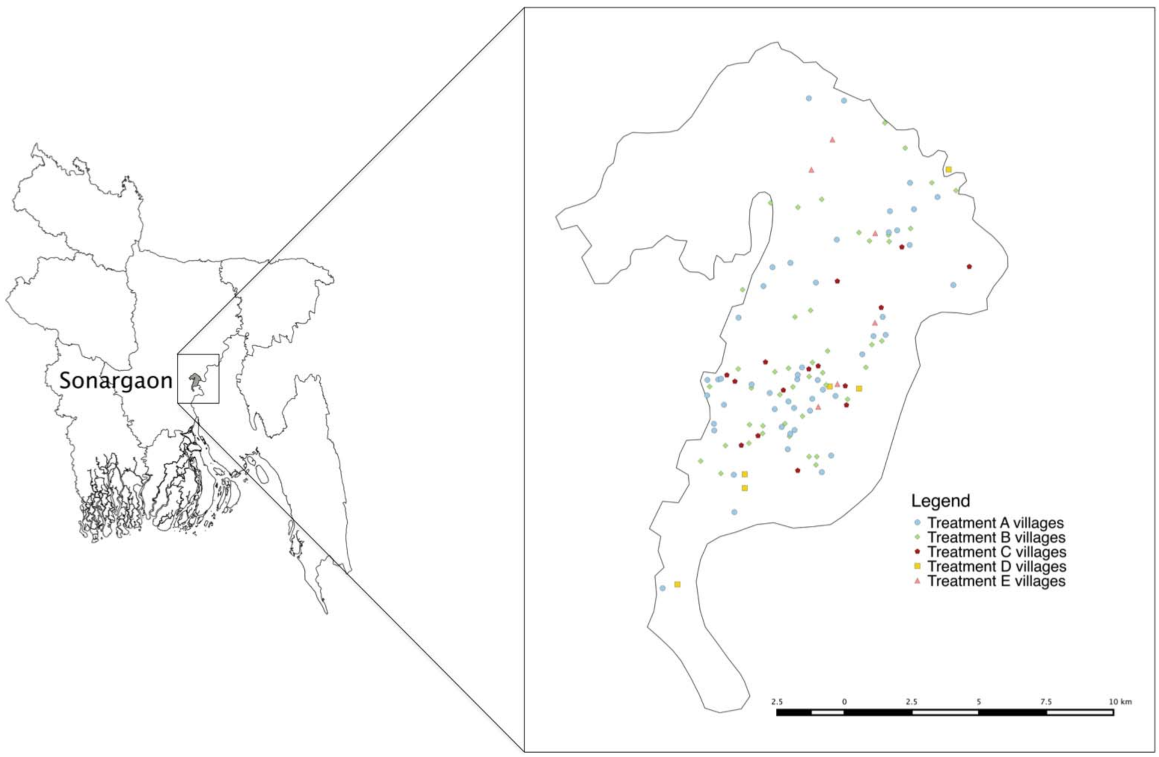

Figure 2: Study area and treatment assignment

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.

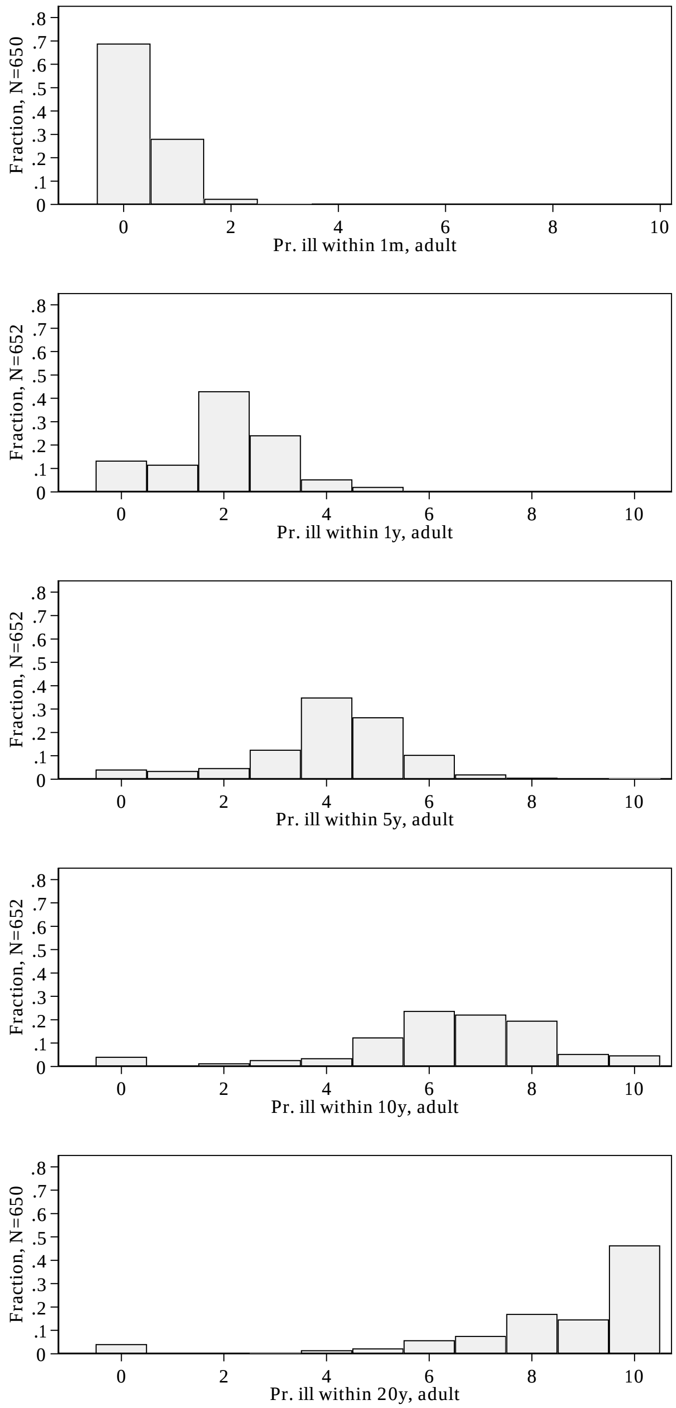

Figure 3: Subjective perceptions about Arsenic health risk

Source: Authors' estimations with 2008 data from Araihazar sub-district (bordering Sonargaon to the North). Each graph shows histograms of subjective probabilities, elicited from respondents, of an adult developing 'serious health conditions' as a consequence of drinking from an As-contaminated well, within a time horizons ranging from one month (top) to 20 years (bottom). Beliefs were measured in discrete steps on a scale from 0 (event perceived as impossible) to 10 (perceived as certain).

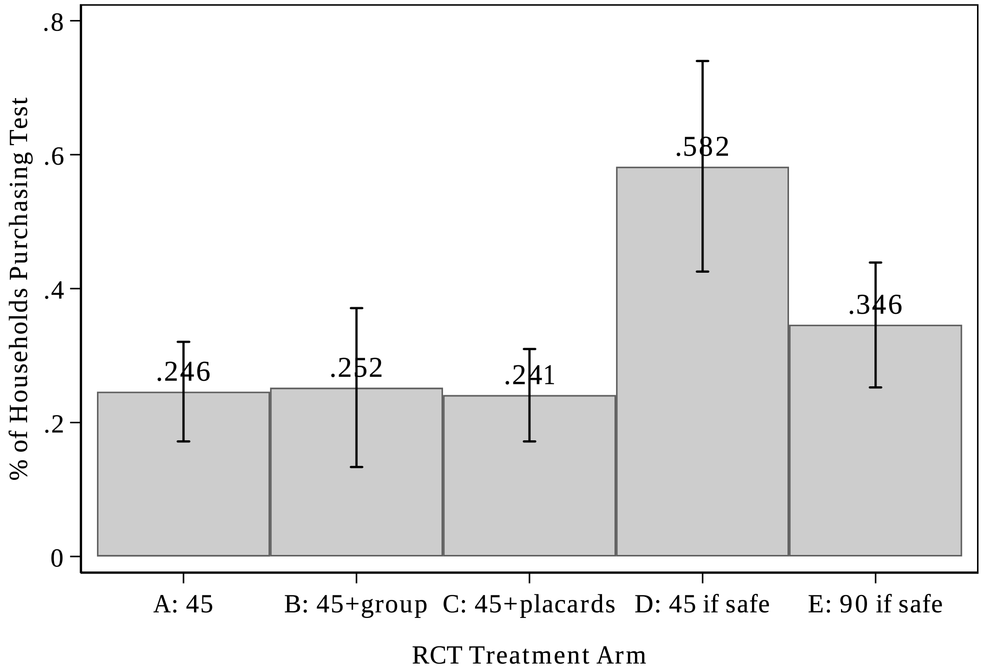

Figure 4: Demand for Tests of Arsenic Concentration in Well Water

Source: Authors' estimations from baseline data (December 2015 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 used for the five bars are, from left to right, $n = 5,164$ (A), 4,697 (B), 1,549 (C), 788 (D), and 486 (E).

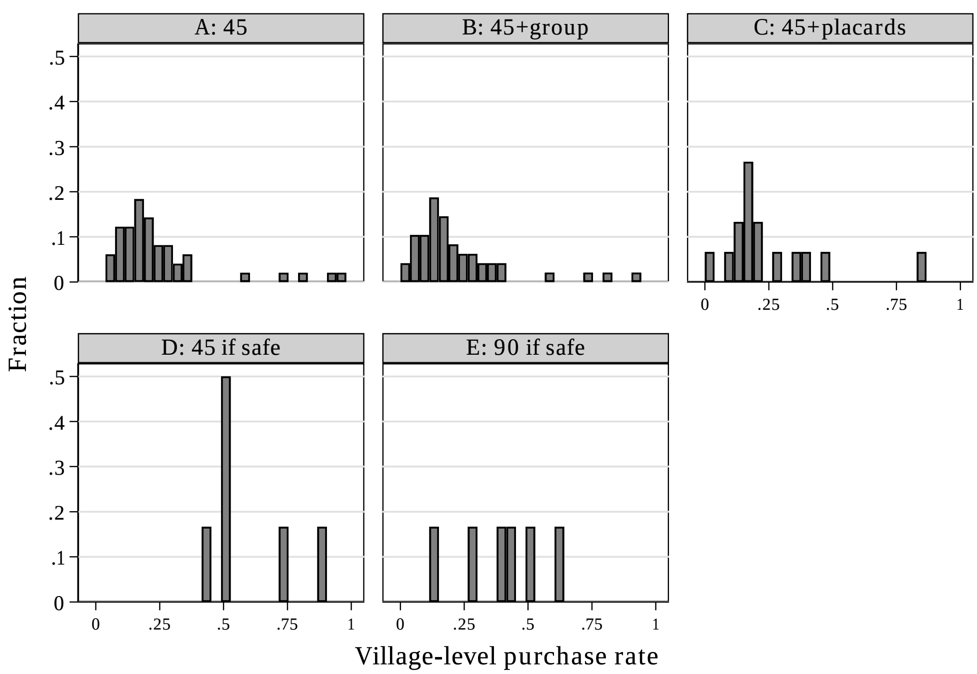

Figure 5: Distribution of purchase rates across villages by experimental arm

Source: Authors' estimations from baseline data (December 2015 to June 2016). Each figure shows an arm-specific histogram of village-level purchase rates of tests. The number of villages in each arm was 49 (A), 48 (B), 15 (C), 6 (D), and 6 (E).

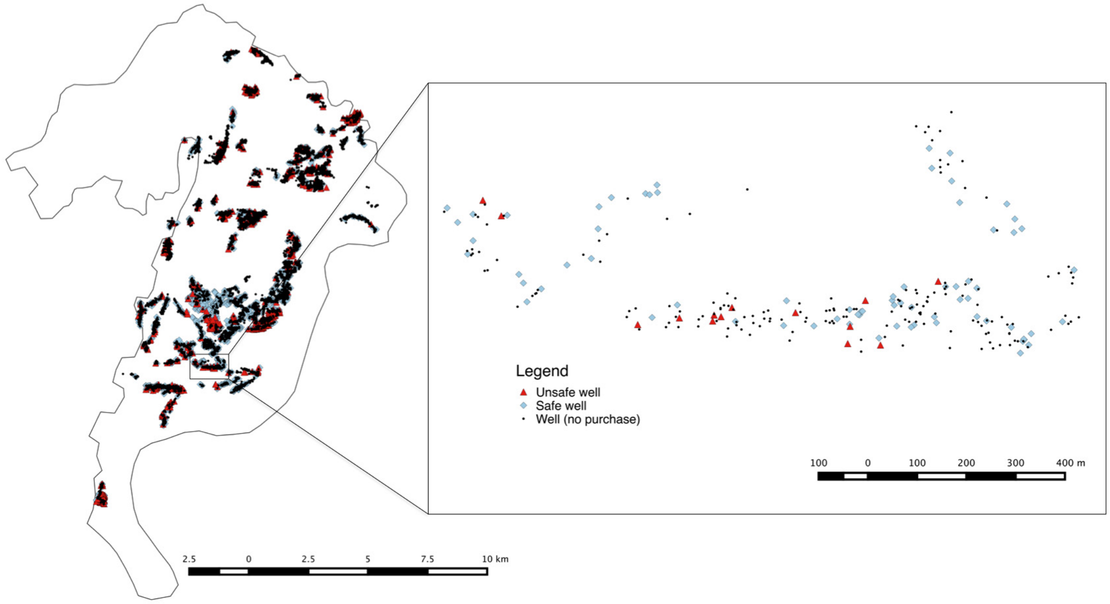

Figure 6: Test uptake and results

Notes: Author's illustrations from baseline data (December 2015 to June 2016). The figure in the box illustrates test uptake and results in one randomly selected village who was part of experimental arm A.

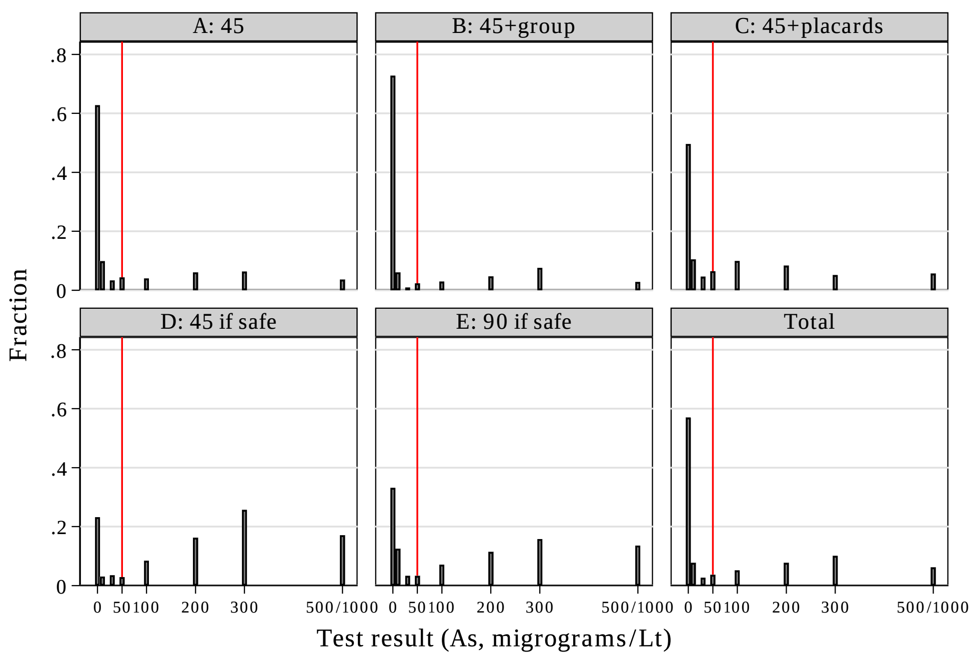

Figure 7: Distribution of Arsenic by experimental arm

Source: Authors' estimations from the result of the tests purchased at baseline (December 2015 to June 2016). Each figure shows an arm-specific histogram of arsenic (in ppb, or micrograms per litre). The field tests identified the As level as a value in the set $\{0, 10, 25, 50, 100, 200, 300, 500, 1000\}$. Results of As=1000 were rare and hence we pool 500 and 1000 together.

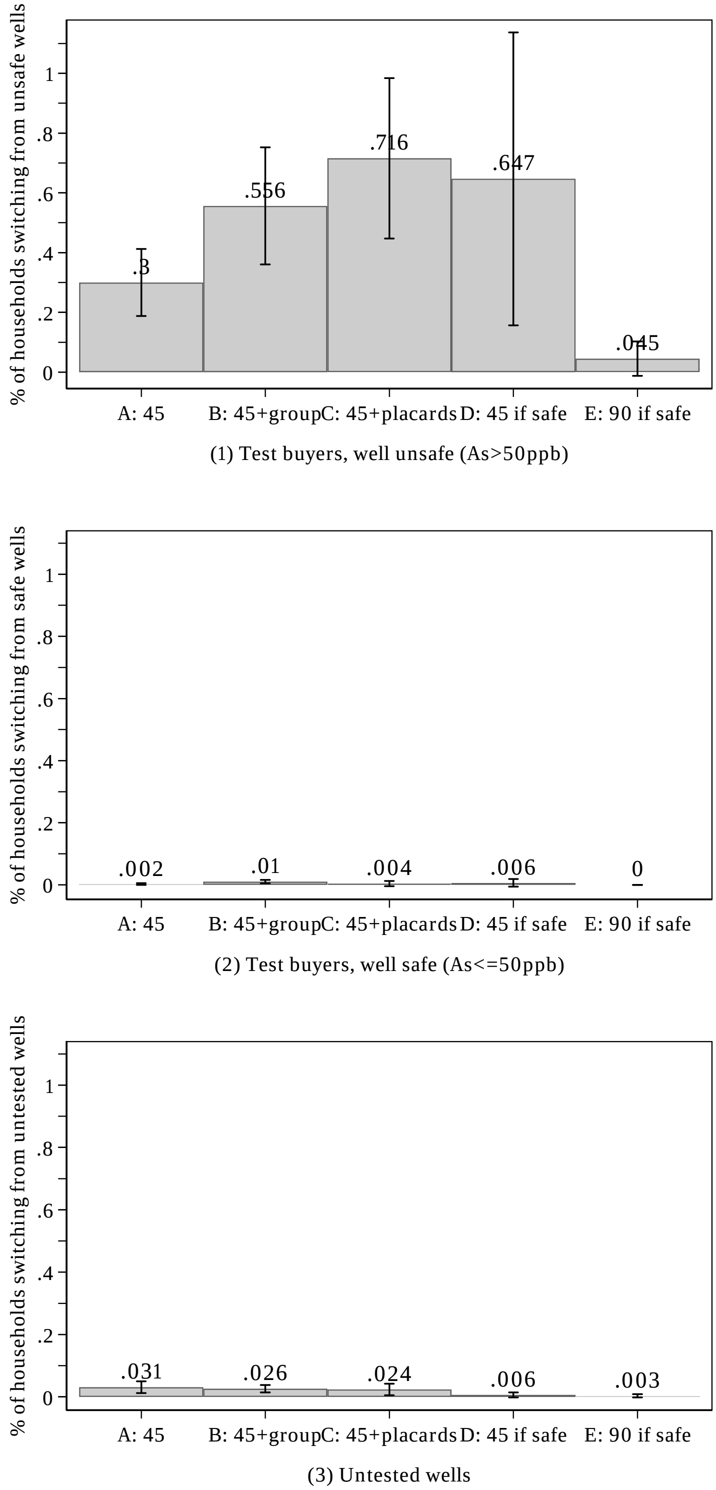

Figure 8: Switching rates by arm and by test result

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 wells that tested unsafe (graph 1 on top), safe (2, middle) and for wells that were not tested because the test had not been purchased (3, bottom). The vertical intervals within each bar are 95% confidence intervals robust to intra-village correlation. The number of wells n_T, $T \in \{A, B, C, D, E\}$ used in each bar are as follows: unsafe wells (graph 1 on top), $n_A = 200$, $n_B = 160$, $n_C = 95$, $n_D = 286$, $n_E = 67$; safe wells (graph 2), $n_A = 840$, $n_B = 869$, $n_C = 253$, $n_D = 155$, $n_E = 92$; untested wells (graph 3), $n_A = 3639$, $n_B = 3252$, $n_C = 1104$, $n_D = 318$, $n_E = 305$.

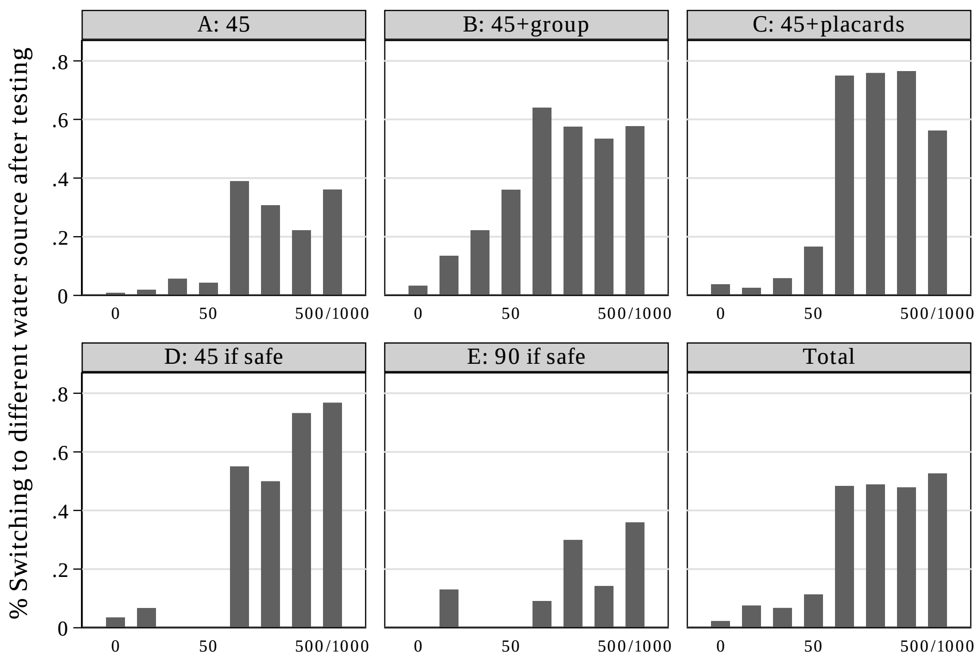

Figure 9: Switching rates from tested wells, by As level and experimental arm

Source: Authors' estimations from baseline (December 2015 to June 2016) and endline (August 2016 to January 2017) data. Each figure shows an arm-specific histogram of arsenic (in ppb, or micrograms per litre). The field tests identified the As level as a value in the set $As \in \{0, 10, 25, 50, 100, 200, 300, 500, 1000\}$. Results of As=1000 were rare and hence we pooled 500 and 1000 together. 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)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Overall			Means by experimental arm					p-values	
	Obs.	Mean	St.Dev.	A	B	C	D	E	A=B=C=D=E	A=B=C
Drink from well at baseline	13986	0.907	0.291	0.930	0.884	0.891	0.931	0.905	0.478	0.251
Household head is male	12684	0.846	0.361	0.843	0.850	0.860	0.812	0.846	0.916	0.921
Household head wage worker	12050	0.274	0.446	0.247	0.298	0.378	0.154	0.176	0.048**	0.426
Household head self-employed	12050	0.424	0.494	0.440	0.438	0.285	0.568	0.327	0.149	0.339
Household head no schooling	12050	0.212	0.409	0.262	0.163	0.054	0.370	0.409	<0.001***	<0.001***
Household head primary school	12050	0.323	0.468	0.317	0.343	0.319	0.304	0.224	0.320	0.790
Heard about As in well water	12684	0.996	0.064	0.996	0.996	0.997	0.996	0.99	0.652	0.742
Aware of health risks of As	12684	0.998	0.044	0.997	0.998	0.999	0.999	0.996	0.342	0.294
House has concrete roof	12505	0.171	0.377	0.175	0.184	0.133	0.122	0.209	0.391	0.405
Household members	12297	3.66	1.34	3.63	3.57	3.59	3.99	4.48	<0.001***	0.826
Number of Children	12297	1.5	1.06	1.48	1.43	1.47	1.72	2.04	<0.001***	0.691
Well As status unknown (belief)	11718	0.757	0.429	0.684	0.789	0.903	0.836	0.600	<0.001***	0.0005***
Well As status unsafe (belief)	11718	0.074	0.261	0.097	0.042	0.057	0.054	0.217	0.162	0.258
Well As status safe (belief)	11718	0.170	0.375	0.220	0.168	0.041	0.110	0.183	<0.001***	<0.001***
Well labeled safe	11718	0.003	0.051	0.003	0.001	0.001	0.003	0.025	0.045**	0.101
Wells within 50m	11528	12.100	14.900	10.400	14.500	12.100	11.700	8.410	0.055*	0.252
Wells within 50m labeled safe	11528	0.022	0.161	0.028	0.003	0.007	0.022	0.173	0.014***	0.030***
Share unsafe wells (BAMWSP)	12684	0.747	0.128	0.759	0.720	0.782	0.764	0.732	0.652	0.412
Well is privately owned	12684	0.986	0.118	0.980	0.991	0.992	0.990	0.973	0.196	0.168
Well depth (×100 feet)	12684	1.760	1.080	1.830	1.770	1.730	1.420	1.600	0.622	0.880
Well age (years)	12684	9.120	7.580	8.860	9.490	9.000	9.130	8.830	0.051*	0.010**
Well cost (×10000 BDT)	12684	0.748	0.633	0.763	0.764	0.707	0.634	0.744	0.468	0.734
Persons drinking from well	12616	8.76	10.9	8.88	8.49	9.75	8.68	6.88	0.002**	0.571
Attrition	12684	0.083	0.275	0.094	0.089	0.063	0.037	0.045	0.085*	0.314
Lost after baseline	12684	0.053	0.224	0.058	0.055	0.045	0.034	0.037	0.568	0.599
Duplicate I.D. at baseline	12684	0.030	0.169	0.036	0.033	0.018	0.003	0.008	0.014**	0.416

Notes: Author's calculations from baseline data (December 2015 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, $n = 5,551$ wells), 48 (B, $n = 5,316$), 15 (C, $n = 1,739$), 6 (D, $n = 846$), and 6 (E, $n = 537$). Except for the first variable ("Drinks from well at baseline") all variables are summarized for household who used the specific well for cooking an drinking at baseline. Differences in the number of observations across these variables are explained by missing entries during the data collection. The exception is "Persons drinking from well", which was recorded only when the well water was being used for drinking by the respondent's household. The p-values in columns 9 and 10 are for tests of equality of means between treatment arms (robust to intra-village correlation), for the null of equality across all five arms or for the null of equality of arms A, B and C, respectively. Asterisks denote test significance: *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Demand for tests

	(1)	(2)	(3)	(4)	(5)
Arm B: 45+group	0.006	-0.002	-0.009	0.000	-0.022
	(0.071)	(0.037)	(0.018)	(0.018)	(0.021)
Arm C: 45+placards	-0.005	0.038	0.023	0.061*	0.018
	(0.051)	(0.032)	(0.030)	(0.035)	(0.034)
Arm D: 45 if safe	0.336***	0.337***	0.331***	0.331***	0.353***
	(0.083)	(0.057)	(0.060)	(0.061)	(0.063)
Arm E: 90 if safe	0.100*	0.110*	0.116***	0.158***	0.085*
	(0.058)	(0.060)	(0.043)	(0.040)	(0.049)
Household head is male			-0.068***		-0.071***
			(0.018)		(0.017)
Household head works for wage			-0.001		0.001
			(0.016)		(0.016)
Household head self-employed			0.029**		0.033**
			(0.014)		(0.014)
Household head has no schooling			-0.134***		-0.135***
			(0.014)		(0.014)
Household head has primary only			-0.072***		-0.072***
			(0.011)		(0.011)
Concrete house roof			0.061***		0.059***
			(0.014)		(0.014)
No. household members besides children			0.043***		0.044***
			(0.008)		(0.008)
No. of children of head in household			0.052***		0.052***
			(0.006)		(0.006)
No. of wells within 100m			-0.000		-0.000
			(0.000)		(0.000)
No. of visibly safe wells within 100m			0.043		0.050**
			(0.026)		(0.025)
% of unsafe wells in village (BAMWSP)			0.116		0.137
			(0.090)		(0.091)
Well depth ('00 feet)			0.028***		0.026***
			(0.010)		(0.009)
Well age (years)			-0.000		-0.000
			(0.001)		(0.001)
Well cost ('0000 BDT)			0.028**		0.029**
			(0.014)		(0.014)
Believes well is safe			-0.132***		-0.134***
			(0.018)		(0.023)
Believes well is unsafe			-0.029		-0.066***
			(0.022)		(0.024)
B× believes well is safe					0.046
					(0.036)
B× believes well is unsafe					0.036
					(0.035)
C× believes well is safe					0.035
					(0.047)
C× believes well is unsafe					0.007
					(0.052)
D× believes well is safe					-0.288***
					(0.069)
D× believes well is unsafe					0.185*
					(0.097)
E× believes well is safe					-0.028
					(0.073)
E× believes well is unsafe					0.172
					(0.139)
Observations	12,684	12,684	9,988	9,988	9,988
R-squared	0.034	0.145	0.151	0.090	0.156
Controls	No	No	Yes	No	Yes
Strata FE	No	Yes	Yes	Yes	Yes
Mean in A	0.246	0.246	0.246	0.246	0.246
Clusters	124	124	113	113	113

Source: Authors' estimations from baseline data (December 2015 to June 2016). The dependent variable is a binary variable = 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 village (estimated by BAMWSP) was below the median. Standard errors are clustered at the village level. 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	
	Total wells	Total	Proportion Unsafe	# Safe Switched	# Safe Did not Switch	# Unsafe Switched	# Unsafe Did not Switch	Total	Switched
A: BDT45	4,679	1040	0.192 (0.031)	2	838	60	140	3,639	113
B: BDT45+Group	4,281	1029	0.155 (0.048)	9	860	89	71	3,252	85
C: BDT45+Placards	1,452	348	0.273 (0.061)	1	252	68	27	1,104	26
D: BDT45 if safe	759	441	0.649 (0.085)	1	154	185	101	318	2
E: BDT90 if safe	464	159	0.421 (0.150)	0	93	3	64	305	1
Total	11,635	3,017	0.268 (0.046)	13	2,197	405	403	8,618	227
Tests of equality (p-value)									
H_0 : A=B=C=D=E			< 0.001						
H_0 : A=B=C			0.3169						
H_0 : D=E			0.1905						

Notes: Authors' calculations using information from a total of 11,635 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) All wells
	Unsafe wells (As> 50ppb)					
B: BDT45+group	0.256**	0.282**	0.202**	0.374***	0.270***	0.013
	(0.113)	(0.111)	(0.097)	(0.106)	(0.098)	(0.014)
C: BDT45+placards	0.416***	0.413***	0.359***	0.333***	0.318***	0.041**
	(0.142)	(0.108)	(0.099)	(0.104)	(0.089)	(0.016)
D: BDT45 if safe	0.347	0.211	0.227*	0.293*	0.214	0.159*
	(0.236)	(0.139)	(0.129)	(0.157)	(0.132)	(0.088)
E: BDT90 if safe	-0.255***	-0.060	-0.095	0.098	-0.121	-0.019
	(0.062)	(0.163)	(0.146)	(0.160)	(0.164)	(0.024)
Household head is male			-0.005	-0.006		
			(0.055)	(0.057)		
Household head works for wage			-0.066	-0.030		
			(0.044)	(0.044)		
Household head self-employed			0.096*	0.072		
			(0.056)	(0.049)		
Household head has no schooling			0.052	0.092		
			(0.054)	(0.058)		
Household head has primary only			0.034	0.041		
			(0.039)	(0.042)		
Concrete house roof			0.018	0.041		
			(0.065)	(0.067)		
No. household members besides children			-0.037**	-0.031*		
			(0.016)	(0.018)		
No. of children of head in household			-0.008	-0.012		
			(0.018)	(0.019)		
Well depth ('00 feet)			0.015	-0.006		
			(0.027)	(0.033)		
Well age (years)			-0.001	0.001		
			(0.002)	(0.002)		
Well cost ('0000 BDT)			-0.061*	-0.055		
			(0.035)	(0.036)		
Believes well is unsafe			0.028	-0.007		
			(0.063)	(0.065)		
Believes well is safe			-0.161*	-0.174*		
			(0.081)	(0.089)		
As = 200ppb			-0.056	-0.031		
			(0.059)	(0.064)		
As = 300ppb			-0.080*	-0.073		
			(0.047)	(0.051)		
As = 500 or 1000ppb			-0.060	-0.036		
			(0.051)	(0.049)		
Number of wells within 50m				0.001		
				(0.003)		
There is at least one safe well within 50m				0.211***		
				(0.075)		
B × at least one safe well within 50m				-0.199*		
				(0.108)		
C × at least one safe well within 50m				-0.043		
				(0.108)		
D × at least one safe well within 50m				-0.160		
				(0.115)		
E × at least one safe well within 50m				-0.352***		
				(0.086)		
Observations	808	808	719	666	666	11,635
R-squared	0.163	0.377	0.392	0.427	0.391	0.076
Controls	No	No	Yes	Yes	No	No
Strata FE	No	Yes	Yes	Yes	Yes	Yes
Mean in A	0.300	0.300	0.300	0.300	0.300	0.0374
Clusters	87	87	81	76	76	124

Notes: Authors' estimations from baseline and endline data. All regressions in columns 1-5 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 As in the water ($As > 50$ppb). In column 3, the decrease in the number of observations is due to missing values in one or more of the controls, and similarly in column 4 a number of observations are lost because the GPS location was not recorded correctly. The model in column 5 is the same as in column 2 but uses only observations used in the results in column 4. The results in column 6 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 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.