Working paper

International Growth Centre

Demand for environmental quality information and households' response

Evidence from well-water arsenic testing

Prabhat Barnwal
Alexander van Geen
Jan von der Goltz
Chander Kumar Singh

January 2017

When citing this paper, please use the title and the following reference number:
E-35128-INC-1

DIRECTED BY LSE UNIVERSITY OF OXFORD

FUNDED BY UKaid from the British people

Demand for environmental quality information and households' response: Evidence from well-water arsenic testing

Prabhat Barnwal [†] Alexander van Geen[‡] Jan von der Goltz[§]
Chander Kumar Singh[¶]

January 26, 2017

Abstract

Access to information about environmental quality may facilitate low cost preventive measures. In this paper, we study demand for information about environmental quality and the behavioral response to the information provided. With a field experiment conducted in Bihar (India), we estimate the price sensitivity of demand for diagnostic testing of drinking water wells for arsenic of natural origin - a serious threat to the health of tens of millions of villagers across South and Southeast Asia. Demand is substantial, but highly sensitive to price; uptake falls from 69% to 22% of households over our price range (Rs. 10 to Rs. 50 – about equivalent to daily per capita income). We further assess how households respond to information regarding the contamination level in their wells. We find that about one-third of households with unsafe wells switch to a safer water source. There is no indication that households who bought the test at higher prices were more likely to respond by switching to a neighboring well. Finally, we demonstrate that households that received adverse test outcomes are more likely to selectively forget test results and proactively remove evidence of their wells' status. Our results highlight the importance of enabling households to take action on information in an effective and socially acceptable way.

JEL Codes: D12; I12; O12; Q50

Keywords: *Environmental quality information, Willingness to pay, Avoidance, Health, Arsenic, Groundwater*

[*]The authors especially thank the Government of Bihar, India for its support, and the International Growth Center for funding. This paper was earlier circulated with the title "Cost-sharing in environmental health products: evidence from arsenic testing of drinking-water wells in Bihar, India". We thank Joseph Graziano, Joseph Herriges, Wojciech Kopczuk, Dilip Mookherjee, John Mutter, Cristian Pop-Eleches, Eric Verhoogen, and seminar participants at Columbia University, IGC South Asia Growth Conference and ACEGD – ISI Delhi 2016. Remaining errors are ours.

[†]Department of Economics, Michigan State University. prabhat@msu.edu.

[‡]Lamont-Doherty Earth Observatory, Columbia University. avangeen@ldeo.columbia.edu.

[§]World Bank. jvondergoltz@worldbankgroup.org.

[¶]TERI University. chanderkumarsingh@gmail.com.

1 Introduction

There is pronounced policy interest in assessing demand for information about environmental quality that is relevant to health outcomes, and in understanding how households react to this information (Pattanayak et al., 2009; Somanathan, 2010; Greenstone and Jack, 2015). Previous research has chiefly focused on the former issue, and asked how subsidies and fees affect access. However, in addition to the question of access, what matters for policy outcomes is how products are used. This is particularly important in case of diagnostic tests which do not offer a tangible product with clear uses and instead, purely provide information on environmental and health quality that can facilitate low cost preventive measures. In this paper, we study the demand for information about environmental quality in the case of well-water contamination with arsenic, and investigate whether the price paid and the information content affects how this information is used.

The health impact of poor environmental quality is particularly important in developing countries. Willingness to pay for information is low and environmental monitoring, weak. At the same time, where those lacking information about environmental quality fail to protect themselves and suffer health consequences, productivity of those affected may be decreased, with potential adverse impacts on economic development if health problems are wide-spread. Hence, similar to preventive health products, such as insecticide-treated bed nets to prevent malaria infection (ITNs), or technologies to remove microbial pathogens from drinking water (Ahuja et al., 2010; Sachs and Malaney, 2002), high social benefits are likely to be associated with provision of information on environmental quality in low income settings. There are two important questions, which we study in this paper. The first relates to the goal of increasing access. To investigate it, we assess how price sensitive is the demand for information on environmental quality. This question is relatively well studied in the context of cost-sharing in the provision of some common preventive health products such as ITNs and water filters (Dupas, 2014a; Kremer and Miguel, 2007; Tarozzi et al., 2014).[1] Yet, given their very distinct nature, it remains important to test whether these findings hold for informational products. For instance, in contrast to a body of evidence establishing the high price sensitivity of de-

[1]Despite the potential of high social benefits, it has proven difficult to chart a path – through private or public provision – to ensure sustainability in access to preventive health products. Given the flaws of both private and public provision, cost-sharing is often suggested as a way to reduce dependency on public programs, without exposing consumers to the full cost of market provision. However, even relatively limited fees have been shown to significantly reduce take-up (Bates et al., 2012; Dupas, 2014a; Kremer and Miguel, 2007).

mand for preventive health care products such as ITNs, Cohen et al. (2015) document a lack of price sensitivity for rapid diagnostic test for malaria. Secondly, it is important to study how households respond to the information about environmental quality revealed by diagnostic products. One, it is essential to assess whether testing has the intended effect: does information provision lead to effective preventive measures? Two, is the effect of information sensitive to price, as screening or sunk cost models would suggest? Three, are there unintended adverse socio-economic implications of environmental quality information revelation, and does revealing environmental quality run counter to social norms, impose stigma, or affect asset values? We assess these questions in the context of households' responses to information on arsenic contamination in their well water.

Arsenic tests for drinking water wells share important product traits with other highly efficient preventive health interventions (Pattanayak et al., 2009). Firstly, in that they offer a potentially effective way of avoiding a significant public health threat. Naturally elevated arsenic concentrations in well water were first reported in the mid-1980s in West Bengal and subsequently shown to extend over a much broader area (Ahmed et al., 2006; Chakraborti et al., 2003; Fendorf et al., 2010). In areas where arsenic contamination is prevalent, tests are essential in that they provide information that is not substitutable. Because the distribution of arsenic incidence in groundwater is difficult to predict, and varies greatly even over small distances, the safety of a well cannot be predicted without a test (van Geen et al., 2002). A well that meets the WHO guidelines for arsenic in drinking water may be found in immediate neighborhood of a very unsafe well. Nor is there an easy way to design wells to be both safe and affordable: within shallow (< 100 m) aquifers tapped by most private wells, there is no systematic and predictable relationship between and arsenic and well depth.[2] At the same time, precisely because arsenic contamination varies greatly over small distances and does not vary substantially over time, well tests make available an effective way to avoid exposure, namely by switching to nearby safe wells. In previous interventions, about one-quarter to two-thirds of households with contaminated wells have been found to switch to safer sources (see, e.g., Ahmed et al. (2006); Chen et al. (2007); Madajewicz et al. (2007)).

Much like other basic preventive health products, arsenic tests are also very cost efficient. The

[2] Arsenic concentrations in well water generally do not vary substantially over time as well, and early concerns that arsenic levels might be rising systematically have not been confirmed (Fendorf et al. 2010). In the context of our study, this means that one time purchase of arsenic testing should be sufficient to reveal the arsenic level in water from a specific well, but it tells little about arsenic level in nearby wells.

cost of goods and services (COGS) for a test provided through our program was a mere USD 2.30, excluding cost purely related to data collection. (There is, of course, a potentially significant inconvenience cost to switching wells.) By stark contrast, the health consequences of chronic arsenic exposure are dramatic. Argos et al. (2010) conducted a large cohort study in an area of Bangladesh where arsenic contamination was representative of the national distribution, and estimated that 21% of all-cause deaths were due to chronic exposure by drinking water at arsenic levels above 10μg/l (the 60th percentile of the arsenic distribution in our sample). Arsenic in tubewell water has also been associated with impaired intellectual and motor function in children (Parvez et al., 2011; Wasserman et al., 2004) and lower mental health in adults (Chowdhury et al., 2015). In consequence, there are significant effects on income and labor supply: Pitt et al. (2015) estimate that lowering the amount of retained arsenic among adult men in Bangladesh to levels encountered in uncontaminated countries would increase earnings by 9%. Matching households to arsenic exposure, Carson et al. (2011) find that overall household labor supply is 8% smaller due to arsenic exposure. Chowdhury et al. (2015) estimate the mental health burden of arsenic contamination for affected individuals alone can be as high as the annual household income in Bangladesh.

Because of their low cost and important health benefits, well tests for arsenic have been provided free of charge at large scale. A number of large-scale testing campaigns have been carried out through public provision in rural communities across the Indo-Gangetic Plain (Ahmed et al., 2006; Fendorf et al., 2010). However, these important programs have not come close to comprehensively covering the geographic area where arsenic is of concern – including in our study area. Due to the continuing installation of new wells and the replacement of malfunctioning or dried up wells, they may also need complementing where they have once been carried out. Thus, after a single blanket testing covering five million wells by the government of Bangladesh in 2000-2005, no further country-wide public programs have been undertaken as of the time of writing. In consequence, recent estimates suggest that more than half of currently used tube wells in Bangladesh have never been tested for arsenic (van Geen et al., 2014). Public provision has hence not fully met the need for testing, and a permanent network of test providers may be required to ensure coverage. This prompts the question whether cost-shared private provision might provide a sustainable complement to public provision, and whether there is the prospect of a market for arsenic tests in which local entrepreneurs would have an incentive to seek out untested wells.

In this paper, we conduct a randomized control trial conducted in 26 villages in Bihar, India, from 2012-2015. In order to elicit demand, we offered tests at prices between Rs. 10 to Rs. 50, randomized at the village level. The highest price level (Rs. 50) was slightly less than one day of per capita income in Bhojpur district in 2011-12 (Rs. 58)., or one-third of the full cost of goods and services.[3]

We find that there is a considerable demand for arsenic testing: at the mean across price groups, and over the duration of our intervention, 45% of households purchase the test. However, demand drops steeply with price, in line with demand elasticities found in other studies of highly effective preventive health care products (Cohen and Dupas, 2010; Kremer and Miguel, 2007).[4][5] We repeat the sales offer two years after the initial campaign, at the same (nominal) sales price and record additional demand , with overall coverage rising from 27% to 45%.[6]

Our study further contributes to the literature by investigating how households respond to the information on environmental quality. We use the quasi-experimental variation caused by the stochastic incidence of arsenic to identify the behavioral responses of households. In a follow-up survey conducted three months after the first wave of test offers, about one third of households whose wells had unsafe levels of arsenic reported having switched to a safer tube well for their drinking and cooking water needs. This avoidance rate is in line with previously reported switching rates, though at the lower end of the spectrum (Ahmed et al., 2006; Bennear et al., 2013; Chen et al., 2007; George et al., 2012a; Madajewicz et al., 2007; Opar et al., 2007). Evidence on significant switching in response to subsidized diagnostic test for arsenic stands in contrast to limited evidence on behavioral responses (i.e. seeking malaria treatment) to the information provided by subsidized diagnostic test for malaria in Kenya (Cohen et al., 2015). We find no effect of price paid for testing on the probability of switching to safer water sources, which is an important finding in assessing

[3]Daily per capita income is calculated by dividing annual per capita income by 365 days. Per capita income in Bhojpur district in 2011-12 was about 14% less than the state average. Data is available at `http://www.finance.bih.nic.in/Documents/Reports/Economic-Survey-2016-EN.pdf`

[4]To our knowledge, no study has previously estimated the demand curve for diagnostic testing of water source quality for arsenic. One related study by George et al. (2013) considers demand for arsenic testing at a single fixed price in Bangladesh, and shows that education and media campaigns increased adoption.

[5]Due to limitations in the data collection, we prefer to use the recall data on sales offers and purchases to estimate demand. We look into the reliability of the sales offer and purchase recalls in our demand estimates by analyzing it extensively in Appendix A.

[6]The observed additional demand is remarkable because the opportunities for learning are somewhat circumscribed by the fact that arsenic tests are an experience good only in a very limited sense. Thus, once some consumers buy tests, others may observe that neighboring wells test positive for arsenic, and may learn about opportunities to switch – but because the health impact of arsenic are slow in onset, health benefits are not immediately observable.

cost and benefit of programs that provide information on environmental quality.

In a novel finding, we find strong evidence of selective recall and concealing of test results. About half of the households whose wells tested *unsafe* were unable to recall their well status correctly (with no significant difference in case of safe wells). We also document that households actively conceal information on their well's arsenic level when tests revealed their well water to be high in arsenic, by discarding placards attached to high arsenic wells. Stigma, concerns over reduced property value, or obstacles to switching might explain this choice. We present evidence that wealthier households are more likely to hide adverse information.

Two limitations arising from the study's implementation are worth noting. A review of the field work finds that in the first phase of test sales, enumerators did not systematically collect data from all households approached with a sales offer. To mitigate the resulting obstacles for demand estimation, we collected recall data on sales offers and purchases during the second offer phase. Secondly, an attempt to create a well owner-level panel to link households across the two rounds of test offers (about two years apart) was unsuccessful, since well tags attached during the first phase proved to be far less durable than expected, and could not be comprehensively tracked.

The remainder of the paper is structured as follows. Section 2 discusses the details of the experiment, data, and empirical specifications. Results are presented in Section 4, and Section 5 concludes.

2 Details on Experiment, Data and Methodology

2.1 Study setting and sample

Our study is set in a region in the Indo-Gangetic plains in Bihar, India, where arsenic levels are elevated in a significant proportion of drinking water wells. Chakraborti et al. (2003) first documented that a large number of wells in the region showed elevated arsenic levels by extending their testing campaign upstream along the Ganges from the state of West Bengal. Arsenic testing is a new service in the study area: tests are not available in the private market (nor are they elsewhere in South Asia), and while Nickson et al. (2007) report that about 5,000 wells have been previously tested in the general area, it has not previously been covered by any government-sponsored blanket testing of wells.[7] Within the general study area, we selected Bhojpur district to conduct our intervention. Within this large district (1,045 villages are recorded in the Census), we select a study area of four blocks (sub-districts) adjacent to the village where arsenic was first reported in Bihar (Chakraborti et al., 2003). We discuss external validity of our results below. Within these, we choose 26 villages of moderate size (50-400 households) for this study, based on a high probability of arsenic incidence, as indicated by distance from the river.[8] Our endline survey identifies 4,084 well-owner households in total.[9]

To elicit demand, we used a simple revealed preference approach – namely, making take-it-or-leave-it offers of arsenic tests at a certain price to households in the sample villages. As is obvious, a take-it-or-leave-it offer elicits only a bound on each household's willingness to pay. For instance, if a household accepts to purchase a test at Rs. 30, we can only infer that its willingness to pay was at least Rs. 30. Similarly, rejection only suggests that willingness to pay was less than the asking price.

We randomly assigned each village to one of five price levels at which households were offered arsenic tests for purchase, rising from Rs. 10 to Rs. 50, in increments of ten. It was felt that offering different prices to households *within* a given village would be seen as violating fairness norms, and would deter purchases.[10] We therefore chose not to randomize our prices within villages. The highest price (Rs. 50) was chosen based on initial local focus group discussions; it is slightly lesser than the average daily per capita income of Rs. 58 in Bhojpur district in 2011-12. Revenue from test sales was used to partially cover the enumerators' salaries and travel cost. The cost of the test kits alone was about USD 0.35 (about Rs. 21 at January 2014 exchange rates); the COGS for

[7]Nickson et al. (2007) report arsenic testing of about 5,000 wells in six out of 14 sub-districts of our study district. The sub-districts were not identified in the study, and it is hence not possible to precisely compare the number of wells tested to the number of local wells. However, the share of wells tested was certainly a small fraction of the 335,000 wells reported in the 2011 Census for the entire study district. 26% of wells tested unsafe.

[8]The original intention was to work in a sample of 25 villages, i.e., five villages in each of our five price groups. However, enumerators erroneously visited two villages of the same name during initial field work. We included the additional village as the 26th for the rest of the program.

[9]We cross-checked the number of households recorded in our study against 2011 Census data for 21 out of 26 villages that could be matched to the census. For these villages, the census shows 4,497 households that own a hand pump, whereas we record 3,322 attempted sales in the same 21 villages - that is, 74% of the census population. The discrepancy is in significant part due to the failure to include entire parts of a few villages, because enumerators believed these to be distinct villages.

[10]This consideration obviated the use of alternative techniques for eliciting willingness to pay, such as the Becker-DeGroot-Marschak (BDM) mechanism and other auction-based methods. In any case, auctions would have been unlikely to be efficient mechanisms, given the potential buyers' uncertain and likely correlated beliefs over the value of arsenic tests.

testing, including wages, quality control, and test result placards amounted to USD 2.26 (Rs. 136). Metal well tags intended purely for data collection added an additional USD 0.48 (Rs. 29). The highest price charged therefore more than covered the cost of the test kits, and about one-third of the entire COGS. We did not add a treatment arm that would have offered tests free of charge, because of a strong expectation that take-up would be near-universal at zero cost. This expectation was based on prior experience in arsenic testing campaigns, and was confirmed further when free tests were offered with near-complete take-up in four pilot villages visited for the design of our experiment. It is also in line with broader evidence from the lab (Shampanier et al., 2007) and from field experiments (Cohen and Dupas, 2010; Kremer and Miguel, 2007).

2.2 Implementation – testing campaign and surveys

We used Arsenic Econo-Quick field test kit which is considered as a cost-effective and time-saving alternative to lab-based testing. Previous laboratory inter-calibrations have shown that the kit correctly determines the status of about 90% of wells with respect to the WHO guideline ($>$ 50μg/l arsenic) (van Geen et al., 2014; George et al., 2012b). Testers were locally recruited from among college graduates, and trained prior to the roll-out of the campaign. Testing then proceeded in two waves. The first wave of testing was conducted in 2012-13. Approximately three months after testing was completed, a follow-up survey was conducted to record whether households had switched to a new well. In this follow-up, we attempted to interview all households who purchased test in the first round and we could record switching for about 90% of the sample. The second wave was conducted in 2014-15, about two years later Tests were offered again in the sample villages and all the households in sample villages were surveyed. The timeline of field work is provided in Table 1– henceforth, for simplicity, we refer to the first round of testing as having taken place in 2012, and the second round, in 2014.

The first wave of testing began with focus group meetings in each village. To increase awareness of the arsenic issue, a large poster was put on display, showing a satellite image of a pilot village along with color markers indicating the arsenic status of tested wells (Figure 2). The poster served the additional purpose of making tangible the great spatial variation in arsenic contamination, and the resulting opportunities for well switching. Following the focus group meetings, testers began to offer tests door-to-door; where a sale was made, tests were conducted using a reliable field kit that requires approximately 15 minutes per test (van Geen et al., 2014). The protocol foresaw that for all households approached with a test offer, GPS locations and basic data on the household would be collected. However, in contrast with what was intended, testers did not record data from *all* households that did not purchase a test. We discuss the resulting challenges for demand estimation, and our solution approach, in detail in Appendix A.

During the initial wave of test offers, a total of 1,212 tests were sold across the 26 sample villages (Table A1, Column 3). At the time of testing, and during the pre-testing focus group discussions, two arsenic cutoffs were systematically conveyed to the households verbally, explaining the arsenic safe, moderate and high values with the color code - Blue, Green and Red, respectively. The results of each test were posted on the pump-head of the well that was tested, with an easy-to-read metal placard, color coded red for unsafe wells ($> 50\mu$g/l arsenic), green for 'borderline safe' wells where arsenic is of some concern ($>$ 10-50μg/l), and blue for safe wells ($\leq 10\mu$g/l) (Figure 3). The cut-off values were chosen to correspond with the Indian national safety standard for arsenic of 50μg/l that was current as of the time of the test campaign, and the WHO guideline of 10μg/l (the government of India – unlike the government of Bangladesh – has since matched its standard to the WHO guideline). The choice of placard color and design was based on the Bangladesh government's blanket testing program which tested wells for arsenic across the country during 2000-2005 (Ahmed et al., 2006) and recent public health interventions on arsenic contamination (van Geen et al., 2014, 2016). Unique well ID tags were also attached to each pump-head in anticipation of a future response survey. Regrettably, well ID tags proved to be less durable than hoped, and only less than 5% of tags placed in 2012 were still attached in 2014.[11] Hence, it was hence not possible to reliably link wells across survey rounds.

Immediately after the first wave of arsenic testing was completed, village-level maps were exhibited in each village, showing the approximate geo-locations of safe, borderline safe and unsafe wells, with the goal of illustrating, where relevant, that the proximity of safe wells would make well-switching feasible. Geo-locations were jittered to preserve anonymity. During home visits,

[11]We fixed a thin strip of steel on the head of the wells at households we visited the household first time (this is shown in Figure 3). It contained a unique well ID and was fixed to the well-head with a metal wire. Over the two year period between surveys, most of these well ID tags disappeared. We assume that this was due to a lack of durability. Moreover, unlike the arsenic test result placards, these well ID tags did not provide any information about water quality so households would have had little reason to actively seek to keep them, and may have removed them if they proved an inconvenience.

households were alerted to the fact that switching from unsafe or borderline safe wells to neighboring safe wells would be an effective way to avoid arsenic exposure. The first phase of the project concluded with a follow-up visit conducted approximately three months after testing was completed. Enumerators visited all households who had purchased the test and collected information on their current source of water for drinking and cooking purposes.

In our sample, about half of the wells are not visible from the outside. However, well sharing with others is readily possible since houses are close to each other and people interact on a regular basis in a small village economy, even if property rights on these wells are well defined. There are a small number of communal/public wells in about half of the villages (no more than one or two wells at a maximum), e.g., wells within the premises of a temple or school. We tested all these wells for free and if people switched to a safe community well in response to high arsenic outcome in their private wells, it was captured in our data collection. Since arsenic incidence is spatially stochastic, it is unlikely that a household could successfully predict his own well type by looking at test outcomes of a nearby well.

In a second phase, commencing in 2014 – some two years after the initial visits – we offered the tests again in the same set of villages, and at the same nominal price assigned initially.[12] Across the 26 villages, a total of 4,084 households were approached with the intention of making a sales offer (Table 4, Column 4). In the second round, data were collected systematically from every household where a respondent could be interviewed, including from households that did not wish to buy the tests. Each house was visited at least two times to ensure high coverage. After two visits, about 14% of households could not be surveyed because no adult member was present or willing to answer questions; sales offers could be completed in 3,528 households. The enumerators reported that, to avoid embarrassment, some households who were unwilling to purchase tests at the asking price avoided being interviewed. For a conservative demand estimate, we therefore work throughout with the number of households approached for sales, rather than the number of households where a sales offer could be completed. A total of 719 tests were sold in this second phase (Column 5). The household survey administered in the second round gathered socio-economic and demographic information, along with GPS locations of the wells. It also collected information on recall of tests

[12] Considering inflation in rural Bihar during this period, the lowest price of Rs. 10 and highest price Rs. 50 during the second round would be equivalent to Rs. 8 and Rs. 41, respectively, in the first round. As we argue in Appendix B, this may partly explain additional demand at the time of the repeat offer.

being offered and purchased in 2012, along with recall of test results. This recall data allows us to work around some of the constraints posed by the implementation issues encountered during the first wave of offers.

2.3 Summary statistics

Summary statistics from the 2014 survey show modestly well-off village communities (Table 2). Households are of moderate size (3.9 members on average). Most (89%) own at least one mobile phone, and most (70%) live in houses made from durable building materials ('pucca'). Ownership of bikes (68%) and cows (67%) is common, though fewer households own consumer durables or have access to sanitation, and very few own cars.

Table 2 also shows a randomization check on observables. We calculate a normalized asset index with house characteristics and assets information using standard principal components approach (Filmer and Pritchett (2001)), and estimated coefficients are provided in Column 4. As Table 2 shows, price category dummies are jointly significant at the 90% level for two out of the eleven variables tested. The two instances where there are significant differences (ownership of cars and access to sanitation) appear isolated, and would suggest opposite signs in a relationship between price and ownership. There is therefore no indication that the price groups in question are systematically any more or less wealthy than the other groups.[13]

To give a sense of the external validity of our results, Table 3 compares household wealth proxies in the 2011 Census for our sample villages, the four blocks that nest them, Bhojpur district, and the state of Bihar. As is evident, households in our sample villages are similarly well-off as the mean household in the blocks (Panel A) and Bhojpur district (Panel B). They are, however, better off than the average household in Bihar, with a far higher share of houses made from durable materials, greater literacy, and ownership of household assets up to 10pp higher for many categories (Panel C). While we show below (Table 6) that purchase decisions at high price levels does not correlate with assets, we might expect demand in our sample villages to be representative of Bhojpur district, but at weakly higher than in Bihar at large.

[13]Note in Table 4 that the total number of households varies significantly across price groups, with larger villages in the low-price groups. However, Table 2 demonstrates that other demographic characteristics and asset ownership were similar across villages in different price groups. We also find no correlation between mean asset index and village size in additional tests.

2.4 Empirical specification

We have two sources of exogenous variation in this study – experimental variation in prices and quasi-experimental variation in arsenic incidence. We use the village-level price variation to estimate the demand, and the household-level arsenic incidence to analyze the behavioral response to the information revealed by tests. Throughout this paper, we analyses data using OLS.

We estimate the demand for arsenic testing with the following three specifications using (1) a continuous price, (2) a dummy variable indicator for high price and (3) price level indicators, respectively (Eq. 1 – Eq. 3)

$$Purchase_{iv} = \beta_0 + \beta_1 price_v + \epsilon_{iv} \tag{1}$$

$$Purchase_{iv} = \beta_0 + \beta_1 \mathbb{1}(price_v \geq 40) + \epsilon_{iv} \tag{2}$$

$$Purchase_{iv} = \beta_0 + \boldsymbol{\beta \alpha_p} + \epsilon_{iv} \tag{3}$$

Here, $Purchase_{iv}$ is a binary variable showing whether household i in village v purchased the test, when offered at a price p_v ($p \in P\{Rs.10, Rs.20, Rs.30, Rs.40, Rs.50\}$). $price$ denotes a continuous price variable, while $\mathbb{1}(price_v \geq 40)$ and α_p represent high price dummy and a set of price level dummy variables, respectively. Our estimator of price sensitivity to demand is the coefficient on the price variable. ϵ_{iv} is the error term.

Next, we estimate a model of avoidance behavior, where the binary outcome variable $Switched_{iv}$ shows whether a household i in village v switched to a safe well or not. $TestOutcome_{iv}$ shows the arsenic status of the baseline well of the respondent households. We estimate the effect of information provided by the diagnostic test with the coefficient on the $TestOutcome_{iv}$.

$$Switched_{iv} = \beta_0 + \beta_j \mathbb{1}(TestOutcome_{iv} = HighArsenic) + \epsilon_{iv} \tag{4}$$

Using a similar specification with a price and asset index interaction term, we test whether behavioral response depends on the price paid to obtain the information i.e. whether switching is correlated to the price paid by households.

Our final investigation is about concealing and selective recall of test outcomes – where households fail to retain the physical marker displaying arsenic test outcome or fail to recall the test result correctly. Since we cannot link households across two years (i.e. first and second round), we pool the cross section data from both rounds (i.e. actual measurement in round 1 and recorded evidence/reported result in round 2).

We estimate the concealing and selective recall for each test outcome category, by regressing test outcome dummy indicator (pooled from round 1 and round 2) on round 2 dummy. This regression is equivalent to a t-test on the equality of proportion of corresponding arsenic test outcomes in two groups - (1) as tested in the first round i.e. in 2012, and (2) as found with evidence or as reported by the households in the second round i.e. in 2014.

$$TestOutcome_i = \beta_0 + \beta_1 Round2_i + \epsilon_i \tag{5}$$

where $TestOutcome_i \in T\{High, Moderate, Safe\}$

β_1 denotes the change in the proportion of particular test outcome from round 1 to round 2. β_0 denotes the proportion of that particular test outcome in round 1. We of course limit the sample to households who purchased the test in round 1, since we do not know the arsenic status of wells in households who did not purchase the test. With a similar specification, we use interaction of $Round2_i$ with asset ownership to test whether concealing and selective recall of test outcomes is correlated with asset ownership.

In all regressions, we report cluster bootstrapped standard errors to account for randomization at the village level. For estimated coefficients in the demand equations, we also calculate wild bootstrap-t p-values as a robustness check (Cameron et al., 2008).

3 Results

3.1 Demand for well arsenic testing

Demand for fee-based arsenic tests in the study area is substantial. Overall, a total of 1,857 tests were sold at randomly assigned prices across the 26 sample villages over the entire duration of the program (2012-2015). This implies that arsenic testing covered about 45% of households approached for sales (Table 4, Column 10).[14] An example of test results in one village is provided in Figure 1; a map displaying the proportion of safe, unsafe, and untested wells in each village is shown in Figure 4. It pools results from the first and second test phase. In total, using the national and WHO thresholds of 50 and 10μg/l, respectively, 50% of wells tested 'safe' ('blue'), 31% tested 'borderline safe', and 19% tested 'unsafe' ('red'). As expected, test results varied over small distances, and there is a wide spread in the shares of unsafe wells across villages, ranging from 2% to 77%.

Demand in the first round of sales alone was 27% across price groups in our preferred recall estimate (Column 7). Demand at the time of the second offer was 18%, after adjusting for repeat purchases (Column 8). As noted, demand estimation for the first round of sales is complicated by incomplete data collection. In Appendix A, we discuss how we address the problem, and assess robustness. In the following, we work with recall data systematically collected during the second test wave to determine 2012 demand, both because it is more internally consistent, and because it yields more conservative estimates (overall demand was 30% using an alternative approach of imputing demand from 2012 sales and the 2014 sample size).[15]

In line with prior research on preventive health products, we find that demand for arsenic testing is highly sensitive to price (Figure 5, Table 5). When we test for the price effect on demand using dummies for each of the five price levels offered, we find the expected negative signs, but are unable to reject equivalence in all cases. However, estimated coefficients for continuous price and high price dummy variables are statistically significant and provide additional confidence in our results. The mean elasticity across sales at different price levels in our data is -0.36 in the first round, and -0.47 in the second round. At the lowest price of Rs. 10 (USD 0.15 at market rates at the time of the repeat offer), 40% of households purchase the test after one offer, and 69% after two offers (Table 4, Columns 7 and 10). While our experiment did not include an arm with zero price offer, uptake of free tests can be assumed to be nearly 100% (as discussed in Section 2.1). Thus, while there

[14] To estimate total coverage after two offers, we add first and second-round coverage, correcting for repeat purchases. We define second-round purchases to have been repeat purchases in 74 instances where households recall having bought the test in 2012, and purchased another test in 2014. Households had been advised that, since arsenic levels in ground water are stable over time, wells need not be tested repeatedly.

[15] Note that the recall data appears to show steeper demand than would be implied by 2012 actual sales divided by 2014 sample size (Figure A1). Relative differences in the propensity to recall test purchase across price levels might bias our estimate of first-round demand (Column 1, Table 5), if households in the lowest price bins recalled sales more accurately. Because there is apparent higher recall in lower price groups, we conduct a sensitivity analysis and confirm that our estimate is robust to excluding the lowest two price levels. Results are available upon request.

is significant demand at Rs. 10, charging this small amount, rather than offering the test for free, reduces coverage after two sales offers by about one-third. Demand further drops precipitously at higher prices, and at Rs. 50, reduces to less than one-sixth of households after one offer, and less than one-quarter after two offers.

This pronounced sensitivity is in line with demand behavior observed in other recent studies of preventive health products such as ITNs or rubber shoes in developing countries (Cohen and Dupas, 2010; Dupas, 2014b; Kremer and Miguel, 2007; Meredith et al., 2013). The fact that arsenic tests arguably were less well-known to consumers than products studied elsewhere was not reflected in distinctly higher price elasticity.[16] This is comparable to outcomes in our experiment at a price of Rs. 50 and after one sales offer: demand of 15% at a price equivalent to 111% of average daily income, and 30% of the full cost of goods and services.

Our demand estimates compare well with results shown by George et al. (2013), who estimate demand for arsenic tests in Bangladesh at a single price point of USD 0.28 in 2011 – the equivalent of about Rs. 10 in 2014 in our setting. George et al. find 53% uptake in the control group, where no dedicated awareness campaign is conducted, and 93% uptake in each of two treatment arms with an awareness campaign. Our demand estimate at Rs. 10 is in between these two values after two offers, but far below after a single offer. This is perhaps intuitive: arsenic test were not widely known in our intervention area, while George et al. (2013) worked in Bangladesh, where government-sponsored blanket testing and many other interventions have significantly raised awareness of arsenic.

In each village, the initial test offer was followed by a repeat offer after some two years had elapsed – at the same (nominal) sales price. Our purpose in re-offering the arsenic test was to assess whether additional demand (i.e. from households who did not purchase in the first phase) could be elicited. We repeated the offer *at the same nominal price charged initially*, as opposed to repeating it at a *uniform* price as in Dupas (2014b). This allows us to study the (reduced-form) effect of making a repeat offer at different price levels, a question of immediate policy interest. We find that repeating the offer after a two-year delay did indeed generate substantial additional demand.

[16] Perhaps the most natural comparison in terms of the nature of products offered is to Berry et al. (2012), who study willingness to pay for water filters to remove pathogens in northern Ghana. Berry et al. report that, while 95% of respondents had non-zero willingness to pay (an analogue of near-universal take-up at zero cost), charging a price equivalent to 116% of daily income (or 30% of the filter's cost) reduced demand to 21%.(Demand figures from Dupas (2014a). Figures are not directly reported in Berry et al. (2012).) Share of income is based on USD 4.20 (GHS 3) price and 2010 (current) per capita GDP of USD 1,323.

Thus, purchases at the time of the second offer raise total coverage by some 18 percentage points (pp), from 27% to 45% (Table 4, Columns 7 and 10). Demand is more price-sensitive than at the first offer (Figure 5). However, we observe an effect of repeating the sales offer on coverage at any price level, with increases ranging from 70% of the original sales at Rs. 10 to 19% at Rs. 40. The per capita real income in Bihar rose at a rate of about 10% per year between 2012 and 2014, and thus the 2014 prices were lower in real terms. However, real price difference alone does not seem sufficient to explain additional demand, especially at lower prices. We provide a detailed discussion on the choice of keeping nominal price constant and two potential channels explaining additional demand in Appendix B.

3.1.1 No buyer selection at different price levels

We test whether wealthier households are more likely to purchase the test at higher prices, by regressing purchase decision on a set of interactions of price and asset index. To address concerns about low statistical power, we first run this analysis with continuous price as well as high price dummy variables. Table 6 shows that, independently of the asking price, wealthier households were more likely to buy. However, the interaction terms between the continuous price variable and asset index are statistically insignificant and small in magnitude (Column 1): a two standard-deviation increase in the asset index attenuates the main effect of price on demand by only about one-tenth. We find consistent results when using high price dummies (Column 2 and 3) or our main specification using dummies for each price level. Hence, purchase decisions at higher price did not correlate with wealth. In all three specifications, coefficient on the interaction term is not only not significant, it is also small. For instance, in Column 1, even at 95% of the asset index distribution, the magnitude of the estimated interaction term would be less than 10% of the price effect

To investigate further, we test how sales price correlates with buyer characteristics in terms of different dimensions of the asset index - that is, different household wealth proxies. Appendix Table C1 shows regression results for buyers who purchased the test in either round. As is evident, few asset categories are correlated with sales price. For those that do correlate, selection was limited to the two highest price levels. Given the large drop in demand associated with a price increase from Rs. 10 to Rs. 20 (13pp, or 45% in relative terms), it is perhaps surprising that there is virtually no distinction in observed asset ownership between households that buy at these price levels.

The absence of a wealth pattern suggests that, either, purchasing decisions were driven by different valuation of the product among similar households, or marginal utility of consumption differed in ways that do not correlate with characteristics we observe. As shown in Column 3 in Appendix Table C1, investment in sanitation – i.e. having a latrine facility in the house – is correlated with purchase decisions at high price levels (about one household in three among those who buy at Rs. 10 owns a latrine, but two in three do among those who buy at Rs. 50). This result might well speak to a concern over hygiene and health driving both investments.

3.1.2 No residential sorting

We test whether households can predict arsenic contamination, and potentially, sort accordingly in choosing their residence. As noted, the distribution of arsenic in groundwater wells is hard to predict; it would be surprising if we were to observe sorting. Appendix Table C2 confirms this notion, in keeping with findings in Madajewicz et al. (2007). There is no relationship between well characteristics (age, depth, and cost) and the probability of high contamination – that is, households do not appear to specify well design to effectively avoid arsenic (Column 1). Nor is there a distinct relationship between asset ownership and arsenic status of wells that would suggest residential sorting (Column 3 and 4). We also show that there is little correlation between test price and well quality (Column 2).[17]

3.2 Behavioral response to arsenic content information: well switching

We next consider how households use the information revealed by arsenic testing, leveraging the quasi-experimental variation induced in the type of information revealed by the spatially stochastic arsenic incidence. Particular importance attaches to whether households switch from highly contaminated wells to safe water sources. Within the context of the wider literature on preventive health products, this can be viewed as equivalent to behavioral issues surrounding the use of information. Thus, it is the act of switching to a safe water source that brings about health benefits after the purchase of a test – and switching imposes further inconvenience cost. Similarly, after the purchase of an ITN or a drinking water filter, it is the act of sleeping under the net or filtering water

[17]Given the small number of high-arsenic wells, tests are run separately for each asset category to avoid over-fitting (Column 4). Due to multiple hypothesis testing, the standard errors reported in Appendix Table C2 are too small. We omit any adjustment because the absence of sorting emerges even when precision is overstated.

that generates health benefits, and each may be associated with inconvenience to a degree specific to the particular context.

Among households that purchased the test in 2012, high arsenic well owners reported 30.5% (percentage points) higher switching to a safer drinking water well, when compared with - very rare - baseline switching among households whose well turned out to be safe. Table 7 estimates the behavioral response to the information provided by arsenic testing in terms of switching from high arsenic wells (red) to other safe (blue) or moderately contaminated (green) wells. Column 1 shows that 24% of households whose wells tested high or moderate in arsenic switched to a safe well; 28% of well-owners switched when we only consider high arsenic wells. The switching rate from moderate arsenic to safe wells is thus lower than the switching rate from high arsenic to safe wells, suggesting that the behavioral response to information depends on the level of contamination, as observed inMadajewicz et al. (2007). Columns 3 and 4 show estimates for switching to a well which is either safe or contains only moderate level of arsenic.). Note that there is little switching reported from safe wells (only 2 out of 633 households with a safe well switched to another safe well i.e. 0.3%).

Overall, this is a low switching rate, but not an atypical response. A number of similar studies in Bangladesh have reported switching rates of 26-39% (Ahmed et al., 2006; Bennear et al., 2013; Chen et al., 2007), although others find higher rates, in between one-half and two-thirds of affected households (George et al., 2012a; Madajewicz et al., 2007; Opar et al., 2007). In line with prior evidence (Chen et al., 2007; Opar et al., 2007), we find that distance to safer wells is an important predictor of switching (Figure 6). The somewhat subdued response to information could be related to the limited number of wells identified to be safe, because of lower take-up of the for-fee service, as opposed to blanket testing.[18] Relatively lower switching in this study could also plausibly be due to restrictions on sharing water based on caste affiliation and religion. – Among households in our survey, 90% report that they prefer to exchange water within their own caste or group of relatives. Similarly, in Uttar Pradesh, a state adjacent to Bihar, caste in particular has been found to be a major factor in impending water trade within a village (Anderson, 2011). We also note that the margin of effort in switching after the information is revealed by arsenic testing may be significantly higher than it is in using many health products. Our setting may be closer to the

[18]This also highlights the potential for a positive externality where arsenic tests are accessible to all well owners.

context of encouraging households to purchase anti-malaria therapy after a rapid diagnostic test for malaria (Cohen et al., 2015).

3.3 Price paid for information and behavioral response

We further find that the propensity to switch does not depend on the purchase price (Table 8). That is, in the case of arsenic testing, the behavioral response to environmental quality information does not vary with the price paid to obtain the information. To guard against concerns that the tests for individual price categories shown in Table 8 might be under-powered, we confirm that there are no significant differences when we regress on continuous price as well as on a dummy variable for 'high' price level. This finding implies an absence of screening or sunk cost effects. Both effects would tend to increase usage with price, and imply that highly subsidized provision might lead to 'overinclusion' of those who do not sufficiently value the information provided.[19] Our result further bolsters recent findings that have suggested that, for preventive health care products, there is little empirical evidence of overinclusion in subsidized provision (Cohen and Dupas (2010); Dupas (2014a) – see Berry et al. (2012) and Ashraf et al. (2007) for experimental evidence of screening, but not sunk cost effects).

3.4 Concealing and selective recall of high arsenic result

We find strong evidence of selective recall, and find that households not only avoid reporting adverse arsenic test outcomes, but take direct action to remove markers of unwelcome results. When visited at the time of the second sales offer, households who purchased a test when the first sales offer was made two years earlier were asked "Do you know the status of this well with respect to arsenic?". About 26% of households responded that their water was not fully safe (and about 15% stated that they could not recall). However, the actual test outcome distribution in the first round of tests showed that the proportion of highly and moderately contaminated wells was about 50%.

Table 9 offers a test for selective recall that builds upon this observation. It compares the proportion of test outcome in each category of arsenic contamination levels (Red/high, Green/moderate, and Blue/safe) observed in first-round tests recorded in 2012 to the proportion of corresponding test

[19]In our setting, the respective arguments are as follows: 'those who decided to buy at high price care more about health from the outset, and will therefore be more likely to switch wells'; and 'those who buy at high prices have invested more in the test, and will hence more highly value the information it yields'.

outcome *recalled* in 2014. We adduce the information on arsenic status of a well in three different ways – namely, (1) those households where the test placard was still affixed to the well; (2) those where the placard had been removed from the well, but was still kept in the house; and (3) those where the placard was neither on the well nor kept in house, but the respondent reported being able to remember the arsenic contamination status.

As is evident, the proportion of respondents who purchased a test in the first round and believed their wells to be unsafe when visited during the second survey round was consistently some nine to eleven percentage points lower than the true proportion of red tests recorded in the first round (Columns 1, 4, 7, and 10). It is particularly striking that such a discrepancy exists even among households where the test placard was still attached to the well: since it is inconceivable that red tags are more likely to be accidentally lost than others, this is clear evidence of intent either to hide the well's status, or to avoid being reminded of it (Column 1). The magnitude of the effect is very substantial: 20% of wells tested 'red' in 2012 – and hence, a decrease of the share of 'red' wells by about 9-11pp implies that about half of the households with wells that were high in arsenic intentionally sought to hide the test outcome. We also note that respondents who did not produce a placard tended to preferentially indicate that wells were tested 'green' – suggesting that households prefer to claim a moderate arsenic level in their highly contaminated wells (Column 8). Conversely, as Appendix Table D1 shows, wells in households that opted to repeat the arsenic test in 2014 were more likely to have tested 'green' than those only tested once. It is possible that some households opted to purchase another test because they could not recall the result of the earlier test. However, more specifically, the higher proportion of repeat purchases among 'green' wells that tested borderline safe may suggest that some households who initially received 'mixed news' sought to resolve any uncertainty, and hence, were more likely to purchase the test again than those who received clear 'good' (i.e. blue) or 'bad' news (i.e. red).

These findings are consistent with general theoretical and experimental evidence of 'self-serving bias' and 'over-confidence' (see, e.g., Eil and Rao (2011)). More practically, we note that efforts to hide unsafe well status could be related to low well switching rates in various ways. It could be that well owners hide bad news because there is (for unrelated reasons) a high private or social cost to take action to remedy the situation, as evidenced by the relatively low switching rates reported above. It is also possible that both the reluctance to share and the propensity to hide bad news speak to a social stigma or material loss (e.g., in house value – for the United States, Boyle et al. (2010) find a temporary 1% reduction in residential sales values associated with a 10μg/l increment in arsenic levels) being attached to owning an unsafe well. We note that there is some indication that wealthier households may be more likely to hide adverse test results, potentially because of greater concerns over stigma or material loss. To show this, we compare test results and recall as above for high arsenic outcome – but distinguish between households that owned and did not own consumer durables (the one asset ownership indicator collected consistently in both survey rounds) (Table 10). As is evident, while all households under-report, households that do own durables are about twice as likely to do so; the difference is significant for the larger samples.

We add two caveats regarding our evidence on concealing and selective recall of adverse outcomes. First, these estimates in Tables 9 and 10 represent concealing and selective recall of adverse test outcomes by households who first revealed their preference for knowing the arsenic status of their well, since we cannot analyze households who did not purchase the test. Secondly, while we cannot correct for attrition during the second-round survey and due to the imperfect recall of test purchase itself, attrition would pose little threat to our results qualitatively: attrition would bias the observed proportion of adverse outcome downward if attrition is correlated with adverse test results. But such a correlation is in itself evidence of selective recall.

4 Summary and Policy Discussion

We have shown experimental evidence from Bihar, India, on the demand for and use of environmental information relevant to health. There is substantial demand for testing wells for arsenic, but it is highly sensitive to price. Compared to the near-universal adoption found under free provision, two-thirds of households purchased tests at the lowest price, and about one-third at the highest price over the duration of the project. We also find that a repeat offer made within two years of the original offer is met with significant demand, raising total coverage by 18pp, from 27% to 45%.

Our results confirm that subsidies remain critical in ensuring high coverage of environmental health information. However, cost-shared provision might still have a useful role to play in providing an ongoing testing service in the absence of or in between public testing campaigns. In particular, one could imagine a business model in which independent testers generate their own wages, while NGOs conduct awareness campaigns, provide test kits, train testers, and implement quality control (for instance, GIS tracking and re-testing of a subsample of wells). Yet, market demand was not quite sufficient to cover wages. In 2012, expected daily revenue was about Rs. 200 (revenue per offer made was highest in the Rs. 30-50 price range, at about Rs. 8; on average, testers visited about 25 households per day). By way of contrast, under local labor market conditions, testers might have expected a daily wage in the range of Rs. 300-400.

Through a follow-up survey conducted after the first wave of sales, we assessed how households respond to the environmental health information furnished through well testing. About one-third of households with unsafe wells switch to less perilous water sources. This is in the lower range of switching rates found in other studies of arsenic testing. Preferences for sharing within caste groups may have limited opportunities to draw water from safer sources – an important consideration for future arsenic testing campaigns in Bihar. We further explore two important and policy relevant aspects of the provision of environmental quality information. First, the probability of switching did not depend on the price paid for the test, implying that in our setting, willingness to pay for information on environmental quality had little impact on the behavioral response to such information.

Secondly, by comparing the share of wells with safe and unsafe arsenic levels between test results collected in 2012 and results recalled in 2014, we show that households avoid reporting adverse test results, and indeed, recall test outcomes strategically or even remove well tags indicating arsenic contamination. This may speak to discomfort with knowledge of well status in the context of low switching rates, stigma, or concerns over property value. The reaction is certainly policy relevant – in particular when allowing for the possibility that the *ex ante* decision to purchase a test might be affected by any motivation to avoid bad news. Secondly, in many settings, local environmental health information generally remains private and strategic revealing by households may defeat mitigation efforts and elevate the damage to others who cannot readily access this information.

References

Ahmed, M., Ahuja, S., Alauddin, M., Hug, S., Lloyd, J., Pfaff, A., Pichler, T., Saltikov, C., Stute, M., and Van Geen, A. (2006). Ensuring safe drinking water in Bangladesh. *Science*, 314(5806):1687.

Ahuja, A., Kremer, M., and Zwane, A. P. (2010). Providing safe water: Evidence from randomized evaluations. *Annu. Rev. Resour. Econ.*, 2(1):237–256.

Anderson, S. (2011). Caste as an impediment to trade. *American Economic Journal: Applied Economics*, 3(1):239–263.

Argos, M., Kalra, T., Rathouz, P. J., Chen, Y., Pierce, B., Parvez, F., Islam, T., Ahmed, A., Rakibuz-Zaman, M., Hasan, R., et al. (2010). Arsenic exposure from drinking water, and all-cause and chronic-disease mortalities in bangladesh (HEALS): a prospective cohort study. *The Lancet*, 376(9737):252–258.

Ashraf, N., Berry, J., and Shapiro, J. M. (2007). Can higher prices stimulate product use? evidence from a field experiment in Zambia. Technical report, National Bureau of Economic Research.

Bates, M. A., Glennerster, R., Gumede, K., and Duflo, E. (2012). The price is wrong. *Field Actions Science Reports. The journal of field actions*, (Special Issue 4).

Bennear, L., Tarozzi, A., Pfaff, A., Balasubramanya, S., Ahmed, K. M., and Van Geen, A. (2013). Impact of a randomized controlled trial in arsenic risk communication on household water-source choices in Bangladesh. *Journal of Environmental Economics and Management*, 65(2):225–240.

Berry, J., Fischer, G., and Guiteras, R. (2012). Eliciting and utilizing willingness to pay: evidence from field trials in Northern Ghana. *Unpublished manuscript.*

Boyle, K. J., Kuminoff, N. V., Zhang, C., Devanney, M., and Bell, K. P. (2010). Does a property-specific environmental health risk create a "neighborhood" housing price stigma? arsenic in private well water. *Water resources research*, 46(3).

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*, 90(3):414–427.

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

Chakraborti, D., Mukherjee, S. C., Pati, S., Sengupta, M. K., Rahman, M. M., Chowdhury, U. K., Lodh, D., Chanda, C. R., Chakraborti, A. K., and Basu, G. K. (2003). Arsenic groundwater contamination in middle ganga plain, bihar, india: a future danger? *Environmental Health Perspectives*, 111(9):1194.

Chen, Y., van Geen, A., Graziano, J. H., Pfaff, A., Madajewicz, M., Parvez, F., Hussain, A. I., Slavkovich, V., Islam, T., and Ahsan, H. (2007). Reduction in urinary arsenic levels in response to arsenic mitigation efforts in araihazar, bangladesh. *Environmental Health Perspectives*, pages 917–923.

Chowdhury, S., Krause, A., and Zimmermann, K. F. (2015). Arsenic contamination of drinking water and mental health. *CEPR Discussion Paper No. DP10978.*

Cohen, J. and Dupas, P. (2010). Free distribution or cost-sharing? Evidence from a randomized malaria prevention experiment. *The Quarterly Journal of Economics*, 125(1):1–45.

Cohen, J., Dupas, P., and Schaner, S. (2015). Price subsidies, diagnostic tests, and targeting of malaria treatment: evidence from a randomized controlled trial. *The American Economic Review*, 105(2):609–645.

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.

Eil, D. and Rao, J. M. (2011). The good news-bad news effect: asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics*, 3(2):114–138.

Fendorf, S., Michael, H. A., and Van Geen, A. (2010). Spatial and temporal variations of groundwater arsenic in South and Southeast Asia. *Science*, 328(5982):1123–1127.

Filmer, D. and Pritchett, L. H. (2001). Estimating wealth effects without expenditure data – or tears: an application to educational enrollments in states of india. *Demography*, 38(1):115–132.

George, C. M., Inauen, J., Rahman, S. M., and Zheng, Y. (2013). The effectiveness of educational interventions to enhance the adoption of fee-based arsenic testing in Bangladesh: A cluster randomized controlled trial. *The American Journal of Tropical Medicine and Hygiene*, 89(1):138–144.

George, C. M., van Geen, A., Slavkovich, V., Singha, A., Levy, D., Islam, T., Ahmed, K. M., Moon-Howard, J., Tarozzi, A., Liu, X., et al. (2012a). A cluster-based randomized controlled trial promoting community participation in arsenic mitigation efforts in Bangladesh. *Environmental Health*, 11(1):1–10.

George, C. M., Zheng, Y., Graziano, J. H., Rasul, S. B., Hossain, Z., Mey, J. L., and van Geen, A. (2012b). Evaluation of an arsenic test kit for rapid well screening in bangladesh. *Environmental science & technology*, 46(20):11213–11219.

Greenstone, M. and Jack, B. K. (2015). Envirodevonomics: A research agenda for an emerging field. *Journal of Economic Literature*, 53(1):5–42.

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

Madajewicz, M., Pfaff, A., Van Geen, A., Graziano, J., Hussein, I., Momotaj, H., Sylvi, R., and Ahsan, H. (2007). Can information alone change behavior? Response to arsenic contamination of groundwater in Bangladesh. *Journal of Development Economics*, 84(2):731–754.

Meredith, J., Robinson, J., Walker, S., and Wydick, B. (2013). Keeping the doctor away: Experimental evidence on investment in preventative health products. *Journal of Development Economics*, 105:196–210.

Nickson, R., Sengupta, C., Mitra, P., Dave, S., Banerjee, A., Bhattacharya, A., Basu, S., Kakoti, N., Moorthy, N., Wasuja, M., et al. (2007). Current knowledge on the distribution of arsenic in groundwater in five states of India. *Journal of Environmental Science and Health Part A*, 42(12):1707–1718.

Opar, A., Pfaff, A., Seddique, A., Ahmed, K., Graziano, J., and Van Geen, A. (2007). Responses of 6,500 households to arsenic mitigation in Araihazar, Bangladesh. *Health & Place*, 13(1):164–172.

Parvez, F., Wasserman, G. A., Factor-Litvak, P., Liu, X., Slavkovich, V., Siddique, A. B., Sultana, R., Sultana, R., Islam, T., Levy, D., et al. (2011). Arsenic exposure and motor function among children in Bangladesh. *Environmental Health Perspectives*, 119(11):1665.

Pattanayak, S. K., Pfaff, A., et al. (2009). Behavior, environment, and health in developing countries: evaluation and valuation. *Annual Review of Resource Economics*, 1(1):183–217.

Pitt, M., Rosenzweig, M. R., and Hassan, N. (2015). Identifying the hidden costs of a public health success: arsenic well water contamination and productivity in Bangladesh. *NBER Working Paper*, (w21741).

Sachs, J. and Malaney, P. (2002). The economic and social burden of malaria. *Nature*, 415(6872):680–685.

Shampanier, K., Mazar, N., and Ariely, D. (2007). Zero as a special price: The true value of free products. *Marketing Science*, 26(6):742–757.

Somanathan, E. (2010). Effects of information on environmental quality in developing countries. *Review of Environmental Economics and Policy*, 4(2):275–292.

Tarozzi, A., Mahajan, A., Blackburn, B., Kopf, D., Krishnan, L., and Yoong, J. (2014). Micro-loans, insecticide-treated bednets, and malaria: Evidence from a randomized controlled trial in orissa, india. *American Economic Review*, 104(7):1909–41.

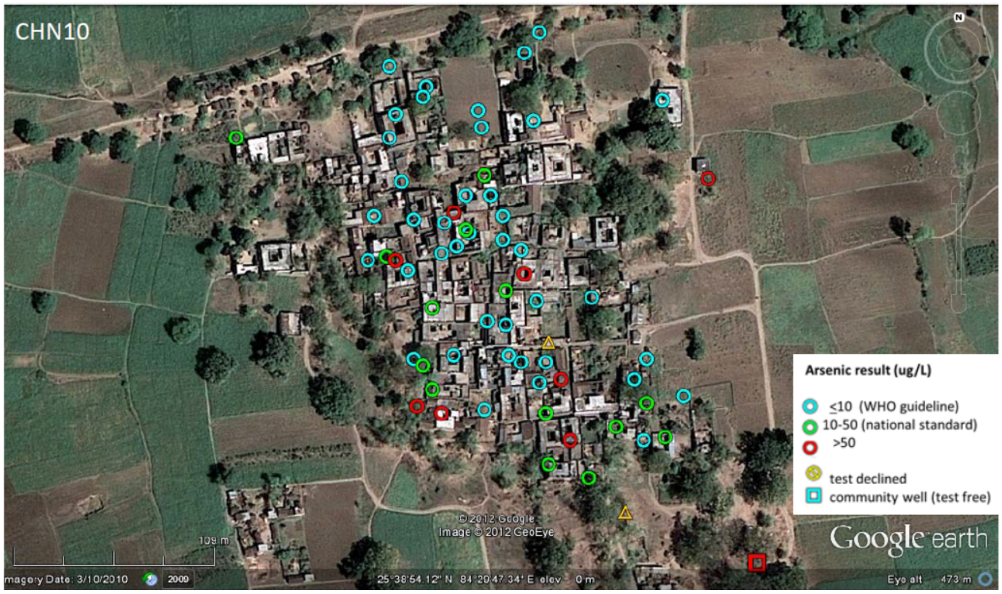
van Geen, A., Ahmed, E. B., Pitcher, L., Mey, J. L., Ahsan, H., Graziano, J. H., and Ahmed, K. M. (2014). Comparison of two blanket surveys of arsenic in tubewells conducted 12 years apart in a 25km^2 area of Bangladesh. *Science of the Total Environment*, 488:484–492.

van Geen, A., Ahmed, K., Ahmed, E., Choudhury, I., Mozumder, M., Bostick, B., and Mailloux, B. (2016). Inequitable allocation of deep community wells for reducing arsenic exposure in bangladesh. *Journal of Water Sanitation and Hygiene for Development*, 6(1):142–150.

van Geen, A., Ahsan, H., Horneman, A. H., Dhar, R. K., Zheng, Y., Hussain, I., Ahmed, K. M., Gelman, A., Stute, M., Simpson, H. J., et al. (2002). Promotion of well-switching to mitigate the current arsenic crisis in bangladesh. *Bulletin of the World Health Organization*, 80(9):732–737.

Wasserman, G. A., Liu, X., Parvez, F., Ahsan, H., Factor-Litvak, P., van Geen, A., Slavkovich, V., Lolacono, N. J., Cheng, Z., Hussain, I., et al. (2004). Water arsenic exposure and children's intellectual function in Araihazar, Bangladesh. *Environmental Health Perspectives*, pages 1329–1333.

Figure 1: Example of well arsenic distribution in a village in Bhojpur district, Bihar (India)



Note: a sample village map from the study is shown with the outcomes of arsenic testing. Red circles denote drinking water wells that are highly contaminated with arsenic; green circles show wells with intermediate arsenic levels; blue circles show wells that are low in arsenic and safe to drink from.

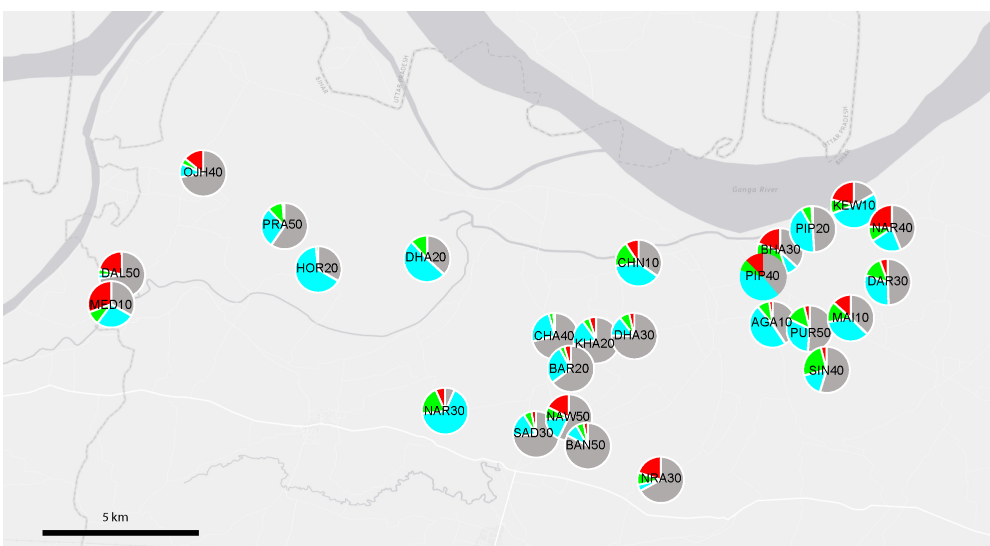
Figure 2: Satellite maps from nearby villages were shown in focus group meetings

Note: village meetings and exhibition of posters showing safe and unsafe wells from near by villages. The geo-location of wells were jittered because of privacy concerns.

Figure 3: Metal Placard showing arsenic status after testing

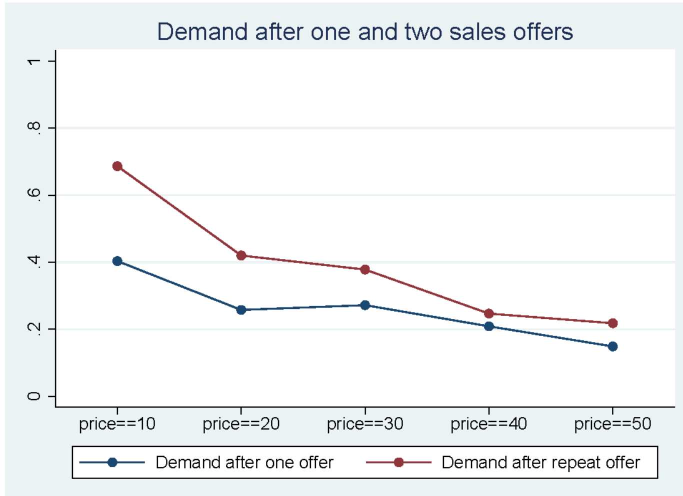
Note: red (Arsenic high), green (Arsenic moderate) and blue (Arsenic low) placards were fixed on the tubewells after arsenic testing.

Figure 4: Map showing village locations with the arsenic test outcomes



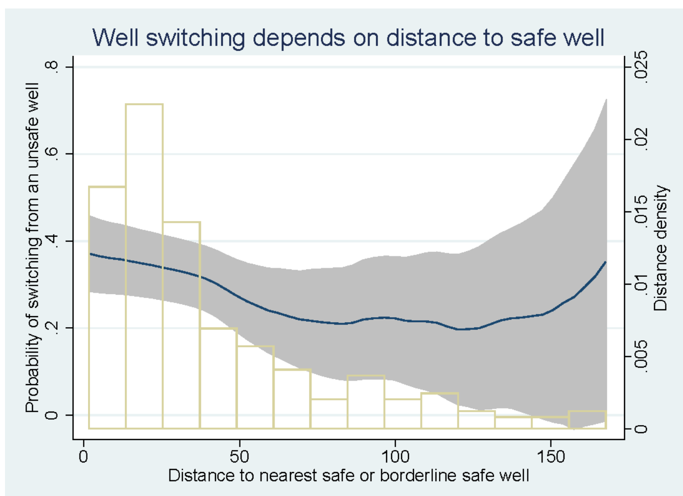
Note: the map shows the location of villages, take-up and outcome of the arsenic testing in subject area. Red (Arsenic high), Green (Arsenic moderate) and Blue (Arsenic safe) colors show the outcome of arsenic testing. Grey color shows the proportion of untested wells.

Figure 5: Demand curves after one and two sales offers



Note: the plot shows demand patterns after one offer (2012) and after two offers. 2012 demand estimates are obtained from recall of sales offers and purchases as measured in the 2014 survey. See Appendix A for discussion.

Figure 6: Switching conditional on distance to blue/green



Note: the graph shows the probability that household whose wells tested 'red' (high arsenic) in 2012 switched to a safer ('blue' or 'green') well, conditional on distance (in metres) to the nearest safer well. Local polynomial fit with confidence interval; histogram of distances overlaid.

Table 1: Fieldwork timeline

August 2012	Arsenic testing in pilot villages
November 2012 - February 2013	First round of arsenic testing
February 2013 - May 2013	Follow-up survey of well switching
November 2014 - January 2015	Second round of arsenic testing

Table 2: Summary statistics and randomization balance

	Household members			Asset Index	Asset ownership								
	Adults (1)	Infants (2)	Children (3)	(4)	Pucca (5)	Has Latrine (6)	Cow (7)	Whitegood (8)	Cell (9)	TV (10)	Bike (11)	Motorbike (12)	Car (13)
Panel A: Continuous price													
Price	0.00120	0.00401	0.000543	0.00193	4.29e-05	0.00729***	0.00117	-0.000568	0.00177	-0.000995	-0.00247	0.00197*	-0.000252
	(0.0197)	(0.00298)	(0.00632)	(0.00549)	(0.00285)	(0.00231)	(0.00188)	(0.00224)	(0.00143)	(0.00227)	(0.00199)	(0.00104)	(0.000424)
Panel B: High price													
High Price (>= Rs. 40)	0.273	0.108	0.0481	0.0344	0.00562	0.233***	-0.0126	-0.0267	0.0412	-0.0401	-0.0779	0.0653**	-0.00722
	(0.567)	(0.0975)	(0.169)	(0.202)	(0.0840)	(0.0642)	(0.0491)	(0.0795)	(0.0435)	(0.0732)	(0.0749)	(0.0295)	(0.00999)
Panel C: Price levels													
Price=Rs. 20	0.678	0.0830	0.238	-0.104	-0.227	-0.0667	-0.00546	0.0262	0.0124	0.0308	-0.0277	-0.0515	-0.0110
	(0.673)	(0.0992)	(0.233)	(0.240)	(0.142)	(0.0895)	(0.0903)	(0.109)	(0.0804)	(0.103)	(0.0564)	(0.0563)	(0.0189)
Price=Rs. 30	-0.729	0.0618	-0.134	0.0444	-0.0372	0.0257	0.125	0.0112	0.0532	0.0214	-0.0469	0.00206	-0.0127
	(0.580)	(0.149)	(0.217)	(0.297)	(0.0994)	(0.119)	(0.0819)	(0.116)	(0.0707)	(0.114)	(0.106)	(0.0411)	(0.0175)
Price=Rs. 40	0.268	0.141	0.0633	-0.0582	-0.142	0.166	0.00104	-0.0180	0.0623	-0.00814	-0.137	0.0297	-0.0276**
	(0.696)	(0.114)	(0.254)	(0.334)	(0.0989)	(0.111)	(0.0817)	(0.141)	(0.0612)	(0.137)	(0.135)	(0.0379)	(0.0141)
Price=Rs. 50	0.439	0.176	0.157	0.0818	-0.0304	0.270**	0.0387	-0.00802	0.0576	-0.0392	-0.0583	0.0644	0.000127
	(1.023)	(0.160)	(0.300)	(0.218)	(0.104)	(0.114)	(0.0970)	(0.106)	(0.0668)	(0.0812)	(0.0766)	(0.0461)	(0.0221)
Mean at Price=Rs. 10 (Constant)	3.741	0.242	0.492	0.018	0.795	0.278	0.638	0.209	0.855	0.198	0.722	0.214	0.038
Mean across price groups	3.893	0.322	0.564	5.44e-09	0.700	0.326	0.665	0.215	0.885	0.204	0.676	0.213	0.0286
R-squared	0.040	0.004	0.019	0.004	0.040	0.059	0.011	0.001	0.007	0.003	0.009	0.009	0.003
Observations	3,526	3,528	3,522	3,229	3,758	3,528	3,527	3,528	3,528	3,528	3,528	3,528	3,527
Joint significance													
Wald chi2(df)	4.766	2.592	3.060	0.776	3.929	17.17	4.613	0.130	1.458	0.761	1.446	4.883	9.929
Prob > chi2	0.312	0.628	0.548	0.942	0.416	0.00179	0.329	0.998	0.834	0.944	0.836	0.300	0.0416

Note: the table shows overall mean values of key demographic and asset variables observed in 2015, alongside regression results showing correlation with (Panel A) continuous price variable, (Panel B) high price dummy indicator, and (Panel C) differences in means across price groups. 'Pucca' denotes concrete houses. Asset index is created with house characteristics and asset ownership information using standard principal components approach (Filmer and Pritchett, 2001). A test for joint significance of the price dummies is reported in the bottom rows. Cluster bootstrap standard errors in parentheses (400 replications). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3: External validity: characteristics of sample villages compared to census block, district, and state means

	Housing characteristics		Asset ownership					Household characteristics		
	'Pucca' (1)	Latrine (2)	Cell phone (3)	Bike (4)	Motorbike (5)	Car (6)	TV (7)	Scheduled caste (8)	Literate (9)	Employed (10)
Sample villages	0.659	0.203	0.617	0.592	0.119	0.0174	0.224	0.168	0.611	0.328
Panel A										
Census blocks where villages are situated	0.598	0.258	0.594	0.525	0.113	0.0182	0.224	0.154	0.589	0.298
Difference	-0.061	0.0547*	-0.0226	-0.0671*	-0.00658	0.00082	-0.000193	-0.0138	-0.0218	-0.0304
Panel B										
Bhojpur district	0.627	0.224	0.598	0.509	0.101	0.0184	0.182	0.162	0.583	0.31
Difference	-0.0315	0.0205	-0.0187	-0.0831**	-0.0184	0.000976	-0.0426	-0.00583	-0.0276	-0.0188
Panel C										
Bihar	0.461	0.19	0.517	0.496	0.0773	0.0161	0.128	0.179	0.505	0.343
Difference	-0.197***	-0.0132	-0.0993*	-0.0957**	-0.0421***	-0.00136	-0.0962***	0.0116	-0.106***	0.0141

Note: the table shows mean values of key demographic and asset variables observed in the 2011 Census, for 21 out of 26 sample villages that could be matched with the census, the four census blocks that nest these villages (Panel A), and the district (Panel B) and state (Panel C) where they are all located. 'Pucca' denotes concrete houses. Mean values are shown for each group, alongside the difference between the mean for the respective group and the mean for our sample villages. Significance of differences obtained from robust standard errors (omitted for readability); *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4: Test offers, sales, and demand

	2012 offers and sales		2014 offers and sales			Demand estimates			
Price (Rs.)	Recalled offers	Recalled sales	Sales offers	Sales	Sales among HHs recalling 2012 offer	2012 demand (recall)	2014 demand	2014 demand given 2012 offer	Coverage after two offers
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
10	615	249	960	288	187	0.40	0.30	0.30	0.69
20	804	206	1,105	183	135	0.26	0.17	0.17	0.42
30	460	125	815	117	74	0.27	0.14	0.16	0.38
40	441	92	653	86	72	0.21	0.13	0.16	0.25
50	350	52	551	45	34	0.15	0.08	0.10	0.22
All	2,670	724	4,084	719	502	0.27	0.18	0.19	0.45

Note: the table summarizes the number of offers and sales in both phases of the experiment, alongside the resulting demand levels. Sales reported in Column (5) include repeat purchases, while coverage after two offers in Column (10) has been adjusted by excluding 74 repeat purchases. See Appendix A for additional results and discussion.

Table 5: Estimated demand

	First-round demand (recall)		Second-round demand	
		Wild bootstrap p-value		Wild bootstrap p-value
	(1)	(2)	(3)	(4)
Panel A: Continuous price				
Price	-0.00551*	0.135	-0.00485***	0.030
	(0.00301)		(0.00162)	
Constant	0.418***	0.005	0.307***	0.000
	(0.113)		(0.0588)	
R-squared	0.028		0.029	
Panel B: High price dummy (>= Rs. 40)				
Price >= Rs. 40	-0.127*	0.065	-0.0954**	0.020
	(0.0655)		(0.0403)	
Constant	0.309***	0.000	0.204***	0.000
	(0.0533)		(0.0349)	
R-squared	0.017		0.013	
Panel C: Breakdown by price levels				
Mean at Price = Rs. 10 (Constant)	0.403**	0.110	0.300***	0.000
	(0.163)		(0.0704)	
Price = Rs. 20	-0.146	0.435	-0.134*	0.120
	(0.190)		(0.0738)	
Price = Rs. 30	-0.132	0.485	-0.156*	0.080
	(0.176)		(0.0915)	
Price = Rs. 40	-0.195	0.405	-0.168**	0.050
	(0.169)		(0.0789)	
Price = Rs. 50	-0.255	0.255	-0.218***	0.015
	(0.182)		(0.0727)	
Observations	2,666		4,084	
R-squared	0.034		0.037	
Mean across Price groups	0.271		0.176	

Note: the table shows estimated demand for each individual round of test offers. We use three different specification of prices (Panel A) continuous price variable, (Panel B) high price dummy variable, and (Panel C) price group dummy variables. Demand for 2012 is estimated based on recall data collected in 2014. See Appendix A for an alternative estimate. Cluster bootstrap standard errors (based on 400 replications) in parentheses. Wild bootstrap p-values are provided in Col (2) and Col (4), respectively (Cameron et al., 2008). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 6: Do purchase decisions at high price levels correlate with wealth?

	Test Purchased		
	(1)	(2)	(3)
Asset Index	0.0456*	0.0550***	0.0509***
	(0.0250)	(0.0193)	(0.0195)
Price	-0.0109***		
	(0.00177)		
Price X Asset Index	0.000642		
	(0.000668)		
High Price (>=Rs. 40)		-0.247***	
		(0.0608)	
High Price (>=Rs. 40) X Asset Index		0.0193	
		(0.0270)	
Price= Rs. 20			-0.215**
			(0.0970)
Price= Rs. 30			-0.292***
			(0.0918)
Price= Rs. 40			-0.378***
			(0.0745)
Price= Rs. 50			-0.444***
			(0.0769)
(Price= Rs.20) X Asset Index			0.00204
			(0.0856)
(Price= Rs.30) X Asset Index			0.0105
			(0.0413)
(Price= Rs.40) X Asset Index			0.0392
			(0.0329)
(Price= Rs.50) X Asset Index			0.00658
			(0.0253)
Constant	0.691***	0.473***	0.635***
	(0.0620)	(0.0480)	(0.0527)
Observations	3,229	3,229	3,229
R-squared	0.104	0.067	0.112
Mean at Price = Rs. 10	0.636	0.636	0.636
Mean across all prices	0.402	0.402	0.402

Note: the table tests whether purchase at higher price levels are correlated with household's wealth. Sample includes all the households who participated in round 2 survey. The dependent variable 'Test Purchased' indicates whether a household has purchased the test in either round. Different specifications include continuous price variable, high price dummy variable, and price group dummies, and their interaction with asset index. Cluster bootstrap standard errors (obtained from 400 replications) in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 7: Behavioral response to arsenic test outcome

	Switched to a safe well		Switched to a safe or moderately contaminated well	
	(1)	(2)	(3)	(4)
Test outcome=High arsenic		0.276***		0.305***
		(0.0621)		(0.0624)
Test outcome= High or moderate arsenic	0.242***		0.259***	
	(0.0435)		(0.0425)	
Safe well (Constant)	0.00316*	0.00316	0.00316*	0.00316*
	(0.00186)	(0.00195)	(0.00186)	(0.00185)
Observations	1,037	844	1,037	844
R-squared	0.158	0.214	0.171	0.239

Note: the table shows the probability that households whose wells had unsafe arsenic levels ('red') switched to safer wells. Arsenic test results from 2012 data; self-reported switching data from 2013 follow-up survey. Column (1) considers switching only to wells with safe ('blue') levels of arsenic; Column (2) and (3) considers switching to safe or moderately contaminated ('green') wells. Cluster bootstrap standard errors (obtained from 400 replications) in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 8: Effect of price paid on behavioral response to information

	Switched from high arsenic well to safe well		Switched from high arsenic well to safe or moderately contaminated well	
	(1)	(2)	(3)	(4)
Mean across price groups	0.280		0.308	
Panel A: Linear Specification				
Price	0.000425		0.00110	
	(0.00347)		(0.00362)	
Constant	0.267**		0.276**	
	(0.116)		(0.118)	
R-squared	0.001		0.001	
Panel B: High price dummy (price>= Rs. 40)				
Price >= Rs. 40	0.0191		0.0260	
	(0.130)		(0.133)	
Constant	0.271***		0.297***	
	(0.0866)		(0.0826)	
R-squared	0.001		0.001	
Panel C: Breakdown by price levels				
Price = Rs. 20		0.242		0.227
		(0.277)		(0.277)
Price = Rs. 30		-0.0326		0.00227
		(0.225)		(0.215)
Price = Rs. 40		0.0254		0.0292
		(0.212)		(0.226)
Price = Rs. 50		0.0424		0.0773
		(0.132)		(0.116)
Constant (mean at Price = Rs. 10)		0.258***		0.273***
		(0.0971)		(0.0971)
R-squared		0.018		0.014
Observations	211	211	211	211
Joint significance				
Wald Chi2		0.096		1.13
Prob > Chi2		0.916		0.889

Note: the table shows the correlation between behavioral response i.e. switching and price paid for arsenic testing. Panel A and Panel B include continuous price variable and high price dummy variable, respectively. Panel C shows regression coefficient for price group level dummy variables. Arsenic test results from 2012-13 data (round 1); self-reported switching data from 2013 follow-up survey. Column (1) and (2) consider switching only to wells with safe ('blue') levels of arsenic; Column (3) and (4) consider switching to safe or moderately contaminated ('green') wells. Cluster bootstrap standard errors (obtained from 400 replications) in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 9: Selective recall of arsenic test outcomes

	Round 2 Sample: Fixed on well			Round 2 Sample: Kept in house			Round 2 Sample: Recall of placard color			Round 2 Sample: All three combined		
Placard color	Red (1)	Green (2)	Blue (3)	Red (4)	Green (5)	Blue (6)	Red (7)	Green (8)	Blue (9)	Red (10)	Green (11)	Blue (12)
Difference in proportion (between round 1 and round 2)	-0.0942***	0.0584	0.0358	-0.0925**	0.155***	-0.0621	-0.116***	0.0555	0.0601	-0.0955***	0.118***	-0.0221
	(0.0239)	(0.0355)	(0.0402)	(0.0400)	(0.0504)	(0.0738)	(0.0285)	(0.0357)	(0.0466)	(0.0252)	(0.0323)	(0.0389)
Actual proportion	0.21	0.18	0.61	0.21	0.18	0.61	0.21	0.18	0.61	0.21	0.18	0.61
Recorded proportion	0.12	0.24	0.64	0.12	0.34	0.54	0.09	0.24	0.67	0.11	0.30	0.58
Observations	1,529	1,529	1,529	1,379	1,379	1,379	1,762	1,762	1,762	1,840	1,840	1,840
R-squared	0.010	0.004	0.001	0.006	0.016	0.002	0.020	0.004	0.003	0.014	0.018	0.000

Note: the table compares the proportion of 'red' (unsafe), 'green' (moderately contaminated) and 'blue' (safe) wells in the recorded results of tests conducted in 2012 (as measured), and in household recall or retained placards obtained in the 2014 survey. Top row headings denote subsamples from round 2 survey– (1) "fixed on well" - the placard still fixed on the well (Columns 1-3),(2) "kept in house"- removed from the well but still kept in the house (Columns 4-6), and (3) "recall of placard color"- the proportion of red, green and blue recall (Columns 7-9), respectively. Columns (10-12) pool information on well status from all test outcome recall and retained placards in round 2. The coefficient on 'Difference in proportion' reflects the difference in shares of each test result category in round 2, when we compare corresponding subsamples from round 2 with actual measurements in round 1. We estimate the 'Difference in proportion' by regressing test outcome dummy indicator (pooled from round 1 and round 2) on round 2 dummy, for each sub-sample category. "Actual proportion" displayed are the actual measured test outcomes in round 1 (i.e. constant in the regression). "Recorded proportion" indicate observed or recalled test outcome during round 2 within each subsample. The sample size in respective columns reflects the sum of all tests recorded in 2012, along with the number of households for which information in a given category was available in 2014. Cluster bootstrap standard errors (obtained from 400 replications) in parentheses. *** $p<0.01$, ** $p<0.05$, * $p < 0.1$.

Table 10: Selective recall and household assets

	Placard color red			
	Sample:Fixed On well (1)	Sample:Kept in house (2)	Sample:Recalled (3)	Sample:All (4)
Second phase	-0.0831***	-0.0688	-0.0919***	-0.0760***
	(0.0285)	(0.0507)	(0.0286)	(0.0256)
HH owns consumer durables	0.0423	0.0423	0.0423	0.0423
	(0.0402)	(0.0405)	(0.0406)	(0.0397)
Second phase * HH owns consumer durables	-0.0571	-0.0661	-0.0903**	-0.0728*
	(0.0495)	(0.0662)	(0.0409)	(0.0407)
Observations	1,497	1,350	1,730	1,808
R-squared	0.012	0.007	0.023	0.016

Note: the table shows differences in the share of 'red' wells in 2012 tests and 2014 recall as in Table C, but conditional on ownership of (any) consumer durables. The coefficient on 'HH owns consumer durables' is the same across all four samples by construction: it is only the composition of the 2014 recall sample that changes, not the composition of the 2012 test sample. Cluster bootstrap standard errors (obtained from 400 replications) in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

A Comparison of 2012 demand estimates based on recorded and recall sales data

As noted in the main body of the paper, during the first offer phase in 2012, enumerators did not systematically collect data from all households - chiefly, some households that did not want to purchase the test were omitted. (This is evident in the comparison of Columns 2-4 in Table A1.) In addition, anecdotal evidence raises a concern that enumerators may have offered tests less systematically in parts of the villages where people showed strong reservations against the idea of arsenic tests being offered for a fee (rather than free of charge) during focus group meetings.

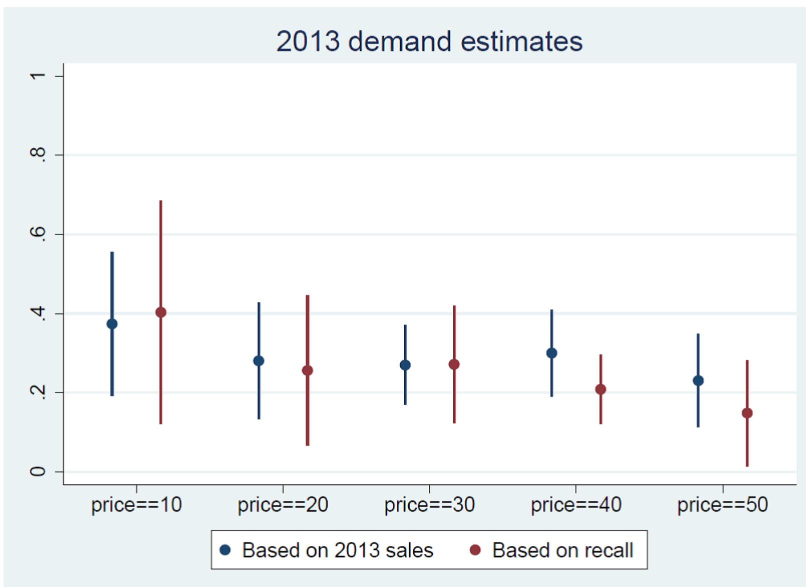
We hence face a considerable challenge in reliably assessing baseline demand, since the number of households to whom the test was offered in 2012 cannot be completely ascertained. We address this challenge with the following strategy. (1) We first compute demand based on recall data collected in the 2014 follow-up survey (i) on whether households were offered the test at baseline, and (ii) on whether they purchased the test at baseline. (Table A1, Columns 5-6.) This estimate is correct to the degree that there is no correlation between the decision to purchase in 2012 and recalling the offer when surveyed in 2014.

To assess whether the recall-based estimate is reasonable, we also (2) estimate demand from the 2012 sales (Column 3), based on the assumption that as many households were approached during the 2012 campaign as during the 2014 campaign (Column 4). This estimate is correct to the degree that (i) sales approaches were comprehensive in 2012 (while numerators neglected to keep records of some visits), and (ii) the number of households has remained constant between survey rounds.

Reassuringly, as is evident from Table A1 and Figure A1, the estimates obtained by recall and by imputing the number of sales offers are well-aligned in the aggregate (27% and 30%, respectively) and in the Rs. 10-30 groups. They diverge more at higher prices, though never significantly so. As a corollary, there is a good match between the ratio of *recalled* 2012 sales to *recorded* 2012 sales (0.65) on the one hand, and the ratio between *recalled* 2012 offers and *recorded* 2014 sample size on the other (0.60). This suggests that recall error is similarly likely for offers and sales, and provides at least some reassurance that the 2012 data is affected by failure to record unsuccessful sales attempts, rather than selective sales attempts.

Although first-round data collection did not follow protocol completely, we are hence able to offer two sensible demand estimates, and show that they match up well with each other. In the main body of the paper, we discuss results based on recall data – arguably, the more internally consistent approach, as well as the more conservative demand estimate. It would be a potential concern, if our demand estimates are biased by a differing impact of adverse test outcomes on test purchase recalls. However, note that we find little correlation between offered price and high arsenic outcomes (Column 2, Table C2). Moreover, a lower recall of high arsenic well affects only a small share of total number of wells, and is also almost fully compensated by a higher recall of moderate arsenic wells (Table 9).

Figure A1: Comparison of demand estimate from first phase data and recall



Note: the plot shows demand estimates obtained by scaling recorded sales in the first round of offers (2012) to 2014 sample size, and from offers and sales recalled in 2014.

Table A1: Test offers, sales, and demand

	Recorded 2012 offers and sales			Recalled 2012 offers and sales		Demand estimates	
Price (Rs.) (1)	Recorded offers (2)	Recorded sales (3)	*Sample 2014* (4)	Recalled offers (5)	Recalled Sales (6)	2012 demand (recorded sales) (7)	2012 demand (recall data) (8)
10	431	361	*960*	615	249	0.38	0.40
20	423	310	*1105*	804	206	0.28	0.26
30	352	218	*815*	460	125	0.27	0.27
40	327	196	*653*	441	92	0.30	0.21
50	289	127	*551*	350	52	0.23	0.15
All	1822	1212	*4084*	2670	724	0.30	0.27

Note: the table summarizes data used in computing the 2012 demand estimates shown in Figure A1.

B Why is there substantial demand at the time of the repeat offer?

We find that repeating the offer after a two-year delay generate substantial additional demand and raise total coverage by some 18 percentage points (pp), from 27% to 45% (Table 4, Columns 7 and 10). Demand is more price-sensitive than at the first offer (Figure 5). However, we observe an effect of repeating the sales offer on coverage at any price level, with increases ranging from 70% of the original sales at Rs. 10 to 19% at Rs. 40. To study the (reduced form) effect of making a repeat offer, we keep price constant within a village. This, in turn, limits our ability to directly test for learning as a specific mechanism driving demand at the time of the second offer. The reason why we cannot assess learning as in Dupas (2014b) is as follows. Our product is distinct from the ITNs offered in Dupas (2014b) in that there is no reason for households to repeat arsenic tests, whereas there is reason to purchase ITNs again after some time. Still, if we had made the second sales offer at a uniform price, we might have tested for learning by using first-round price to instrument for first-round demand, and then study the effect of first-round demand on second-round demand through peer learning. This is not possible, however, when price levels are the same in the first and second round: as an instrument, price would clearly violate the exclusion restriction.

From a policy perspective, the effect of making a repeat offer is remarkable: price matters greatly for demand, but at any price level considered here, repeating the offer meaningfully increases coverage (and from a business perspective, sales). Irrespective of the channels – learning, income growth, or marketing intensity, this simple finding underscores the need for a more careful assessment of experimental evidence generated with offers available only for a short period. Because we lack a household panel, and because there may be some error in recall of first-round tests, we cannot completely rule out the concern that some of the demand at the second offer may be driven by households that may not have been approached during the first offer phase in 2012. However, the observable evidence offers significant reassurance. About 70% of the new purchases in 2014 are made by households who recall being offered the test in 2012, but did not purchase (Table 4, Columns 5-6). Perhaps most compellingly, the pattern of 2014 demand is very similar among those who recall having been made an earlier offer and the overall sample (Column 10).

It is intriguing to ask why there is a high level of demand when a repeat offer is made within the relatively short time frame of two years. However, our data does not allow us to conclusively assess this question; we present some suggestive evidence in this Appendix. (i) Strong state-level growth in nominal income between survey rounds suggests that changes in wealth between the first and second offer may have played a role; our survey data on asset ownership is consistent with this mechanism, but not conclusive. The absence of a correlation between wealth and price among buyers is at odds with this explanation (see Section 3.1.1). (ii) Learning may have lead households to adjust their valuation of arsenic testing. The product's characteristics were not familiar to potential customers at the time of the first offer, and the initial wave of tests may have allowed households to change their beliefs about the possibility of contamination, and opportunities to switch, although the health benefits of switching cannot be observed within two years. We obtain the 'expected' sign in a test with a credibly causal interpretation, but the results are not significant (i.e. a positive but insignificant effect of 'arsenic unsafe' outcome in the first phase on the demand for arsenic testing during the second phase). (iii) In the absence of conclusive evidence on wealth or learning effects, one could speculate about a direct effect of repeating the offer – what one might call a 'marketing' or 'nudge' effect. We consider it a priority for further work to assess the importance of such an effect. This appendix summarizes evidence on what might explain demand at the time of the repeat offer. On balance, the evidence is inconclusive. Patterns in wealth proxies are consistent with a contribution of growing income and wealth. We note, however, that this is at odds with the absence of a correlation of wealth proxies with sales price among buyers shown above. A test for learning that allows for a sound causal interpretation is consistent in sign, but not significant.

B.1 Wealth effects

There is mixed evidence on increased wealth as a driver of repeat offer demand. As reported above, we find that observable wealth does not correlate systematically with willingness to pay. Indeed, one of the two wealth proxies that does correlate – ownership of a latrine – can be read as a marker of difference in concern over health that might affect valuation of the arsenic test as much as it may speak to lower marginal utility of consumption.

Still, there are some good reasons to ask whether rising wealth may have to some degree contributed to generating additional demand.

The most important piece of *prima facie* evidence is the rapid economic growth Bihar experienced between sales rounds. Per capita real income rose precipitously, at a rate of about 10% per year between 2012 and 2014.[20] In line with such a favorable development, ownership of consumer durables among households who purchased tests in the first round of offers (the one asset category we can reliably compare among both survey rounds, and the one group of consumers sampled in a consistent way) rose by 5pp from a baseline value of 23% between 2012 and 2014 (result not shown). Because the tests were offered at the same *nominal* price in both phases, inflation further reinforced this effect. In total, nominal per capita income grew by some 38% between the two offers.

Secondly, patterns in asset ownership among buyers groups and across time are consistent with a wealth effects – though they do not offer a very powerful test. Our data allows in principle for two tests to reject wealth effects (at the mean). Most obviously, we can compare wealth among the two groups of buyers *at the time of purchase*, that is, in 2012 and 2014, respectively. This comparison could furnish some evidence against wealth effects if it were to emerge that second-round buyers were less well-off at the time of purchase than first-round buyers were at the time their wells were tested (with the assumption that the two groups initially had the same valuation of the tests). We can only draw this comparison on the ownership of (any) consumer durables; questions used to collect ownership information for all other categories of assets differed too much between the 2012 and 2014 surveys. For consumer durables, there is no significant difference between buyer groups, and the coefficient is centered near zero (Panel A in Table B1). This finding is consistent with wealth effects (new buyers catching up in wealth to original buyers), but also does not exclude a contribution of learning.

Beyond the ownership of consumer durables, we are constrained to comparing wealth as observed in the year 2014: among households that bought in 2012 and households that bought in 2014. This comparison could also reject wealth effects, namely if second-round buyers were weakly better off in 2014 than first-round buyers (and we were willing to assume that growth in wealth among the two groups was such that the ranking was not reversed since 2012 – which would then imply, less appealingly, that the wealthier group initially had a lower valuation of the tests). Our data suggests quite clearly that the opposite was the case: first-round buyers were better off than second-round buyers when surveyed in 2014 (Table B1). Difference in ownership of durables such as TV and consumer durables are significant, second round buyers have significantly less education than first

[20] State GDP growth for India from `http://planningcommission.nic.in/data/datatable/data_2312/`

round buyers, and there are notable differences in caste composition.[21]

B.2 Learning

Arsenic tests in themselves are distinctly a non-experience good: a one-off application which does not directly affect the consumer. It is therefore most plausible to suggest that learning might be chiefly driven by increased awareness of the probability of arsenic contamination, and of opportunities to switch to safe wells.

We test in the following way for evidence of learning after the first wave of tests. Because the distribution of arsenic in ground water varies substantially and unpredictably over small distances, variation in the results of first-round tests is exogenous. We posit that different distributions of first-round results at the village level may induce differential effects on second-round demand. In particular, we speculate that, when a high share of wells tested 'unsafe' during the first wave, concern in the village community over arsenic contamination might have been raised, translating into learning – namely, greater awareness of the health risks associated with arsenic, and the benefits of testing and well-switching. Empirically, the relationship between second-phase purchases and the share of wells tested 'unsafe' in the first phase has the expected sign, across a range of specifications (Table B2). However, results are not significant with cluster bootstrap standard errors. Furthermore, we have considerably low statistical power to detect any learning effect in Table B2 because there are only 26 villages in our sample.

[21] We note that, strictly speaking, we are comparing between one group observed pre-treatment (2014 buyers) and one group observed post-treatment (2012 buyers). However, since the health effects of Arsenic are long-term, one would not expect a strong treatment effect a mere two years after the test, even conditional on households effectively avoiding exposure. We acknowledge that in principle, Arsenic testing could have had effects upon wealth through conduits other than health – for instance, a change in the value of houses with wells tested safe/unsafe, or a change in social status with implications for future wealth.

Table B1: Household characteristics of first and second phase buyers

	Panel A: as observed at time of purchase		
	2014 buyers	2012 buyers	2014 vs. 2012
	(1)	(2)	(1) - (2)
HH has consumer durables	0.225	0.226	-0.00135
	(0.0404)	(0.0276)	(0.0392)

	Panel B: as observed in 2014		
	2014 buyers	2012 recall	2014 vs. 2012 recall
	(1)	(2)	(1) - (2)
Household characteristics			
Number of HH members	4.919	4.311	0.608
	(0.367)	(0.325)	(0.382)
Infant living in HH	0.302	0.223	0.0798**
	(0.0459)	(0.0246)	(0.0370)
Child living in HH	0.488	0.438	0.0497
	(0.0585)	(0.0618)	(0.0657)
Housing characteristics			
House pucca	0.701	0.756	-0.0553
	(0.0556)	(0.0504)	(0.0391)
Has latrine	0.330	0.408	-0.0778
	(0.0551)	(0.0496)	(0.0553)
Asset ownership			
HH has consumer durables	0.225	0.301	-0.0766*
	(0.0404)	(0.0563)	(0.0405)
Has cell phone	0.912	0.861	0.0507
	(0.0230)	(0.0578)	(0.0460)
Has TV	0.208	0.298	-0.0905**
	(0.0372)	(0.0573)	(0.0424)
Has bicycle	0.783	0.811	-0.0285
	(0.0187)	(0.0402)	(0.0382)
Has motorbike	0.248	0.261	-0.0131
	(0.0254)	(0.0243)	(0.0260)
Has cow	0.680	0.680	6.24e-05
	(0.0417)	(0.0319)	(0.0353)
Caste			
Scheduled caste or tribe	0.0163	0.0386	-0.0223
	(0.00852)	(0.0240)	(0.0226)
Other backward caste	0.227	0.127	0.0995**
	(0.0518)	(0.0298)	(0.0411)
Kshatriya	0.0767	0.124	-0.0473
	(0.0309)	(0.0455)	(0.0371)
Brahmin	0.251	0.388	-0.137***
	(0.0658)	(0.0646)	(0.0510)
Baniya	0.297	0.203	0.0940*
	(0.0670)	(0.0446)	(0.0537)

Note: the table shows characteristics of households that bought tests in 2014 (Column 1) and 2012 (Column 2), and the change between the two phases (Column 3). Panel A shows ownership data as observed at the time of purchase; Panel B shows data as observed in 2014 – that is, 2014 values for those who buy in 2014 in Column (1), and 2014 values for those who recall having purchased in 2012 in Column (2). Cluster bootstrap standard errors (obtained from 400 replications) in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table B2: Do first-round test results relate to second-round demand?

	Demand in Second Phase				
	(1)	(2)	(3)	(4)	(5)
Share of wells in village tested	0.0384	0.0699	0.0437	0.0933	0.117
arsenic high (red) in first round	(0.112)	(0.125)	(0.107)	(0.114)	(0.130)
	[0.0301]	[0.0384]	[0.0301]	[0.0326]	[0.0404]
Controls					
Price	Yes	Yes	Yes	Yes	Yes
First-round demand	No	No	Linear	Quadratic	Quadratic
Wealth proxies	No	Yes	No	No	Yes
N	4,084	3,002	4,084	4,084	3,002
R-squared	0.037	0.060	0.051	0.059	0.082

Note: the table summarizes the correlation between arsenic test outcomes in the first phase and the demand in second phase. In each column, the dependent variable is demand for well tests in the second phase of offers, and the coefficient of interest is the share of wells that tested 'red' (high arsenic) among wells tested in the first offer phase. All models include price controls; Columns 3-5 control for first-round demand, and Column 5 controls for wealth proxies. We consider Column 4 to show the preferred specification. Cluster bootstrap standard errors (400 replications) in parentheses, classical standard errors in square brackets. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C1: Do purchase decisions at high price levels correlate with wealth?

		House characteristics		Asset ownership							
	Asset Index	Pucca	Latrine	Cow	Whitegoods	Cell	TV	Bike	Motorbike	Car
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Linear Specification										
Price	0.00719	-0.00198	0.00726**	0.000165	0.00165	0.00141	0.00163	-0.000502	0.00290***	0.000138
	(0.00742)	(0.00309)	(0.00310)	(0.00216)	(0.00334)	(0.00167)	(0.00315)	(0.00203)	(0.000926)	(0.000390)
Panel B: High price dummy (price>= Rs. 30)										
Price >= Rs. 30	0.208	0.00586	0.165*	0.0481	0.0350	0.0502	0.0362	0.0118	0.0591*	0.00521
	(0.222)	(0.0858)	(0.0878)	(0.0536)	(0.0963)	(0.0565)	(0.0945)	(0.0613)	(0.0352)	(0.0106)
Panel C: High price dummy (price>= Rs. 40)	0.233	-0.0392	0.291***	-0.0323	0.0471	0.0700	0.0378	-0.0469	0.0947***	-0.00150
Price >= Rs. 40	(0.251)	(0.0860)	(0.0696)	(0.0562)	(0.115)	(0.0515)	(0.115)	(0.0890)	(0.0293)	(0.0123)
Panel D: Breakdown by price levels										
Price = Rs. 20	-0.0562	-0.196	-0.0418	-0.0573	0.0348	-0.0326	0.0475	-0.0281	0.0321	-0.00344
	(0.270)	(0.127)	(0.106)	(0.0867)	(0.120)	(0.113)	(0.125)	(0.0732)	(0.0717)	(0.0154)
Price = Rs. 30	0.130	-0.0409	0.0124	0.0841	0.0335	0.0110	0.0488	0.0479	0.0323	0.00858
	(0.336)	(0.123)	(0.137)	(0.0792)	(0.143)	(0.0583)	(0.145)	(0.0688)	(0.0437)	(0.0179)
Price = Rs. 40	0.241	-0.188	0.254**	-0.0481	0.0841	0.106***	0.0921	-0.0665	0.113***	-0.0123
	(0.397)	(0.121)	(0.114)	(0.104)	(0.182)	(0.0353)	(0.187)	(0.160)	(0.0384)	(0.0138)
Price = Rs. 50	0.238	-0.00340	0.316***	-0.0156	0.0384	-0.00421	0.0218	-0.0180	0.110***	0.0157
	(0.237)	(0.0909)	(0.119)	(0.112)	(0.147)	(0.0680)	(0.123)	(0.0898)	(0.0370)	(0.0217)
Mean at Price = Rs. 10 (Constant)	0.100	0.808	0.333	0.686	0.228	0.887	0.212	0.784	0.220	0.0268
Observations	1,297	1,301	1,366	1,365	1,366	1,366	1,366	1,366	1,366	1,365

Note: the table shows correlations between purchase price and wealth proxies among households that bought a test during either rounds. Dependent variables are asset index and wealth proxies, as mentioned in the header of each column. Each panel shows coefficients of interest from different specifications. Panel A shows results from a linear regression in continuous price variable; Panel B and C shows results from a regression on high price indicator variables; Panel D shows results from a regression on price indicators. Cluster bootstrap standard errors (obtained from 400 replications) in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C2: Sorting on well status

	Dependent variable: High arsenic well				
	Well characteristics	Price		Asset Index	Wealth Proxies
	(1)	(2)	(3)	(4)	(5)
Well age	-0.00234				
	(0.00323)				
Well depth	0.00114				
	(0.00127)				
Well cost	1.48e-06				
	(9.79e-06)				
Price		0.0051			
		(0.00386)			
High price (>= Rs. 40)			0.1012		
			(0.0908)		
Asset Index				0.0212	
				(0.0309)	
					Coefficients from univariate regressions
Car					0.172
					(0.140)
Cell					-0.0148
					(0.0881)
Several Cells					-0.0558
					(0.0800)
TV					-0.00610
					(0.0615)
Bike					0.0626*
					(0.0325)
Motorbike					-0.0285
					(0.0413)
Cow					0.102**
					(0.0438)
Several Cows					0.0529
					(0.0514)
Whitegoods					0.0377
					(0.0679)
Pucca					-0.0255
					(0.0609)
Latrice					0.0981
					(0.0689)
Number of HH members					-0.00480
					(0.00936)
Infants					0.0125
					(0.0212)
Children					-0.00866
					(0.0219)
Observations	677	719	719	676	719
R-squared	0.007	0.022	0.008	0.002	n/a

Note: the table shows correlations among wells tested in 2014, between the probability of a well having high arsenic status (at least 50μg/l) with characteristics of the well (Column 1), price (Column 2 and 3), asset index (Column 4) and the household asset ownership (Column 5). To avoid evident overfitting problems, regression coefficients show in Column 5 were obtained by performing univariate regressions for each characteristic. Cluster bootstrap standard errors (obtained from 400 replications) in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1.

Table D1: Decision to re-test depends on contamination status

	Well contamination status		
	Red	Green	Blue
	(1)	(2)	(3)
Test purchased in both 2012 and 2014	-0.0411	0.172***	-0.130*
	(0.0582)	(0.0598)	(0.0792)
Share among wells tested once only	*0.257*	*0.274*	*0.468*
Observations	719	719	719
R-squared	0.001	0.013	0.006

Note: the table compares the proportion of 'red' (unsafe), 'green' (moderately contaminated) and 'blue' (safe) wells in the recorded results of tests conducted in 2014, among households that recalled preciously purchasing a test, and households that recalled a prior offer, but no purchase. Arsenic levels are stable over time, so test results obtained in 2012 can be assumed to have been identical to those measured in 2014. Cluster bootstrap standard errors (obtained from 400 replications) in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

The International Growth Centre (IGC) aims to promote sustainable growth in developing countries by providing demand-led policy advice based on frontier research.

Find out more about
our work on our website
www.theigc.org

For media or communications
enquiries, please contact
mail@theigc.org

Subscribe to our newsletter
and topic updates
www.theigc.org/newsletter

Follow us on Twitter
@the_igc

Contact us
International Growth Centre,
London School of Economic
and Political Science,
Houghton Street,
London WC2A 2AE

IGC
International Growth Centre

DIRECTED BY
LSE
UNIVERSITY OF OXFORD

FUNDED BY
UKaid from the British people

Designed by soapbox.co.uk