The World Bank Economic Review, 35(4), 2021, 969–998
https://doi.org10.1093/wber/lhab006
Article

OXFORD

Do Community Water Sources Provide Safe Drinking Water? Evidence from a Randomized Experiment in Rural Bangladesh

Serena Cocciolo, Selene Ghisolfi, Md. Ahasan Habib, S.M.A. Rashid, and Anna Tompsett

Abstract

Health, and in turn income and welfare, depend on access to safe drinking water. Although the majority of rural households worldwide obtain drinking water from community water sources, there is limited evidence about how effectively these sources provide safe drinking water. This study combines a randomized experiment with water quality testing to evaluate the impact of a program that provides community deep tubewells in rural Bangladesh. The program reduces exposure to arsenic, a major natural pollutant, but not fecal contamination. Households may use fewer sources with fecal contamination, but any such effects are offset by recontamination through transport and possibly storage. The results suggest that while community deep-tubewell construction programs may reduce exposure to arsenic in Bangladesh, reducing exposure to fecal contamination may require interventions that go beyond community sources.

JEL classification: H41, I14, O18, Q53, Q56

Keywords: safe drinking water, deep tubewells, Bangladesh, arsenic

1. Introduction

Billions of people across the developing world still lack access to safe drinking water, a major determinant of health,[1] and in turn, of welfare and income (Bloom and Canning 2008;

Serena Cocciolo is a Young Professional at the World Bank, Washington, DC, USA. Her email address is scocciolo@worldbank.org. Selene Ghisolfi is a postdoctoral researcher at the Laboratory for Effective Anti-poverty Policies, Bocconi University, Italy. Her email address is ghisolfi@unibocconi.it. Md. Ahasan Habib is Head of Research, Monitoring, and Evaluation at the NGO Forum for Public Health, Dhaka, Bangladesh and a doctoral student at KTH Royal Institute of Technology, Stockholm, Sweden. His email address is habi@kth.se. S. M. A. Rashid is Executive Director of the NGO Forum for Public Health, Dhaka, Bangladesh. His email address is ngof@bangla.net.bd. Anna Tompsett (corresponding author) is an Assistant Professor at Stockholm University, Stockholm, Sweden. Her email address is anna.tompsett@ne.su.se.
This work was supported by the International Initiative for Impact Evaluation (3ie) [DPW1/1006], the Swedish Research Council Development Research [348-2014-3841], and the International Growth Center (the IGC) [31402]. We also acknowledge technical review and support throughout the study by 3ie. AEA pregistration ID: AEARCTR-0002755 (https://www.socialscienceregistry.org/trials/2755). We thank Prosun Bhattacharya, Ram Fishman, Björn Tyrefors Hinnerich, Andreas Madestam, Alexander van Geen, the editor, and three anonymous referees for helpful comments and suggestions. We thank Sabbir Ahmed, Cristina Altomare, Tillmann von Carnap, Merve Demirel, Md. Shahadat Hossain, Jahirul Islam, Md. Tariqul Islam, Md. Rezwanul Karim, Rezaul Karim, Ranjan Kumar Mohonto, Mir Abu Raihan, Lorenzo Schirato, and Tahmid Sharif for outstanding research assistance and support. A supplementary online appendix for this article can be found at *The World Bank Economic Review* website.

1 See Cutler and Miller (2005) as an example from a literature too extensive to review here.

© The Author(s) 2021. Published by Oxford University Press on behalf of the International Bank for Reconstruction and Development / THE WORLD BANK.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

Pitt, Rosenzweig, and Hassan 2020). Learning which interventions are effective in improving access to safe drinking water can be difficult, because the places in which interventions are successfully implemented are rarely comparable to the places in which interventions are not successfully implemented. In particular, although more than half of rural households worldwide depend on community water sources (UNICEF and WHO 2011), little is known with certainty about how successfully these sources provide safe drinking water. This study provides experimental evidence on the impact of new community water source construction on drinking water quality. Experimental assignment ensures that communities that do and do not receive the community water source program are statistically identical with respect to all other characteristics that might affect drinking water quality. Any differences that emerge after new water source construction can thus be attributed to the causal effects of the intervention.[2]

The context for this study is rural Bangladesh. In 2015, when this study began, only 52 percent of Bangladesh's 165 million citizens were using sources of drinking water that were free from both fecal contamination and arsenic, a natural pollutant that is common in shallow groundwater in Bangladesh (BBS and UNICEF 2014). In many parts of rural Bangladesh, the primary technology used to improve access to safe drinking water is the deep tubewell. Deep tubewells draw groundwater from deep aquifers that are isolated from sources of fecal and arsenic contamination. Deep tubewells are expensive and must typically be provided at the community level. However, how effectively deep-tubewell construction programs improve access to safe drinking water is in practice unclear. Take-up rates are uncertain, leading to a range of estimates for cost-effectiveness that vary by a factor of 16 (Jamil et al. 2019). Additionally, providing deep tubewells may not solve the problem of fecal contamination in drinking water if contamination also occurs during transport and storage (Wright, Gundry, and Conroy 2004).

This study evaluates the impact of a community deep-tubewell construction program on arsenic and fecal contamination in household drinking water. The deep-tubewell construction program consists of a package of subsidies and technical advice. A partner nongovernmental organization offered the program in 129 communities of between 50 and 250 households, successfully installing 107 new deep tubewells. The new deep tubewells provide communities with arsenic-safe drinking water, essentially eliminating arsenic contamination at source for those who collect all their drinking and cooking water from the new wells. The deep tubewells also have lower rates of fecal contamination than other wells in the same communities, although a third of the new deep tubewells still test positive for fecal contamination.[3] The most likely mechanism for fecal contamination is accumulation of microbial organisms in the tubewell body during collective use (Ferguson et al. 2011; ICDDRB and UNICEF 2018).

Well take-up is modest, relative to expected take-up based on previous literature (van Geen et al. 2003; Jamil et al. 2019). In communities that successfully installed new wells, the median household is 1.6 minutes walk (around 130 m) from a new well. Among households whose drinking water is contaminated with arsenic at baseline, those that live closest to the well collect, on average, around 40 percent of their drinking and cooking water from the source. Take-up declines by approximately 10 percentage points per minute of walking distance. Among households who do not adopt the well despite having household drinking water that is contaminated with arsenic at baseline, the majority say that women do not feel comfortable using a water source outside their compound and that the wells are too far away, even when the wells are less than a one-minute walk away.

To evaluate the program, enumerators tested water quality for 6,051 study households, constituting a representative sample of households in the treated communities and 42 control communities that did not receive the intervention. Enumerators tested water quality in household drinking water and in all sources used by each household, both before and after the intervention. All samples were tested for contamination

2 See, e.g., Banerjee and Duflo (2009).

3 Among project wells, 34 percent test positive for fecal contamination, compared to 46 percent of other tubewells in the same communities. Other studies also report bacterial contamination in deep tubewells (e.g., Goel et al. 2019). See also Howard et al. (2006).

Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

both with arsenic and with fecal bacteria. At baseline, enumerators also conducted a water quality census in all water sources in each community.

This paper evaluates the average "intent-to-treat" effects of the program using a difference-in-difference approach that compares changes in treated communities to changes in control communities. The paper also reports "scaled" effects that, under additional assumptions, measure the effect of each well successfully installed under the program or, equivalently, the "treatment effect on the treated." Since enumerators informed households and/or water source caretakers of the results of all water quality tests, the evaluation measures the impact of the deep-tubewell construction program combined with information about water source quality, relative to only receiving information about water source quality.

Households in treated communities use sources with lower rates of arsenic contamination. On average, each deep tubewell installed leads 10 households to switch from sources with arsenic contamination at the WHO threshold to sources free from arsenic contamination (90 percent confidence interval: −15 to −4.2 households) and eliminates arsenic contamination in household drinking water for around 5 households, although the confidence interval does not exclude zero (90 percent CI: −9.5 to +0.4). The smaller estimated effects on arsenic in household drinking water may be the consequence of additional passive removal or sedimentation of arsenic during storage of household drinking water.

Households in treated communities use sources with only slightly lower rates of fecal contamination. Each tubewell installed leads about 5 households to switch from sources with fecal contamination to sources free from fecal contamination (90 percent CI: −9.5 to −0.4). Most of the difference in impact on source water quality between the two pollutants is explained by the comparatively high levels of fecal contamination in project tubewells. Even these relatively modest improvements in source water quality are offset by increases in travel time and possibly by changes in storage behaviour. Our best estimates suggest that walking an extra minute to collect drinking water increases the risk of fecal contamination by around 1.8 percentage points (90 percent CI: 0.3 to 3.2), while storing drinking water in the house before consumption increases the risk of fecal contamination by around 7 percentage points (90 percent CI: 5.4 to 8.6). The net estimated effect of each tubewell installed is to introduce fecal contamination into drinking water for around 2 households, although the confidence intervals include small reductions as well as small increases in fecal contamination (90 percent CI: −2.5 to +6.0).

Few previous studies provide causal evidence on the effects of any new water infrastructure on drinking water quality.[4] The closest precedent is Kremer et al. (2011) who evaluate the impact of improvements to community water source quality, through spring protection, on household water quality. Kremer et al. (2011) find that entirely removing source-level microbial contamination reduces household microbial contamination by 24 percent, but cannot causally estimate what explains the remaining contamination. More generally, it is difficult to empirically disentangle the effects of source water contamination, transport, and storage on household drinking water contamination with fecal pathogens.[5] Households who live near safe water sources likely differ from those who live further away, in ways that might also affect household behaviour with respect to water safety and hygiene, such as income or education. Such differences could exaggerate or obscure the relationship between source water contamination, transport, and storage and household drinking water quality in cross-sectional data. The intervention evaluated in this study creates entirely new sources, thus changing both source water quality and travel times, allowing this study to make progress on understanding the determinants of household water quality.

4 Devoto et al. (2012) use a randomized incentive scheme to measure the effect of private household water connection on drinking water quality. A few studies provide causal evidence of the impact of new water infrastructure on other outcomes, but lack direct measurements of water quality (e.g., Bennett 2012; Madajewicz, Tompsett, and Habib 2021). Other studies describe the correlations between infrastructure access and water quality, but cannot separate out the causal effects of water infrastructure from other factors that also affect water quality and may be correlated with the presence of water infrastructure (e.g., Genthe et al. 1997; Hoque et al. 2006; Shields et al. 2015).

5 Nonexperimental evidence draws mixed conclusions (e.g., Fewtrell et al. 2005; Clasen et al. 2015).

Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

The results confirm that deep tubewells can provide arsenic-safe water in rural Bangladesh but suggest that the price per person whose exposure is reduced is between 2 and 4 times previous "best-case scenario" estimates for deep tubewells. These previous estimates assume take-up rates that are higher than those observed in this study.[6] The relatively modest take-up rates observed in this context occur despite the wells being installed using "best practice" participatory approaches shown to be effective in reducing elite capture (Madajewicz et al. 2021), under the supervision of an experienced NGO partner, and with full information about the baseline geography of arsenic contamination. Simulations suggest that increasing take-up would further reduce exposure to arsenic contamination, albeit with diminishing marginal returns if increasing take-up weakens targeting, meaning that new adopters are less likely to be exposed to arsenic at baseline.

The results suggest, however, that deep-tubewell construction programs may be unlikely to substantially reduce exposure to fecal contamination in drinking water in rural Bangladesh. This is partly because of recontamination during transport and storage, but primarily because the deep tubewells appear, in practice, to become contaminated during collective use. Without resolving this problem, simulations suggest that increasing take-up would not further reduce exposure to fecal contamination.

An important caveat is that the measure of fecal contamination used in this study is coarse, meaning that the study can evaluate only whether or not fecal contamination is present and not the level of fecal contamination in project wells. The results may therefore not fully capture the reduction in exposure to fecal contamination at source. Future research is needed to quantify the extent of fecal contamination in deep tubewells and to better understand how deep tubewells become contaminated during collective use.

One particular recent concern is whether interventions designed to reduce exposure to arsenic contamination could have inadvertently increased exposure to fecal contamination (Wu et al. 2011; Buchmann et al. 2019). Information campaigns could lead households to switch to sources that are arsenic-safe but contaminated with fecal bacteria, because the two types of contamination are negatively correlated for hydrogeological reasons. Providing safer but more distant sources could lead to increased contamination through transport and storage. The results in this study rule out large increases in exposure to fecal contamination as a consequence of these interventions, which may assuage these concerns. However, the results of this study leave unresolved the problem of how to reduce exposure to fecal contamination in drinking water in rural Bangladesh.

More broadly, the trade-offs between source quality, transport distance, and storage and the risk of contamination in community water sources are unlikely to be specific to rural Bangladesh. Global elimination of exposure to fecal pathogens in drinking water may require strategies that go beyond improved community water sources.

2. Context

Despite years of effort by the government, nongovernmental organizations, and international aid agencies, progress on safe drinking water in Bangladesh remains elusive. More than 130 million people drink fecally contaminated water and 27 million people drink water that is contaminated with arsenic at international standards (BBS and UNICEF 2019).

Exposure to fecal contamination in drinking water is a longstanding problem in Bangladesh, while the problem of arsenic contamination emerged in the 1990s. Diarrheal disease was a major cause of death during the 1970s and 1980s, when around 20 percent of children died before their fifth birthdays[7] and many households depended on surface water as a source of drinking water. Education campaigns

6 Jamil et al. (2019) assume 60 percent take-up of community deep tubewells among households previously using an unsafe well within 100 m of a new well, drawing on well-switching behaviour described in Madajewicz et al. (2007) and Chen et al. (2007).

7 Authors' calculation based on data provided by the World Bank DataBank.

Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

successfully encouraged people to switch from surface water to shallow, hand-pumped tubewells, likely contributing to a reduction in infant mortality but inadvertently exposing people to arsenic poisoning (Caldwell et al. 2003). Arsenic occurs naturally in shallow groundwater in Bangladesh and is undetectable without chemical tests. Long-term exposure to arsenic leads to a number of serious health impacts (see, e.g., Hong et al. 2014). Daily use of arsenic-contaminated water at the Bangladeshi standard of 50 parts per billion (ppb) is associated with an additional 1 in 100 lifetime risk of cancer, rising to more than 1 in 10 for higher contamination levels (Smith, Lingas, and Rahman 2000). There is also evidence of health impacts from exposure at lower levels, including levels below the more conservative WHO standard of 10 ppb (Rahman et al. 2009). The epidemic of arsenic-related disease in Bangladesh has been called "the largest poisoning of a population in history" (Smith, Lingas, and Rahman 2000).

In many regions of Bangladesh, the primary intervention used to improve access to safe drinking water is the installation of deep tubewells.[8] Deep tubewells extract water from arsenic-safe aquifers, often found at considerable depths below ground, where the aquifers are also isolated from sources of fecal contamination on the surface. However, whether deep-tubewell construction programs successfully improve access to safe drinking water remains uncertain. New well locations may be chosen for political purposes, rather than targeted to those in most need (van Geen et al. 2015).

Additionally, recent studies have raised the concern that some efforts to reduce exposure to arsenic have increased exposure to fecal contamination. This may occur if households switch to other nearby wells that are low in arsenic, which for hydrogeological reasons are more likely to have fecal contamination, or because households that adopt safer but more distant sources increase their exposure to bacterial contamination of drinking water through increased transport or storage times (Buchmann et al. 2019; Wu et al. 2011).

The specific context for this study is north-western Bangladesh, in Shibganj and Sonatala Upazilas, in Bogra District, and in Gobindaganj Upazila, in Gaibandha District, as shown in fig. 1. The study area is not in the epicentre of the arsenic contamination problem, but high arsenic contamination is present in scattered pockets (Daily Observer 2014). The area had few deep tubewells. The implementing partner, the NGO Forum for Public Health (NGO Forum), therefore expected that the impact of providing deep tubewells could be high.

3. The Intervention

The intervention comprises a package of subsidies and technical advice to build deep tubewells.[9] The intervention was developed jointly by the implementing partner, NGO Forum, and the research team, drawing on local best practices and past experience. The deep tubewells installed draw water from aquifers that are sufficiently deep to be isolated from sources of both bacterial and arsenic contamination. Each deep tubewell costs around USD 750[10] to install, of which the project subsidized between 90 and 100 percent.

NGO Forum implemented the program in geographically contiguous communities made up of between 50 and 250 households. Small communities were offered one tubewell and larger communities two tubewells, corresponding on average to 1 well per 118 households.[11]

8 Other technologies that have been proposed have major drawbacks. See, e.g., Howard et al. (2006).

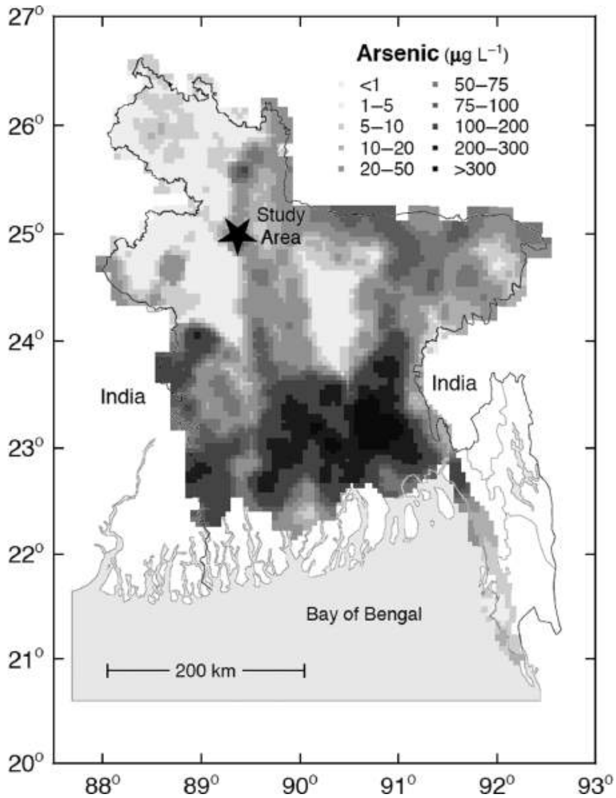
9 Supplementary online appendix S1 provides further details on the intervention and supplementary online appendix S2 describes our monitoring plan.

10 Equivalently, 60,000 Bangladeshi taka (BDT).

11 We designed the tubewell allocation rules to achieve the goals of a parallel study (Cocciolo et al. 2019). Specifically, we implemented one of two rules: (a) we assigned tubewells to villages as a function of village size, then divided these among the designated communities within each village; (b) we assigned tubewells to communities to keep the ratio of households to tubewells as close as possible to 125:1.

Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

Figure 1. Arsenic Contamination in Bangladesh and Study Site Location



Source: Adapted from "Groundwater Studies of Arsenic Contamination in Bangladesh" DPHE/BGS/DFID (2000).
Note: Study area shown with star.

The program uses a participatory approach shown previously to reduce elite capture and increase project impact (Madajewicz et al. 2021). Communities take project decisions by unanimous consensus in open meetings with minimum participation requirements for women and the poor, facilitated by project staff. All households are invited to the meeting, and around half of the households attend. Around two-fifths of participants are women. Households who self-identify as poor or very poor are slightly less likely to attend the meetings than those who self-identify as middle income, but the differences are small. During the meetings, project staff displayed large-scale maps showing all community water sources and their contamination status.

Communities received the program under one of three subsidy arrangements: a third of treated communities were required to contribute USD 75 per well; a third were required to contribute a total of 18 person-days of labour, approximately equivalent in value to the cash contribution requirement when priced at the local unskilled daily wage rate[12]; and a third received the program without a contribution requirement. The focus of this paper is the mean effect of the program on safe drinking water, averaged across the three contribution treatment arms.[13]

12 The cash contribution requirement for each well is 6,000 Bangladeshi taka (BDT) in local currency. Each person-day of labour corresponds to a six-hour shift, consistent with local norms for unskilled labour, and is valued at the local daily unskilled wage rate of BDT 300.

13 We report how the contribution requirement affected program impact in Cocciolo et al. (2020).

Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

Figure 2. Evaluation Timeline

Source: Authors' elaboration.
Note: Bars indicate length of activity.

If communities successfully held a meeting, identified an appropriate site, and met any contribution requirement, contractors attempted to install wells. Contractors used local drilling technology, primarily driven by human manpower, to attempt to reach an arsenic-safe aquifer. A safe aquifer is a layer that is permeable (meaning that water can flow through it relatively freely) but separated from the highly-arsenic-contaminated layers near the surface by an impermeable layer (through which the contaminated water cannot pass). Local drilling technology can penetrate layers of weak or fractured rock, but not solid rock. A field engineer employed by the NGO partner supervised the installation process, and project staff tested all wells to confirm that the water was safe at WHO standards before finalizing the installation.

Program implementation took place between March 2016 and August 2017 (fig. 2), with some piloting beginning in October 2015. The program did not change during the study period, and there were only minor deviations from the study protocol.[14] There was limited scope for implementers to innovate, although the process of facilitating the community meeting required some learning: the only community that failed to reach an agreement was the very first community in which the field team implemented the intervention.

4. Theory of Change

Many programs aim to improve access to safe drinking water by providing communities with new, safe sources of drinking water. A simple theory of change underlies these programs: new, safe sources are built; households adopt the new sources; source water quality improves; and thereby household water quality also improves (fig. 3a). A more nuanced theory of change recognizes that not all households will adopt a new source and that source water quality is only one of the determinants of household water quality (e.g., Wright, Gundry, and Conroy 2004). In particular, longer transport and storage times provide more opportunities for recontamination between the point of collection and the point of use, decreasing household water quality (fig. 3b). These unintended consequences could potentially reverse the effect of the program on household drinking water quality.

5. Study Design

The evaluation is a randomized controlled trial. Figure 4 shows the study design.

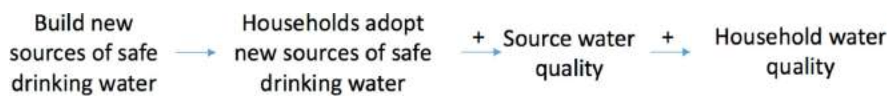
5.1. Study Population and Sample

Field staff recruited communities with high levels of arsenic contamination. We first used preexisting data on arsenic contamination to identify candidate villages. Field staff then obtained lists of resident house-

14 For example, in one community, our community definition protocol was not correctly implemented, resulting in a community consisting of two clusters too geographically distant from one another to hold a single community meeting. As a result, the field staff implemented the project in only one of the two clusters, not the full community. These cases were rare.

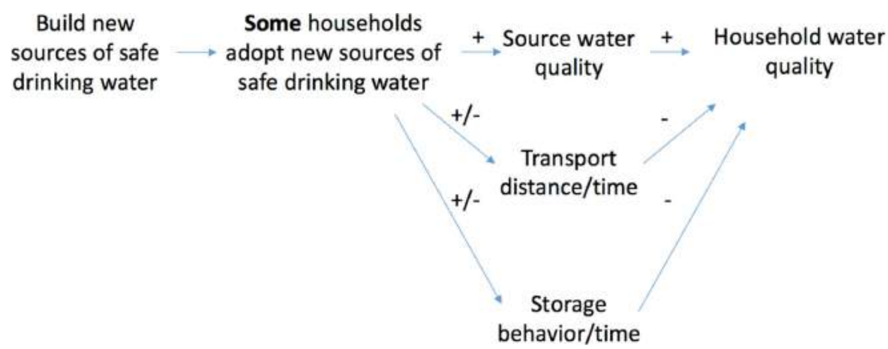
Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

Figure 3a. Theory of Change: Naïve Model



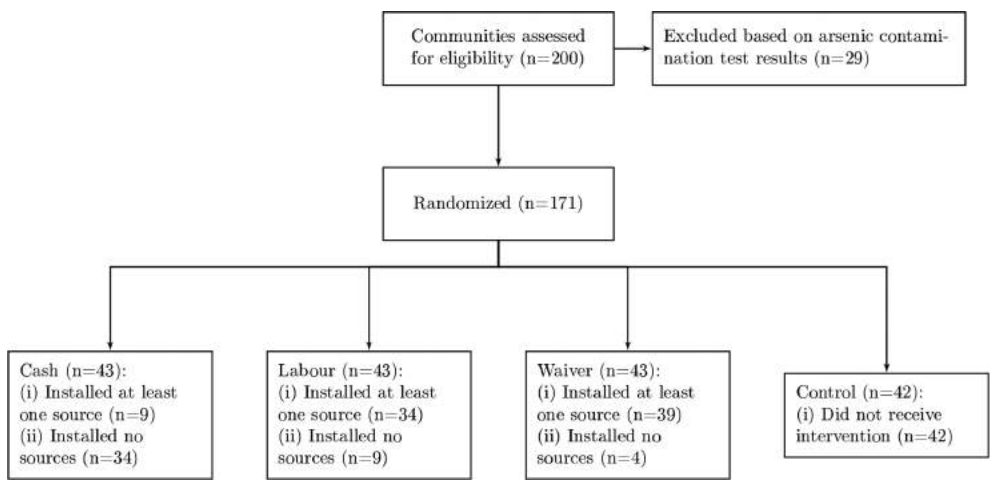
Source: Authors' elaboration.
Note: Arrows indicate casual relationships. +/– signs indicate direction of causal relationships.

Figure 3b. Theory of Change: More Complete Model



Source: Authors' elaboration.
Note: Arrows indicate causal relationships. +/– signs indicate direction of causal relationships.

Figure 4. Evaluation Design Flow Chart



Source: Authors' elaboration.
Note: Figure illustrates evaluation design.

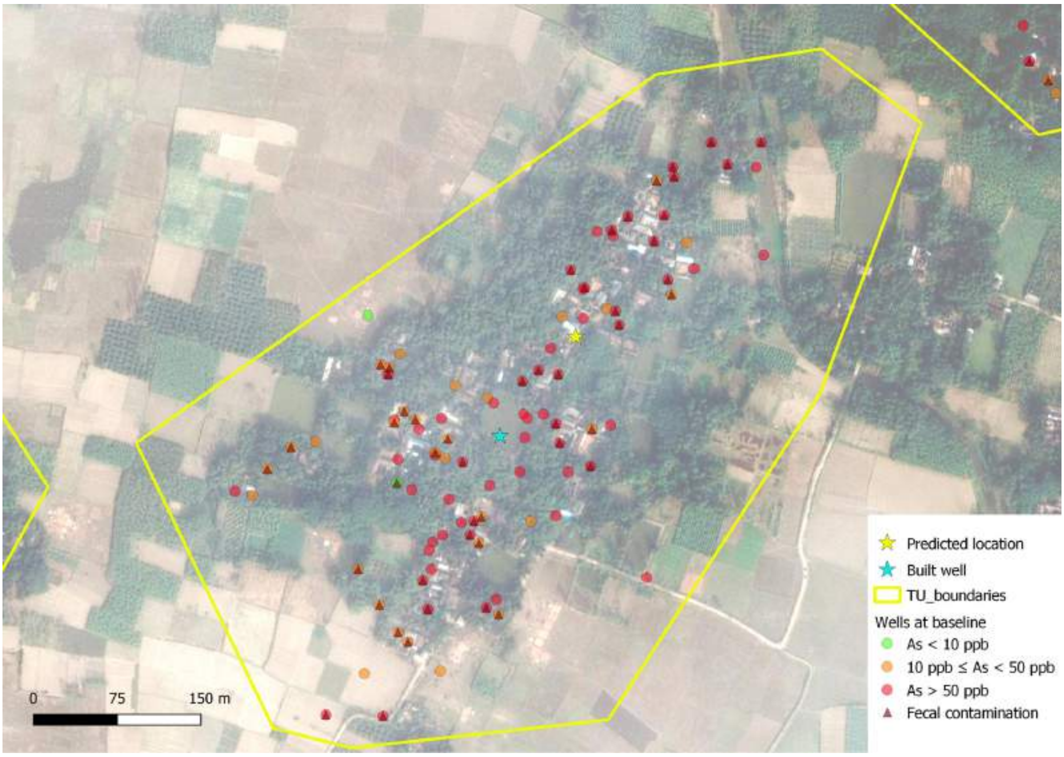
holds from local administrative sources. We excluded villages with fewer than 50 households and divided villages with more than 250 households into several smaller communities along natural boundaries, as illustrated in fig. 5. We then screened these candidate communities for arsenic contamination.

Figure 6a shows the final sample of 171 study communities, as well as the fraction of water sources in each community with arsenic contamination above the WHO threshold at baseline.[15] The final recruitment criteria were that either more than 25 percent of community water sources were contaminated with

15 Supplementary online appendix S3 provides further details.

Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

Figure 5. Example Community



Source: Authors' elaboration based on data sources discussed in the text.
Note: Figure shows one of the study communities, along with the locations of all community wells identified at baseline, the location of the project well, and the predicted location of this well. See discussion in section "How Do Source Water Quality, Transport, and Storage Affect Household Water Quality?"

arsenic, or more than 15 percent of community water sources were contaminated and these sources were spatially clustered.

In each community, we aimed to survey 40 randomly sampled households both before and after the intervention, using the administrative lists of households as a sampling frame.[16] Response rates exceeded 92 percent in both survey rounds, and the household-level attrition rate between baseline and follow-up is 0.7 percent.[17]

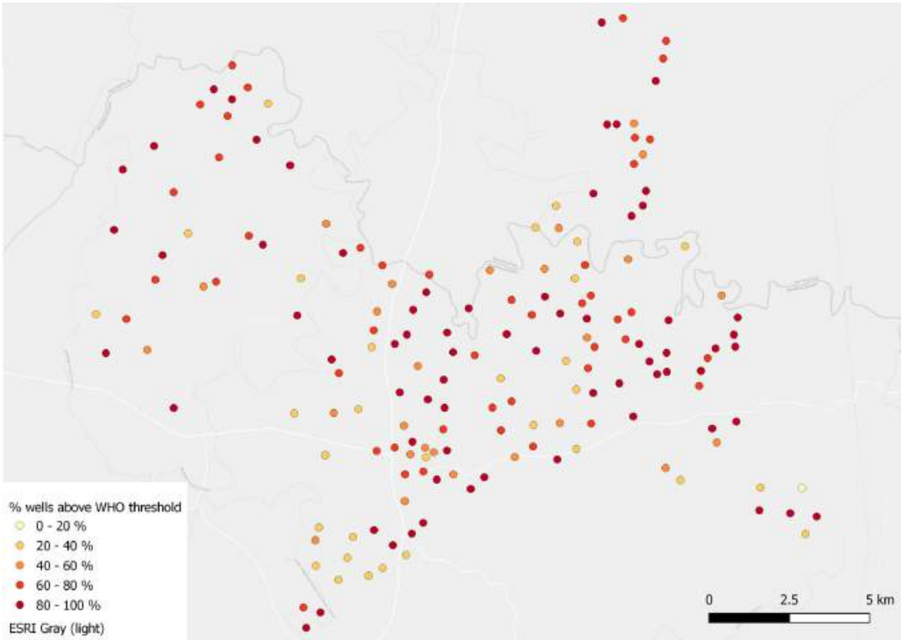
5.2. Data

Enumerators collected data through household surveys and a water quality testing program.[18] During the household survey, enumerators interviewed respondents on household demographics, health, wealth, social networks, and behaviour related to water collection and use. The respondents were household heads and/or their spouses, or another adult member of the household when neither household head nor spouse was available. During the water quality testing program, enumerators collected samples from both household drinking water and water sources used by the household, for testing for both fecal and arsenic contamination. To collect samples of household drinking water, enumerators asked respondents

16 Occasionally, the number of households surveyed in a community was higher or lower than the targeted number. This is because in some cases we had to revise the community definition after completing the baseline household surveys, mostly to correct errors from the administrative lists.
17 Supplementary online appendix S4 provides further details.
18 More details are provided in supplementary online appendix S5.

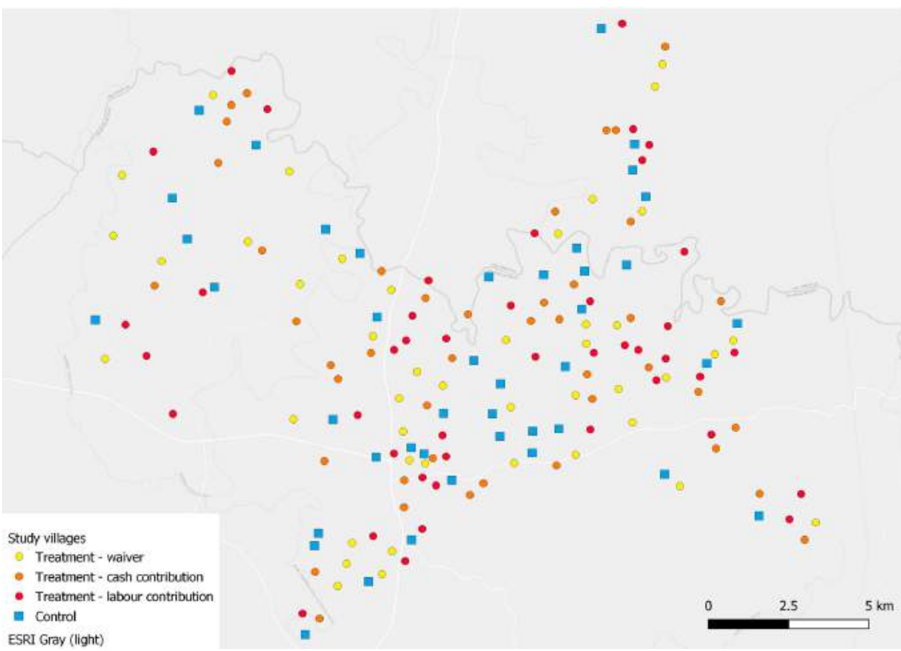
Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

Figure 6a. Baseline Arsenic Contamination in Study Communities



Source: Authors' elaboration based on data sources discussed in the text.
Note: Map shows location of study communities and indicates baseline levels of arsenic contamination.

Figure 6b. Assignment to Treatment



Source: Authors' elaboration based on data sources discussed in the text.
Note: Map shows study communities and their treatment status.

Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

to do as they normally would in order to obtain a glass of water for a household member. Enumerators collected samples from water sources directly. Respondents gave oral informed consent separately for the household survey and for water quality testing.[19] No compensation was provided to respondents, although all respondents were given the opportunity to acquire information about household and water source safety, which the vast majority took up.

Fecal contamination was detected using hydrogen sulphide vials produced by NGO Forum. These tests detect hydrogen-producing bacteria, which are almost exclusively organisms that live in the gut of warm-blooded animals and therefore indicate human or animal fecal contamination.[20] Test results are available after two days. We tracked the samples using barcodes and sent respondents test results by automated SMS.

Arsenic contamination was measured using the Hach EZ Arsenic High Range Field Test Kit, which measures arsenic levels within the range 0–500 ppb. The Hach test kit is low cost, widely available in Bangladesh, and is the test with which the partner NGO had the greatest experience and familiarity.[21] Test results are available after 20 minutes. Enumerators gave respondents the test results after the survey.[22]

When water is stored in an open vessel, arsenic oxidizes and begins to precipitate out, a process known as passive removal or passive sedimentation (see, e.g., Ahmed 2001; Roberts et al. 2004). While passive removal cannot be relied upon as an arsenic removal strategy (Sutherland, Kabir, and Chowdhury 2001), it implies that arsenic concentrations measured in stored drinking water are likely to be lower and more variable than arsenic concentrations measured in wells.[23]

The analysis relies on matching households to water sources and therefore implicitly on households truthfully identifying the water sources they use. Although enumerators ask households to simply list the water sources they use, before discussing any issues related to water safety, social desirability bias might still dissuade households from reporting use of unsafe water sources (Ahuja, Kremer, and Zwane 2010). If underreporting of use of unsafe sources is similar in treated and control communities, it does not affect the estimated effects of the program. However, treatment could hypothetically change reporting behaviour, creating differential reporting bias in treated communities. We discuss this potential concern in the section Effects on Access to Safe Drinking Water.

19 Among households who consented to participate in the survey at baseline, only three did not consent to the testing procedure. At followup, 49 households consented to the survey but not to water source testing. For these households, the test results are set to missing.

20 Gupta et al. (2008) report 88 percent sensitivity and 80 percent specificity to detect E. coli contamination in samples from water sources, while Islam et al. (2017) report 83 percent sensitivity and 49 percent specificity in samples of stored drinking water. Samples should be kept at room temperature for 48 hours, and the test is read as positive if the colour has changed from clear to black. In some cases, particularly at baseline, tests were left for more than 48 hours or entered after fewer than 48 hours. To ensure that data are comparable across rounds, we apply a correction to the data which accounts for variation in how long each test was left before entering the results, using information on the specificity and sensitivity of the test reported in Gupta et al. (2008).

21 Hach reports that the EZ test detects 90 percent of arsenic present as arsenate (As(V)) and 100 percent of arsenic present as arsenite (As(III)) (HACH n.d.). Field performance may be somewhat lower (Steinmaus et al. 2006; van Geen et al. 2005). Before collecting samples to test for arsenic from wells, enumerators pumped the well the same number of times as the depth of the tubewell.

22 van Geen et al. (2005) recommend leaving the test for 40 minutes, instead of the manufacturer-recommended 20 minutes. Using only the manufacturer-recommended 20 minutes may have led us to underestimate arsenic concentrations.

23 Arsenite (As(III)) oxidizes rapidly to arsenate (As(V)) in stored water. Since the Hach test detects arsenate somewhat less effectively than arsenite, this may also cause us to underestimate arsenic concentrations in stored samples of drinking water. Our original study design included a secondary measure of arsenic contamination, using laboratory testing, in a subsample of households. However, we did not have sufficient funds available to complete all the intended laboratory tests at baseline, and correlation is weak between laboratory test results and field test results at follow-up, which may reflect a problem with our tracking systems for the laboratory tests.

Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

Table 1. Access to Safe Drinking Water: Baseline Descriptive Statistics

	Study sample	National rural population
(a) Household water quality tests		
Arsenic contamination (parts per billion)	29	32
	(1.8)	(1.2)
Arsenic contamination (WHO threshold)	0.63	0.61
	(0.017)	(0.009)
Arsenic contamination (Bangladeshi threshold)	0.24	0.17
	(0.016)	(0.006)
Fecal contamination	0.65	0.63
	(0.010)	(0.010)
(b) Water source quality tests (primary water source)		
Arsenic contamination (parts per billion)	37	34
	(2.2)	(1.6)
Arsenic contamination (WHO threshold)	0.69	0.59
	(0.018)	(0.011)
Arsenic contamination (Bangladeshi threshold)	0.31	0.19
	(0.017)	(0.008)
Fecal contamination	0.54	0.39
	(0.010)	(0.011)
(c) Other variables		
Household observed to store drinking water	0.73	0.19
	(0.011)	(0.002)
Household reports treating water before drinking	0.09	0.04
	(0.008)	(0.002)
Time needed to collect water (minutes)	2.2	15
	(0.04)	(0.27)
Water collected per day (litres)	59	79
	(1.2)	(1.2)

Source: Authors' analysis based on data sources discussed in the text.
Note: The table reports means for listed variables in the study sample and a representative sample for rural Bangladesh from BBS and UNICEF (2013). Survey weights applied. Standard errors are clustered by primary sampling unit and shown in parentheses.

Enumerators collected baseline data for most communities in late 2015 and early 2016 (fig. 2). For 16 communities, enumerators collected baseline data in spring 2017, after the study received additional grant funding in 2016. Enumerators collected follow-up data in 2018 in all study communities.

5.3. Sample for Analysis

The final panel sample consists of 6,051 households interviewed at both baseline and follow-up, for which we have household and source water quality data from both rounds.[24] Table 1 describes access to safe drinking water at baseline in this sample, alongside comparable statistics for the national rural

24 We focus on this sample to minimize sample changes between analyses. Some analyses have fewer observations, primarily because we dropped outliers that appeared to be errors. For example, we cleaned the location data of outliers that reflected error in measured GPS coordinates or enumerator error in recording walking times. Enumerators successfully conducted the household survey with 6,529 households at baseline, of which 6,526 also gave consent for water quality testing. At follow-up, enumerators located 6,487 of these households and completed the interview with 6,484 of them, of which 6,434 gave consent for water quality testing. Among these households, we have household drinking water quality data (arsenic and fecal contamination) at both baseline and follow-up for 6,313 households, and source water quality data at both baseline and follow-up for 6,162 households. When data are missing in the final panel dataset, it implies that (a) we could not locate a matching record in the water source survey data or (b) we could not uniquely match the fecal contamination test identifier with our test result database. We used locally produced barcodes, which occasionally contained duplicate ids.

Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

population.[25] At baseline, households almost exclusively obtained water from shallow tubewells, privately owned either by the household or by another close relative living nearby.[26] Arsenic contamination in the study communities is mostly above the national average, reflecting the targeted recruitment of communities that face arsenic contamination problems. Fecal contamination rates are above or similar to the national averages.[27]

5.4. Assignment to Treatment

The study randomly assigned 129 communities to receive the intervention, all of which were offered the safe drinking water program, with 43 communities randomly assigned to each of the three different contribution requirements. The remaining 42 communities were designated a control group and received no intervention, although nothing prevented them from receiving any other interventions or from installing their own safe water sources if they wished to do so.

We assigned communities to the control group and the treatment arms at public lottery meetings, to which we invited representatives from each eligible community. The public lottery meetings provided field staff with an important source of legitimacy if communities questioned why some received the intervention while others did not or why contribution requirements differed across communities. Figure 6b shows the map of communities assigned to the control group and the three treatment arms. Random assignment to treatment was successful in creating statistically comparable groups with respect to baseline characteristics.[28]

We prespecified the analyses we report in this paper (AEA RCT registry, ID number AEARCTR-0002755). At the time, we had analyzed baseline data and completed the intervention but not collected any follow-up data. We follow a pre-analysis plan.[29] When we report additional analyses that we did not prespecify, we indicate that this is the case. The study protocol was approved by Ethical and Independent Review Services.[30]

6. Implementation and Take-Up

Field staff implemented the program in all 129 communities assigned to treatment, which were offered a total of 179 tubewells. The program successfully installed 107 wells in 82 communities. The most common cause of failure, accounting for 44 offered wells in 34 communities, was that communities assigned to the cash contribution requirement did not raise the contribution, despite holding a community meeting and selecting a site. The main reason appears to be the low real value of time in rural communities, which implied that the labour contribution requirement was in practice less onerous than its counterpart in cash, despite their nominal equivalence (Cocciolo et al. 2020). Also, some communities could not identify a suitable site, and some well installations failed for hydrogeological reasons, despite the communities successfully completing all program requirements.[31]

25 Data are from the Multiple Indicator Cluster Survey (MICS) 2012–2013 (BBS and UNICEF 2013).

26 The mean total time required to collect drinking water is around 2 minutes. This is lower than average for Bangladesh's rural population. In some parts of Bangladesh, there is much greater quantitative freshwater scarcity, particularly the Barind Tract region in northwestern Bangladesh, the coastal regions and the hilly regions of Chittagong and Sylhet (Ahmed and Hassan 2012). There were also slight differences in measurement between our study and MICS that may partially account for the differences.

27 Supplementary online appendix S6 provides an extended description of the context in terms of socioeconomic characteristics, along with more details on the variables reported in table 1.

28 Supplementary online appendix S7 provides details.

29 Supplementary online appendix S8 lists the prespecified key research questions and measures of outcome variables.

30 Supplementary online appendix S9 provides further details.

31 No suitable land could be secured for 13 offered wells in 12 communities. Installation failed for hydrogeological reasons in 13 wells in 12 communities. Specifically, installation could not be completed either due to the presence of an impen-

Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

Table 2. Comparison of Project Wells with Other Water Sources

	Fecal contamination (1)	Arsenic contamination (WHO threshold) (2)	Arsenic contamination (Bangladeshi threshold) (3)
Difference	−0.13***	−0.58***	−0.35***
	(0.05)	(0.04)	(0.03)
Mean (project tubewell)	0.34	0.06	0.01
Mean (other tubewells)	0.46	0.63	0.34
N	3,399	3,515	3,515

Source: Authors' analysis based on data sources discussed in the text.
Note: The table reports the regression-estimated difference in contamination rates in project tubewells compared to other water sources in the same communities, from a regression that includes community fixed effects. The table also reports mean contamination levels in project tubewells and other water sources in the same communities. The sample includes water sources that at least one study household reported using for drinking or cooking, in communities in which we installed at least one project well. Standard errors are clustered by community and shown in parentheses. *** $p < 0.01$.

Table 2 summarizes water quality in project tubewells, compared to other wells in the same communities. Project tubewells almost all have arsenic contamination below the WHO threshold. However, project tubewells are only 13 percentage points (28 percent) less likely to test positive for fecal contamination than are other wells in the same communities. The deep tubewells draw water from an aquifer which is isolated from sources of fecal contamination, but contamination could potentially take place through leakage into the pipe system from shallow groundwater or within the pump body itself. Other research points to the potential for contamination through the tubewell mouth (ICDDRB and UNICEF 2018) and shows that the tubewell body may act as a reservoir for microbial organisms (Ferguson et al. 2011).

The study communities are small and relatively compact. In communities in which the project successfully installed at least one tubewell, the median household lives 1.6 minutes walking time from the project source, and 26 percent of households report collecting at least some of their household's drinking and cooking water from a project source. However, the large majority of the households who report using the source also report that they continue to use other sources. Only 13 percent of the households who use the project source report collecting all of their drinking and cooking water from the source. On average, in these communities, 10 percent of drinking and cooking water is reported to be obtained from a project well. Scaling by community size and numbers of installed wells, about 27 households collect at least some water from each installed well and total take-up is equivalent to about 11 households collecting all their water from the well.

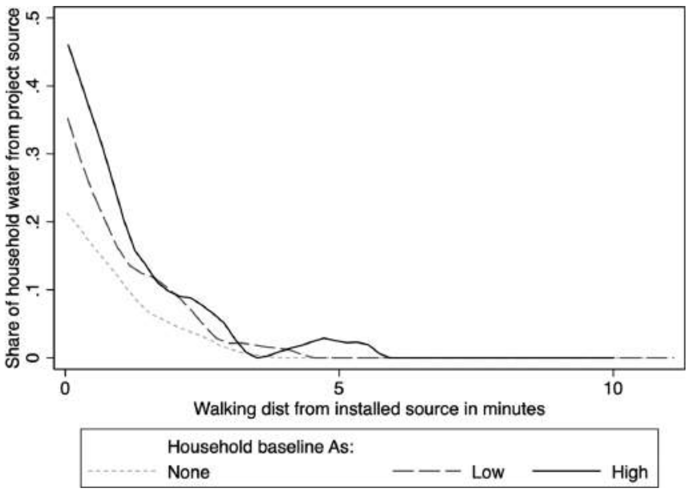
Households who live closer to installed wells are more likely to adopt the new wells (fig. 7a). Even at short distances from installed wells, some households do not adopt the well and many of those that do adopt it continue to collect water from other sources. Take-up declines rapidly with distance.

The most commonly cited reason for nonadoption is that the wells are too far from the household, even when the wells are less than one minute's walking distance away. A majority of nonadopters also report that women from the household feel uncomfortable collecting water outside the compound. Some households do not adopt the source because they believe their own well to be safe. Relatively few households cite waiting times, discomfort with using a well on someone else's private land, or a lack of concern about arsenic as reasons for nonadoption (fig. 7b).

etrable rocky layer or a sandy layer which caused the excavation to collapse before the PVC pipe could be installed. If installation failed for hydrogeological reasons, field staff returned cash contributions to households and compensated households who had contributed labour with BDT 300 per person per shift. Additionally, one community declined to hold a meeting at all, and one community held a meeting but could not agree on a location. See table S17.1 in the supplementary online appendix for details.

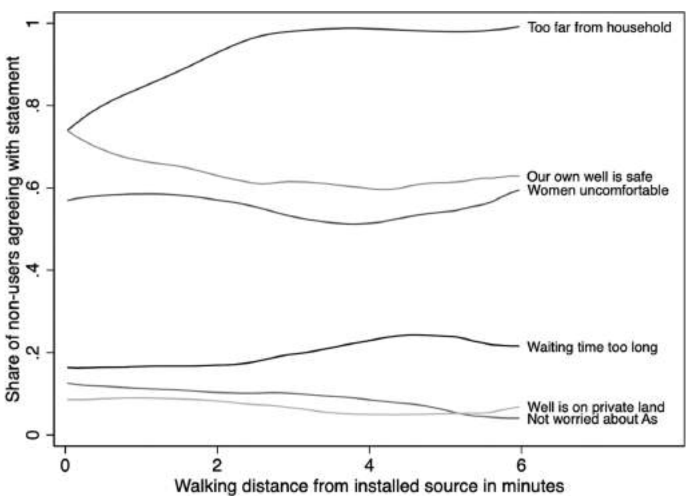
Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

Figure 7a. Take-Up of Project Wells



Source: Authors' elaboration based on data sources discussed in the text.
Note: Graph shows mean share of drinking and cooking water obtained from project well by distance to well and baseline arsenic contamination in household drinking water. Data from communities in which we successfully installed wells.

Figure 7b. Reasons Given by Nonadopters



Source: Authors' elaboration based on data sources discussed in the text.
Note: Graph shows share of households who do not use project well that agree or strongly agree with the listed statement. We report results only for statements with which at least 5 percent of respondent households agreed or strongly agreed. Data from communities in which we successfully installed wells.

7. Effects on Access to Safe Drinking Water

This section describes the average program impacts on changes in behaviour with respect to how households obtain drinking water—the quality of the sources they use, how far they walk to collect water, and whether they store water in the household before drinking—and on the quality of the drinking water at the point of use in the household.

The main estimating equation is

$$\Delta y_{ic} = \alpha + \beta_T T_c + \mathbf{Z}_c \boldsymbol{\gamma} + \epsilon_c,$$

Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

where Δy_{ic} is the change in outcome variable y between baseline and follow-up in household i in community c,[32] T_c is an indicator that takes the value 1 if community c is assigned to treatment and 0 otherwise, and the vector $\boldsymbol{Z}_c$ contains stratification controls.[33] The parameter of interest is β_T, which measures the mean difference-in-differences between treatment and control groups. Because T_c is randomly assigned to communities, it is independent of any other determinants of the outcomes of interest, meaning that β_T has a causal interpretation. We cluster standard errors by community to account for correlated changes in outcome variables within communities (Bertrand, Duflo, and Mullainathan 2004). As prespecified, we also report p values obtained from randomization-based inference by simulating the treatment assignment process 500 times.

The estimated program impacts are "intent-to-treat" effects, averaged across all communities in which the program was implemented regardless of whether or not wells were successfully installed. Since households differ at follow-up in the amount of time they have been exposed to the treatment, the estimated impacts correspond to weighted averages of treatment effects over the first two years following program implementation.[34] We also report analyses that scale the average effects by the number of wells installed per household.[35] We refer to these estimates throughout as the "scaled" estimates. Under the assumption that the effects of the program are proportional to the number of wells installed per household, the scaled estimates can be interpreted as the average effect of installed wells, normalized by the number of households in each community, in the population of communities who successfully installed wells under the program. These analyses were not prespecified.

We report results for all prespecified research questions, using the prespecified main outcome variables.[36] Simulation-based power calculations using observed installation rates suggest that our study was powered to detect mean changes of 3.5 percent and 3.8 percent in arsenic and fecal contamination, respectively, in household drinking water, with greater power to detect changes in use of contaminated sources.[37]

How does the program change the quality of sources that households use? Table 3 shows the estimated effects on water quality in sources from which households obtain water for drinking and/or cooking. When households use multiple sources, we calculate average contamination across the sources used, weighting by the fraction of water collected from each source, as specified in the pre-analysis plan. In this table, as with tables 4 through 6, the coefficient reported as the constant corresponds to the change between

32 We depart from the prespecified approach in one minor respect, and analyze data at the household level, applying weights so that each community counts equally in the analysis, and clustering standard errors at the community level. The prespecified approach was to collapse the data to village-level means. The estimated point effects are mechanically identical when estimated at the household level but are slightly more precisely estimated. This results from making less conservative adjustments to standard errors for the stratification controls.

33 We include controls for each lottery at which treatment was assigned: one lottery in most unions and two lotteries in one of the larger unions. Following Lin (2013), Imbens and Rubin (2015) and Gibbons, Serrato, and Urbancic (2019), we demean lottery fixed effects and include the interaction term between the lottery controls and the treatment dummies, ensuring that β_T consistently estimates the average difference between treated and control villages.

34 Additionally, the time between baseline and follow-up differs for the villages added later to the study. Any difference between these villages and other study villages are absorbed by the stratification controls.

35 Specifically, we use the three contribution treatment dummies as instrumental variables to predict the number of installed wells per household in each community. The estimates thus correspond to a Local Average Treatment Effect (LATE), or effect on the compliers, those that installed wells as a result of the program. Details are in supplementary online appendix S11.

36 Details are in supplementary online appendix S8. We do not report results for one of the prespecified secondary measures of arsenic contamination: arsenic concentration measured using laboratory tests. These data were incomplete and sample tracking did not appear to be reliable.

37 Details are in supplementary online appendix S10.

Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

Table 3. Effect of the Program on Use of Contaminated Sources

	Arsenic contamination (WHO threshold) (1)	Arsenic contamination (Bangladeshi threshold) (2)	Arsenic contamination level (3)	Fecal contamination (4)
Treated	−0.056**	−0.027	−0.155	−0.021
	(0.023)	(0.019)	(3.648)	(0.019)
Constant	−0.068***	−0.008	7.642**	−0.035**
	(0.018)	(0.015)	(3.167)	(0.015)
p-value (analytical)	0.02	0.15	0.97	0.28
p-value (RBI)	0.04	0.23	0.95	0.38
Mean at baseline	0.69	0.30	37.27	0.55
R^2	0.03	0.02	0.07	0.04
N	6,051	6,051	6,051	6,051

Source: Authors' analysis based on data sources discussed in the text.
Note: The table shows estimated average program impact on use of sources with listed contaminant. Regression in first differences, including stratification controls. Analysis is at the household level with weights applied so that each community counts equally in the analysis. Standard errors clustered by community are shown in parentheses. p-values obtained from analytical standard errors and randomization-based inference (RBI). ** $p < 0.05$, *** $p < 0.01$.

Table 4. Effect of the Program on Transport and Storage

	Distance HH-WS (m) (1)	Distance HH-WS (min) (2)	Observed storage (3)	Reported storage (4)
Treated	−0.147	0.065**	0.002	0.003
	(1.041)	(0.027)	(0.028)	(0.032)
Constant	−2.801***	0.012	−0.134***	−0.040
	(0.685)	(0.024)	(0.024)	(0.029)
p-value (analytical)	0.89	0.02	0.95	0.93
p-value (RBI)	0.93	0.02	0.95	0.92
Mean at baseline	31.40	0.81	0.73	0.76
R^2	0.01	0.01	0.02	0.02
N	5,832	5,729	6,050	6,051

Source: Authors' analysis based on data sources discussed in the text.
Note: HH = household; WS = water source. The table shows estimated average program impact on listed measure of water-related practice. Regression in first differences, including stratification controls. Analysis is at the household level with weights applied so that each community counts equally in the analysis. Standard errors clustered by community are shown in parentheses. p-values obtained from analytical standard errors and randomization-based inference (RBI). ** $p < 0.05$, *** $p < 0.01$.

baseline and follow-up in the control group, while the coefficient labelled "treated" corresponds to the estimated treatment effect.

Use of sources with arsenic contamination above the WHO threshold falls by 5.6 percentage points in treated communities relative to control communities (column 1), while use of sources with arsenic contamination above the Bangladeshi threshold falls by 2.7 percentage points (column 2). The weighted-average arsenic contamination level in sources that households use also falls on average, although only slightly: by just under 0.1 ppb (column 3). Figure 8a visualizes the effects across the full distribution of changes in arsenic contamination. The overall effects are driven by changes near the center of the distribution.

There is a comparatively modest decrease in use of sources with fecal contamination, equal to 2.1 percentage points with a confidence interval that does not exclude zero (column 4). The effect on source fecal contamination is just over a third of the effect on arsenic contamination. The smaller effect on fecal contamination is mostly explained by the pattern shown in table 2. While project water sources are almost entirely free of arsenic, they are only 28 percent less likely to have fecal contamination than other wells

Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

Table 5. Effect of the Program on Storage Conditions

	Containers are uncovered (self-reported) (1)	Containers are on the floor (self-reported) (2)	Containers are uncovered (observed) (3)	Containers are on the floor (observed) (4)	Water is scooped (observed) (5)
Treated	0.023	0.008	0.040*	0.025	−0.002
	(0.039)	(0.030)	(0.021)	(0.019)	(0.022)
Constant	−0.351***	−0.092***	0.383***	0.484***	0.439***
	(0.033)	(0.027)	(0.018)	(0.016)	(0.020)
Only endline data	–	–	✓	✓	✓
p-value (analytical)	0.55	0.78	0.06	0.20	0.93
p-value (RBI)	0.63	0.82	0.04	0.24	0.92
Mean at baseline	0.54	0.74	–	–	–
R^2	0.04	0.02	0.02	0.01	0.01
N	6,051	6,051	5,902	5,997	6,029

Source: Authors' analysis based on data sources discussed in the text.
Note: The table shows estimated average program impact on listed measure of storage practice. Regression in first differences, unless otherwise indicated, including stratification controls. Analysis is at the household level with weights applied so that each community counts equally in the analysis. Standard errors clustered by community are shown in parentheses. p-values obtained from analytical standard errors and randomization-based inference (RBI). * $p < 0.1$, *** $p < 0.01$.

Table 6. Effect of the Program on Household Water Quality

	Arsenic contamination (WHO threshold) (1)	Arsenic contamination (Bangladeshi threshold) (2)	Arsenic contamination level (3)	Fecal contamination (4)
Treated	−0.022	−0.004	3.756	0.002
	(0.020)	(0.018)	(2.865)	(0.020)
Constant	−0.096***	−0.023	1.218	−0.033*
	(0.016)	(0.015)	(2.414)	(0.017)
p-value (analytical)	0.27	0.82	0.19	0.93
p-value (RBI)	0.36	0.87	0.25	0.93
Mean at baseline	0.62	0.24	28.97	0.65
R^2	0.02	0.01	0.04	0.03
N	6,051	6,051	6,051	6,051

Source: Authors' analysis based on data sources discussed in the text.
Note: The table shows estimated average program impact on listed household water quality measure. Regression in first differences, including stratification controls. Analysis is at the household level with weights applied so that each community counts equally in the analysis. Standard errors clustered by community are shown in parentheses. p-values obtained from analytical standard errors and randomization-based inference (RBI). * $p < 0.1$, *** $p < 0.01$.

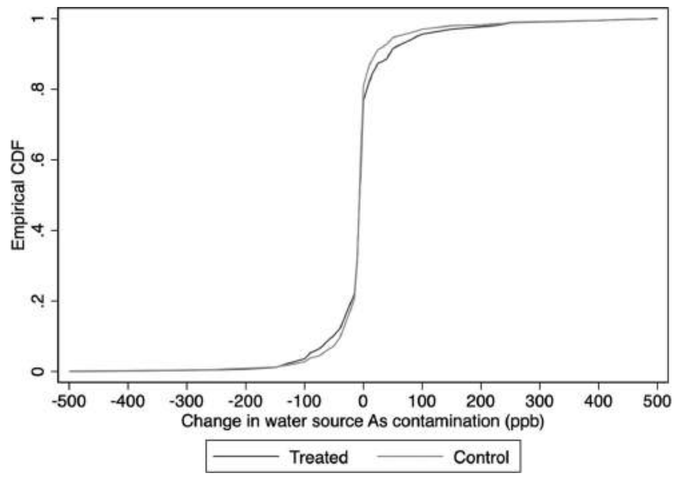
in the same communities. As a result, not all households who adopt project wells reduce their exposure to fecal contamination at source. Had all our project sources been free from fecal contamination and take-up patterns remained unchanged, use of sources with fecal contamination would have fallen by 4.5 percentage points.[38] The reduction in use of sources with fecal contamination would still have been smaller than the reduction in use of sources with arsenic contamination, because the wells are better targeted to households using sources contaminated with arsenic at baseline than households using sources with fecal contamination.

Scaling these estimates by the fraction of offered wells actually installed, the results suggest that the impact of each project well installed is equivalent to inducing 10 households to switch from sources with arsenic contamination at the WHO level to sources free from arsenic contamination and 5 households to switch from sources with fecal contamination to sources free from fecal contamination.

38 See fig. 9a.

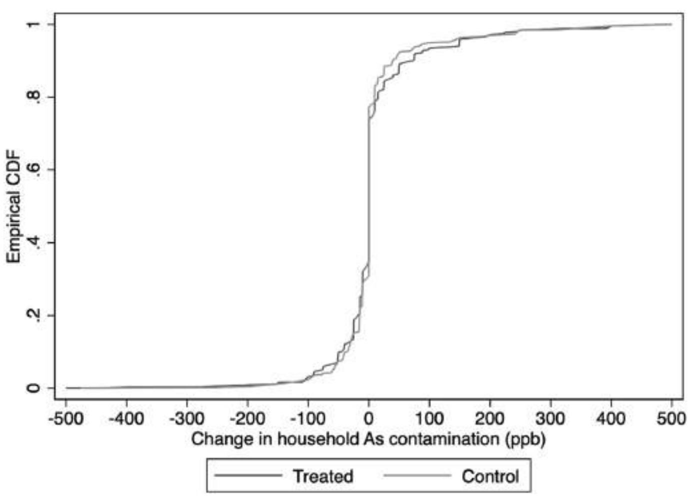
Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

Figure 8a. CDF of Change in Arsenic Contamination: Water Source Quality



Source: Authors' elaboration based on data sources discussed in the text.
Note: Figure show empirical cumulative distribution function (CDF) for changes in mean arsenic contamination between baseline and follow-up in sources used by study households.

Figure 8b. CDF of Change in Arsenic Contamination: Household Drinking Water



Source: Authors' elaboration based on data sources discussed in the text.
Note: Figure show empirical cumulative distribution function (CDF) for changes in mean arsenic contamination between baseline and follow-up in samples from household drinking water.

Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

The estimated program effects would be biased downwards by spillover effects if households in control communities also adopted project wells. However, no household in any control community reported using a project tubewell at follow-up. The absence of spillovers may not be unexpected, given strong preferences for nearby sources.

Use of sources with arsenic contamination at the WHO threshold fell by 6.8 percentage points in the control group between baseline and follow-up. The data suggest that this is the result of fluctuations in arsenic contamination in groundwater over time.[39] In a sample of 6,013 wells for which we have contamination readings at both baseline and follow-up, the rate of arsenic contamination at the WHO standard fell by 6.3 percentage points between baseline and follow-up.[40] This sample constitutes more than 80 percent of the wells used by study households. The decline in contamination in these wells is almost identical in magnitude to the decline in contamination in sources used by households in the control group.[41]

Some other studies document that households respond to information about arsenic contamination by switching away from arsenic-contaminated wells (Madajewicz et al. 2007; Bennear et al. 2013; Balasubramanya et al. 2014; Huhmann et al. 2019).[42] We do not find that well-switching plays an important role in reducing use of arsenic-contaminated sources in the control group in this study. In the control group, 12 percent of households adopt new sources during the study period. Households exposed to arsenic contamination at baseline are slightly more likely (2 to 4 percentage points) to adopt a new well than households with no arsenic contamination. However, most control households who adopt new sources switch to new private shallow wells, which are no more likely to be arsenic safe than other wells in the same communities. Use of a new well is not systematically correlated with a reduction in arsenic contamination for control households.

How does the program change how far households walk to collect drinking water? Table 4 shows the effects of the program on distance walked to collect drinking water. We estimate effects on distance using measured distances between households and water sources, calculated from GPS coordinates (column 1), and walking times reported by households (column 2). The estimated effects on measured distance and reported walking time differ from one another. The estimated effect on measured distance is small, negative, and not statistically different from zero. The estimated effect on reported walking time is positive and statistically significant, equivalent to about 5 m walking distance.

The difference between the two sets of results may result from measurement error in the GPS coordinates, which leads to overestimation of distances (Ranacher et al. 2016).[43] The less precise the measurement, the more distances are overestimated. This potentially affects the results in table 4 in the following way. We measure the location of our installed wells with greater accuracy than the location of other sources that households use, because we average over multiple measures of location and we verify the locations by inspection. As a result, if households adopt our wells, we overestimate the distance between their household and the project well less than we overestimate the distance between their

39 Supplementary online appendix S12 provides an extended descriptive analysis of changes in the control group. We note that it is also possible that the observed changes reflect changes in how our enumerators measured arsenic contamination. However, enumerators followed exactly the same testing process and received the same training at both baseline and follow-up.

40 Other studies document similar magnitude fluctuations in arsenic contamination in wells surveyed repeatedly over similar time frames (e.g., Bhattacharya et al. 2011).

41 We also estimate a fall in use of sources with fecal contamination in the control group. This could also reflect changes in contamination over time or be an artefact of changes in the way we processed the fecal contamination tests. Field staff adhered more closely to the 48 hour protocol for entering test results at follow-up.

42 Related, Tarozzi et al. (2020) report that charging a modest price for arsenic testing also leads to well-switching, albeit at lower rates and lower cost-effectiveness.

43 There are two reasons. First, any measurement error that is orthogonal to the true distance leads to an increase in the measured distance. Second, if the true distance is within the support of the distribution of measurement error, as is often the case in this context, errors that lead to distance overestimation are not cancelled out by errors that lead to distance underestimation. To see this most intuitively, consider two measurements of the location of the same point. The true distance between the point and itself is obviously zero, but in the presence of any measurement error the distance between the two measurements will not be zero.

Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

household and the water source they previously used.[44] Consequently, we underestimate any increase in distance walked. This bias could cancel out any true increase in distance walked, explaining the differences between columns (1) and (2).

This pattern of measurement error could also explain why households in the control group appear to have reduced the distance they walk to collect water between baseline and follow-up when we use measured distance but not when we use self-reported walking time. We used more accurate tablets at follow-up than at baseline. We therefore overestimate distances between households and sources whose location was only measured at baseline more than distances between households and sources whose location was only measured at follow-up. When households switch from a source used only at baseline to a source used only at follow-up, this pattern of measurement error negatively biases the estimated changes in walking distance.

How does the program change whether and how households store drinking water before use? Table 4 also shows the effects of the program on whether or not households store drinking water before use. We report effects on whether enumerators observed households obtaining their samples of drinking water from storage (column 3) and whether households self-report storing drinking water before use (column 4). The differences between treated and control communities are small, suggesting that the intervention does not lead to large changes in storage practice.

Table 5 reports whether the program leads to changes in unsafe storage practice. Across five indicators of unsafe storage, four increase and one decreases in treated communities as compared to control communities, but only one increase is statistically significant at the 10 percent level. Three of these indicators were only measured at follow-up. These results suggest that the program leads, at most, to small changes in unsafe storage behaviour.

Some variables in tables 4 and 5 show substantial changes between baseline and follow-up in the control group. We did not necessarily collect data at the same time of year at baseline and at follow-up, so these changes may simply reflect seasonal differences in behaviour.

How does the program change household water quality? Table 6 reports the effects of the program on contamination in household drinking water.

Arsenic contamination at the WHO standard in household drinking water fell by 2.2 percentage points in treated communities compared to control communities (column 1). This effect is smaller than the estimated reduction in use of arsenic-contaminated sources, although the confidence interval includes both zero and the estimated effect on use of arsenic-contaminated sources. Arsenic contamination at the higher Bangladeshi threshold also falls, albeit by an insignificant amount (column 2). Arsenic test results actually rise slightly (column 3), although the confidence interval includes reductions in arsenic concentrations and it appears that the point estimate is driven by outliers.[45] Figure 8b visualizes the effects across the full distribution of changes in arsenic contamination in household drinking water, showing that the program effects are again driven by changes in the center of the distribution. Scaling these estimates by well installation rates, the data suggest that each project well installed effectively eliminates household arsenic contamination for around 5 households.

One potential explanation for the smaller effects on arsenic contamination in household drinking water than on use of arsenic-contaminated sources could be differential reporting bias in the treated and

44 For all households, and for sources for which we have locations at both baseline and follow-up, we average the recorded locations after excluding locations that are clearly wrong, such as those which place the household in a different community.

45 The distribution of arsenic test results is highly skewed: only 1.5 percent of households have arsenic contamination above 250 ppb and 0.3 percent have arsenic contamination above 500 ppb.

Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

control groups. Exposure to the program could make households less likely to admit to continued use of unsafe sources. Households in treated villages might also overreport use of project sources because of experimenter demand effects. This could lead us to overstate the effects of the program when we evaluate effects on source-level contamination. However, our data suggest that the results are unlikely to be influenced by differential reporting bias. We evaluate the relationship between source and household measures of contamination. Misreporting should attenuate the relationship between source and household contamination and increase the fraction of household variation that is unexplained by source contamination. If treated households change their reporting behaviour in response to the intervention, the relationship between household and source contamination should also change. On the contrary, however, we find that the relationship between household and source contamination is similar in treated and control communities.[46] A more probable explanation for the smaller effects on arsenic contamination in household drinking water is that we are less likely to detect arsenic in household drinking water because of passive removal in stored drinking water.

Ex ante, it was ambiguous whether the program would increase or decrease exposure to fecal contamination in drinking water. The point estimate is a 0.2 percent increase in contamination and the 90 percent confidence interval spans both modest increases (3.0 percent) and decreases (3.4 percent) in contamination (column 4). Scaling by well installation rates, these effects are equivalent to each well installed introducing fecal contamination into household drinking water for around 2 households.

Respondents might take more care to avoid fecal contamination if they know that the sample of water is going to be tested, because of social desirability bias. For example, they might wash their hands and vessels more scrupulously than usual. This might lead us to underestimate the extent of fecal contamination in household drinking water. Since the data collection process is identical in both treatment and control groups, this would only influence the estimated effects if households in treated communities change their behaviour between baseline and follow-up differently to households in control communities. Differential behaviour change might arise if households exposed to a program that is focused on water safety consequently improved their hygiene practices or, conversely, if households in control villages, who received information about water contamination but no interventions to improve access to safe drinking water, exerted more effort to reduce contamination through other channels, for example by improving hygiene (a "John Henry" effect). Given the small point estimates and the modest change in fecal contamination in household drinking water in the control group, any such differential changes in behaviour are unlikely to be important in magnitude.

The real policy objective of the safe drinking water program is improved health. Although we collected data on health outcomes and mortality, we focus on water quality for two reasons. First, our measures of drinking water quality have the benefit of being largely objective measures of project impact and are therefore unlikely to be susceptible to reporting bias. Second, our study is underpowered to detect impacts on health, as adverse health events are uncommon and may take time to manifest, especially in the case of arsenic poisoning.[47]

Heterogeneous effects We prespecified a number of heterogeneity analyses, which in the interests of brevity we primarily report in the supplementary online appendix.[48] One potentially important finding is that effects on use of water sources with fecal contamination vary strongly with poverty level. Middle- and upper-income groups reduce their use of sources with fecal contamination, while the poor increase their use of sources with fecal contamination. The differences are sufficiently large that they would

46 Details are in supplementary online appendix S13.

47 We did not have sufficient funds to test for biomarkers of arsenic exposure (such as urine, as in, e.g., Ahsan et al. 2000) in addition to testing water quality. Supplementary online appendix S14 describes data on health outcomes.

48 See supplementary online appendix S15.

Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

Table 7. Effects of Source Water Quality, Transport and Storage on Household Water Quality

	Drinking water fecal contamination			
	(1)	(2)	(3)	(4)
Source fecal contamination	0.242***	0.224***	0.245***	0.225***
	(0.015)	(0.015)	(0.015)	(0.015)
Travel time HH-WS (mins, reported)	0.015*	0.018**	0.016**	0.018**
	(0.008)	(0.009)	(0.008)	(0.009)
Observed storage	0.081***	0.070***	–	–
	(0.010)	(0.010)		
Constant	−0.012	−0.015***	−0.023**	−0.024***
	(0.010)	(0.002)	(0.010)	(0.001)
Treatment unit fixed effects	No	Yes	No	Yes
R^2	0.07	0.14	0.06	0.14
Obs	5,728	5,728	5,729	5,729

Source: Authors' analysis based on data sources discussed in the text.
Note: HH = household; WS = water source. Regression in first differences. Analysis is at the household level with weights applied so that each community counts equally in the analysis. Community fixed effects where specified. Standard errors clustered by community. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

survive conservative Bonferroni corrections for multiple hypothesis testing. One possible explanation is that wells used by poor households become contaminated more quickly, while deep tubewells used by middle- and higher-income households remain uncontaminated, perhaps because poor households have less information about how to avoid well contamination during use, or less time or fewer resources to dedicate to well cleaning.

8. How Do Source Water Quality, Transport, and Storage Affect Household Water Quality?

This section describes how changes in source water quality, walking distance, and storage influence water quality in household drinking water. We focus on fecal contamination because arsenic contamination only occurs at the source. We estimate a difference-in-difference equation:

$$\Delta \mathrm{FC}_i^b = b_0 + b_1 \Delta \mathrm{FC}_i^w + b_2 \Delta \mathrm{DIST}_i^w + b_3 \Delta \mathrm{STORAGE}_i + \eta_c + \epsilon_i. \tag{1}$$

All variables are changes, denoted Δ, between baseline and follow-up. The variable FC^b is fecal contamination in household drinking water and FC^w is mean fecal contamination in water sources used by the household, weighted by volume obtained from each source. The variable DIST^w is the weighted mean distance travelled by the household to collect water from its water source or sources. Given the concerns about measurement error that we describe in the previous section, we use the self-reported measure of walking time to water sources in this analysis. The variable STORAGE is an indicator variable for whether or not household i stores drinking water, as opposed to collecting drinking water on demand. The term η_c denotes a community-level dummy variable that absorbs community mean changes in both the outcome variable and the right-hand-side variables. We estimate equation (1) with and without these community fixed effects. When we include community fixed effects, we exploit only within-community variation to estimate the relationships of interest.[49] The coefficients of interest are b_1, b_2, and b_3.

Table 7 shows the results for four specifications: with and without community fixed effects, and with and without accounting for changes in storage. The results suggest that switching to a source with fecal contamination increases the risk of contamination in household drinking water by between 22 and 24 percentage points. Increasing travel time by one minute increases the risk of contamination

49 There was no clear ex ante reason to prefer one approach over the other. Either approach might yield higher precision, depending on the exact structure of ϵ_i.

Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

by between 1.5 and 1.8 percentage points. Storing drinking water increases the risk of contamination by between 7 and 8 percentage points. The estimated effects are very stable across all specifications. Omitting the controls for changes in storage practice slightly increases the estimated effects of transport time on drinking water contamination. This is because increasing travel times are associated (weakly) with increasing storage.[50]

Although we did not prespecify how we would measure distance in the analysis described in this section, we originally intended to use the measure of distance calculated from GPS coordinates. Using this measure, the estimated effects of source contamination and storage are very similar to those in table 7, but the estimated effects of transport distance are smaller, the confidence intervals are wider, and the results do not rule out the possibility that increasing travel time has no effect on contamination.[51] The most likely explanation for the difference between the two sets of results is increased measurement error in the measured distances, biasing the effects of distance towards zero through attenuation bias.

The difference-in-difference approach yields causal estimates under the assumption that the changes on the right-hand side of the equation are uncorrelated with other changes in behaviour, such as hygiene practices, that also affect fecal contamination in household drinking water. This assumption may be reasonable, since the program did not provide extensive or differential information on other types of health or hygiene behaviour. Additionally, as table 7 shows, the difference-in-difference results are stable when we include or exclude controls for changes in storage behaviour or community fixed effects, suggesting that the effects of unobserved changes in hygiene behaviour on contamination would have to be considerably larger than the combined effect of all community-level changes and storage to meaningfully affect the results. However, although assignment to the safe drinking water program is random, the wells are not installed in randomly chosen locations. As a result, changes in travel distance or changes in source water contamination could still be correlated with changes in household drinking water contamination through other channels, potentially biasing the estimated effects.

To address these concerns, we prespecified an instrumental variables (IV) analysis exploiting the experimental design of the safe drinking water program. Using baseline data, we predict the choices communities make about where to locate a well, and, in turn, changes in behaviour, in particular changes in use of sources with fecal contamination and changes in distance walked to collect water. The advantage of this approach, in principle, was to eliminate any potential bias in the difference-in-difference analysis arising from the endogeneity of the well location. In practice, however, the instruments were much weaker than expected. We successfully predict the locations of wells, as well as spatial patterns of well adoption. For example, fig. 5 shows the location we predicted as well as the final location of the project well. However, the instrument for changes in source fecal contamination was weak because the rates of fecal contamination in project wells were higher than anticipated. The instrument for changes in distance was also weak. In most specifications, the IV point estimates take the same sign as the difference-in-difference estimates, but the confidence intervals are very wide. Since the IV analyses provide little additional information, we report them in detail only in supplementary online appendix S16.

9. How Would Increasing Take-Up Have Affected Impact?

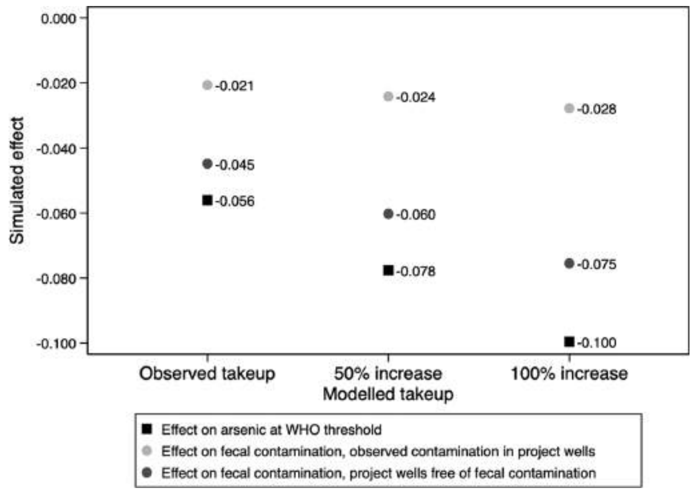
Take-up in this study was lower than anticipated by previous research. In this section, we simulate how increasing take-up might have changed program impact. We focus on a simple scenario. Beginning with observed take-up patterns, we proportionally increase the fraction of water households collect from project wells, always respecting the feasibility constraint that households can collect at most 100 percent of their water from a project well, until the share of the community's water coming from project wells increases by a given factor. This approach respects existing adoption patterns in the data: households living closer

50 Results are available on request.

51 Results are shown in table S17.2 in the supplementary online appendix.

Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

Figure 9a. Simulated Changes in Use of Contaminated Water Sources under Increased Adoption Scenarios



Source: Authors' elaboration based on data sources discussed in the text.
Note: Graph shows simulated effects on use of sources with arsenic and fecal contamination under indicated scenarios.

to the well and households with higher baseline arsenic contamination collect a larger share of their water from the well. We then simulate effects on use of contaminated sources and travel distance and, drawing on the estimates described in the previous section, model how these changes translate into effects on fecal contamination in household drinking water.[52]

Increased take-up further reduces use of arsenic-contaminated water sources, although there are diminishing marginal returns (fig. 9a). Since households with higher arsenic contamination already collect more of their water from the well, there is lower potential for them to increase take-up. Increasing take-up thus weakens the relationship between baseline arsenic contamination and take-up. Doubling take-up increases the reduction in use of arsenic-contaminated water sources by 78 percent, suggesting an elasticity of impact to take-up of 0.78.

Increased take-up has a smaller proportional effect on use of sources with fecal contamination. Doubling take-up, while holding constant observed rates of fecal contamination in project wells, increases the reduction in use of sources with fecal contamination by about 33 percent, an elasticity of impact to take-up of 0.33. Had the project wells been free from fecal contamination, the elasticity of impact to take-up would be closer to that for arsenic contamination.

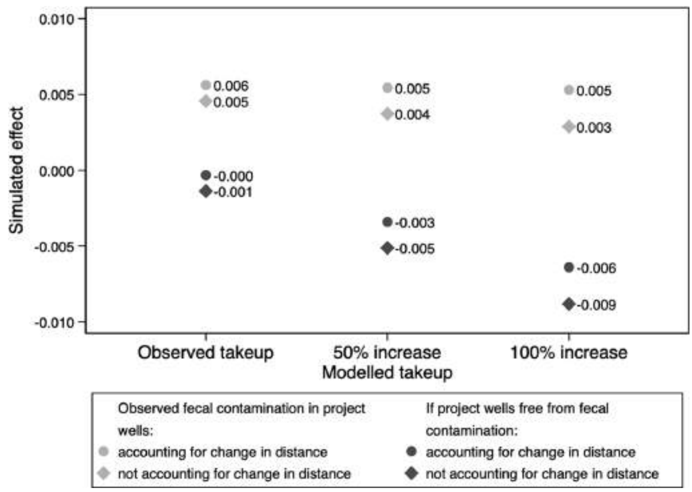
The simulation suggests that fecal contamination in household drinking water would have increased even if take-up doubled (fig. 9b). Although doubling take-up reduces use of sources with fecal contamination, households walk further to collect water, which in turn increases contamination, offsetting the improvements in source water quality.[53] If project wells were entirely free of contamination, but take-up remained as observed, the program also would not lead to improvements in fecal contamination in household drinking water. Only if project wells had been free from fecal contamination and take-up had been

52 Simulated take-up patterns are shown in fig. S17.1 in the supplementary online appendix. We did not prespecify the simulations in this section. We use the estimates from column (3) of table 7 to model fecal contamination in household drinking water because we do not simulate changes in storage behaviour. Because travel distance and storage behaviour are correlated, the effects of changes in storage behaviour are captured by the simulated effects of changes in travel distance.

53 Simulated effects on travel distance are shown in fig. S17.2 in the supplementary online appendix.

Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

Figure 9b. Simulated Changes in Household Drinking Water Quality under Increased Adoption Scenarios



Source: Authors' elaboration based on data sources discussed in the text.
Note: Graph shows simulated effects on fecal contamination in household drinking water under indicated scenarios.

higher do we model a reduction in fecal contamination in household drinking water. Even in this case, the reduction is small.

We focus on a single, plausible scenario for how take-up might have increased. Evaluating a wider range of take-up scenarios is beyond the scope of this paper. However, the conclusions we draw from this exercise may be reasonably general. Many scenarios for increased take-up will lead to lower baseline arsenic concentrations and higher walking distances among the additional adopters, and improving targeting for both arsenic and fecal contamination is difficult because the two are negatively correlated.

10. Discussion

This paper uses a randomized controlled trial to evaluate the impact of a community deep-tubewell construction program on access to safe drinking water in rural Bangladesh.

Each deep tubewell installed leads about 10 households, comprising around 40 individuals, to switch from arsenic-contaminated sources to arsenic-safe sources.[54] Including only the well installation costs, the cost of avoiding arsenic contamination would be BDT 3,080 or USD 18.5 per person whose exposure is reduced.[55] This cost is double previous "best-case scenario" estimates (Jamil et al. 2019), despite implementing the program under local best practice conditions with full information about the distribution of arsenic contamination at baseline. At this cost, providing deep tubewells to all 27 million Bangladeshi citizens who drink water contaminated with arsenic would cost USD 500 million in well installation costs alone. Based on the relationship between source arsenic contamination and cancer reported in Smith et al. (2000), the effects suggest that one death from cancer would be prevented by every 3.3 wells installed.

The deep-tubewell program has little impact on fecal contamination. The net effect is to introduce fecal contamination into household drinking water for two households. About 3 in 10 households contain at least one child under five, and each child exposed to fecal contamination in household drinking water

54 Each household has on average 3.9 members.

55 The average cost of well installation is BDT 60,000 or approximately USD 720 at current exchange rates. Using the estimated effects on household contamination would suggest lower cost-effectiveness, BDT 1,540 or USD 37 per person whose exposure is reduced.

Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

might feasibly experience an additional mortality risk of 0.4 in 1,000, about a 1 percent increase in risk.[56] Deep tubewells have lower but still substantial rates of fecal contamination under current use and maintenance practices. Increasing transport time or storage also heightens the risk of contamination. Simulations suggest that the intervention as implemented would have been unlikely to reduce fecal contamination in household drinking water, even if take-up were substantially increased.

The results suggest that cost-benefit comparisons across safe drinking water technologies should consider effects on fecal contamination as well as arsenic contamination. For example, local piped water systems may cost more than deep tubewells (Jamil et al. 2019). However, local piped water systems might conceivably reduce fecal contamination more effectively than community deep tubewells, because fewer households use each water point, reducing the potential for contamination through collective use, and water points are closer to households, reducing the potential for contamination during transport and storage. Accounting for effects on fecal contamination could alter conclusions about which technologies can best provide safe drinking water in rural Bangladesh.

The fecal contamination tests used in this study only detect the presence or absence of fecal bacteria. Future research is needed to confirm whether deep tubewells have lower concentrations of fecal bacteria when they are present. Additionally, research is needed to confirm the mechanisms via which fecal contamination is introduced into deep tubewells, and to understand whether different cleaning practices could improve water quality and whether such practices would be adopted by caretakers.

A baseline water testing program enabled field staff to selectively recruit communities exposed to arsenic contamination. Communities also used the information from this water testing program to select well locations. Collecting water source census data is relatively costly, but in its absence, the program might have been less successful in targeting households with arsenic contamination. A key question for future research is whether the benefits of this information justify the costs of collecting it.

While the findings are most applicable to the context of rural Bangladesh, they may generalize to other settings. Our results confirm that fecal contamination of household drinking water is difficult to eliminate when households collect water from communal sources and store it before drinking. Eliminating exposure to fecal pathogens in drinking water may require complementary interventions, such as training to improve tubewell maintenance, or alternative approaches to providing safe access to drinking water.

References

Ahmed, M. F. 2001. “An Overview of Arsenic Removal Technologies in Bangladesh and India.” In *Proceedings of BUET-UNU International Workshop on Technologies for Arsenic Removal from Drinking Water, Dhaka* (pp. 5–7).

Ahmed, R., and S. Hassan. 2012. *Hard-to-reach Areas: Providing Water Supply and Sanitation Services to All. Supporting Poor-Inclusive Water Supply and Sanitation Sector Reforms.* Water and Sanitation Program: Guidance Note.

Ahsan, H., M. Perrin, A. Rahman, F. Parvez, M. Stute, Y. Zheng, and A. Hasnat Milton et al. 2000. “Associations between Drinking Water and Urinary Arsenic Levels and Skin Lesions in Bangladesh.” *Journal of Occupational and Environmental Medicine* 42 (12): 1195–1201.

Ahuja, A., M. Kremer, and A. P. Zwane. 2010. “Providing Safe Water: Evidence from Randomized Evaluations.” *Annual Review of Resource Economics* 2 (1): 237–56.

56 Household drinking water testing positive for fecal contamination is associated with a 17 percent increase in diarrhea in under 5s over the following week (Islam et al. 2017). Under-5 mortality in Rajshahi is 37 deaths per 1,000 live births (BBS and UNICEF 2019). Diarrhea is the cause of death for 6 percent of under-5 deaths in Bangladesh (UNICEF 2015). The calculation assumes, following Kremer et al. (2011), that household drinking water quality affects mortality from diarrhea by the same proportion as diarrheal disease.

Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

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

Banerjee, A. V., and E. Duflo. 2009. “The Experimental Approach to Development Economics.” *Annual Review of Economics* 1 (1): 151–78.

Bangladesh Bureau of Statistics (BBS) and UNICEF Bangladesh. 2013. *Bangladesh Multiple Indicator Cluster Survey (MICS) 2012–2013*. BGD_2012_MICS_v01_M. Dataset downloaded from http://www.childinfo.org/mics4_surveys.html. Accessed 17th October 2017.

———. 2014. *Bangladesh Multiple Indicator Cluster Survey 2012–2013: Progotir Pathey: Final Report*. Bangladesh Bureau of Statistics (BBS) and UNICEF Bangladesh, Dhaka, Bangladesh.

———. 2019. *Progotir Pathey, Bangladesh Multiple Indicator Cluster Survey 2019, Survey Findings Report*. Dhaka, Bangladesh: Bangladesh Bureau of Statistics and United Nations Children’s Fund.

Bennear, L., A. Tarozzi, A. Pfaff, S. Balasubramanya, K. M. Ahmed, and A. van Geen. 2013. “Impact of a Randomized Controlled Trial in Arsenic Risk Communication on Household Water-Source Choices in Bangladesh.” *Journal of Environmental Economics and Management* 65 (2): 225–40.

Bennett, D. 2012. “Does Clean Water Make You Dirty? Water Supply and Sanitation in the Philippines.” *Journal of Human Resources* 47 (1): 146–73.

Bertrand, M., E. Duflo, and S. Mullainathan. 2004. “How Much Should We Trust Differences-in-Differences Estimates?” *Quarterly Journal of Economics* 119 (1): 249–75.

Bhattacharya, P., M. Hossain, S. N. Rahman, C. Robinson, B. Nath, M. Rahman, and M. M. Islam et al. 2011. “Temporal and Seasonal Variability of Arsenic in Drinking Water Wells in Matlab, Southeastern Bangladesh: A Preliminary Evaluation on the Basis of a 4 Year Study.” *Journal of Environmental Science and Health, Part A* 46 (11): 1177–84.

Bloom, D. E., and D. Canning. 2008. “Population Health and Economic Growth.” Commission on Growth and Development Working Paper No. 24. World Bank: Washington, DC. https://openknowledge.worldbank.org/handle/10986/28036.

Buchmann, N., E. M. Field, R. Glennerster, and R. N. Hussam. 2019. “Throwing the Baby Out with the Drinking Water: Unintended Consequences of Arsenic Mitigation Efforts in Bangladesh (No. w25729).” National Bureau of Economic Research.

Caldwell, B. K., J. C. Caldwell, S. N. Mitra, and W. Smith. 2003. “Tubewells and Arsenic in Bangladesh: Challenging a Public Health Success Story.” *International Journal of Population Geography* 9 (1): 23–38.

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

Clasen, T. F., K. T. Alexander, D. Sinclair, S. Boisson, R. Peletz, H. H. Chang, and F. Majorin et al. 2015. “Interventions to Improve Water Quality for Preventing Diarrhoea.” *Cochrane Database of Systematic Reviews* 10.

Cocciolo, S., S. Ghisolfi, A. Habib, and A. Tompsett. 2020. “How Do Community Contribution Requirements Affect Local Public Good Provision? Experimental Evidence from Safe Water Sources in Bangladesh.” Working Paper.

Cocciolo, S., A. Habib, and A. Tompsett. 2019. “Group Size and Collective Action: Evidence from Bangladesh.” Working Paper.

Cutler, D., and G. Miller. 2005. “The Role of Public Health Improvements in Health Advances: The Twentieth-Century United States.” *Demography* 42 (1): 1–22.

Daily Observer. 2014. “Tube-Well Water Arsenicated in Bogra: 161 Villages in Eight Upazilas Affected.” Newspaper Article, 18th December. https://www.observerbd.com/2014/12/18/61337.php.

Devoto, F., E. Duflo, P. Dupas, W. Parienté, and V. Pons. 2012. “Happiness on Tap: Piped Water Adoption in Urban Morocco.” *American Economic Journal: Economic Policy* 4 (4): 68–99.

Ferguson, A. S., B. J. Mailloux, K. M. Ahmed, A. van Geen, L. D. McKay, and P. J. Culligan. 2011. “Hand-Pumps as Reservoirs for Microbial Contamination of Well Water.” *Journal of Water and Health* 9 (4): 708–17.

Fewtrell, L., R. B. Kaufmann, D. Kay, W. Enanoria, L. Haller, and J. M. Colford, Jr., 2005. “Water, Sanitation, and Hygiene Interventions to Reduce Diarrhoea in Less Developed Countries: A Systematic Review and Meta-Analysis.” *The Lancet Infectious Disease* 5 (1): 42–52.

Genthe, B., N. Strauss, J. Seager, C. Vundule, F. Maforah, and R. Kfir. 1997. “The Effect of Type of Water Supply on Water Quality in a Developing Community in South Africa.” *Water Science and Technology* 35 (11–12): 35–40.

Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

Gibbons, C. E., J. C. S. Serrato, and M. B. Urbancic. 2014. "Broken or Fixed Effects?," (No. w20342) *National Bureau of Economic Research* 8: (1).

Goel, V., M. S. Islam, M. Yunus, M. T. Ali, A. F. Khan, N. Alam, and A. S. G. Faruque et al. 2019. "Deep Tubewell Microbial Water Quality and Access in Arsenic Mitigation Programs in Rural Bangladesh." *Science of the Total Environment* 659: 1577–84.

Gupta, S. K., M. A. Sheikh, M. S. Islam, K. S. Rahman, N. Jahan, M. M. Rahman, and R. M. Hoekstra et al. 2008. "Usefulness of the Hydrogen Sulfide Test for Assessment of Water Quality in Bangladesh." *Journal of Applied Microbiology* 104 (2): 388–95.

HACH. n.d. "Knowledge Base: Hach Parameter FAQ." https://www.hach.com/cms-view.jsa?page=/cms/static/support/sp-KB-ParameterFAQ-en.jsp#03.

Hong, Y. S., K. H. Song, and J. Y. Chung. 2014. "Health Effects of Chronic Arsenic Exposure." *Journal of Preventative Medicine & Public Health* 47 (5): 245–52.

Hoque, B. A., K. Hallman, J. Levy, H. Bouis, N. Ali, F. Khan, and S. Khanam et al. 2006. "Rural Drinking Water at Supply and Household Levels: Quality and Management." *International Journal of Hygiene and Environmental Health* 209 (5): 451–60.

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

Huhmann, B. L., C. F. Harvey, A. Navas-Acien, J. Graziano, F. Parvez, Y. Chen, and M. Argos et al. 2019. "Changes in Arsenic Exposure in Araihazar, Bangladesh from 2001 through 2015 following a Blanket Well Testing and Education Campaign." *Environment International* 125: 8222–89.

Imbens, G. W., and D. B. Rubin. 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences*. Cambridge University Press.

International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDRB)/United Nations Children's Fund (UNICEF). 2018. "A Bacteriological Water Quality Issue: Rohingya 'Forcefully Displaced Myanmar Citizen' Camp in Cox's Bazar, Bangladesh." Fact Sheet 8. https://reliefweb.int/report/bangladesh/fact-sheet-8-bacteriological-water-quality-issue-rohingya-forcefully-displaced.

Islam, M., A. Ercumen, A. M. Naser, L. Unicomb, M. Rahman, B. F. Arnold, and J. M. Colford Jr et al. 2017. "Effectiveness of the Hydrogen Sulfide Test as a Water Quality Indicator for Diarrhea Risk in Rural Bangladesh." *American Journal of Tropical Medicine and Hygiene* 97 (6): 1867–71.

Jamil, N. B., H. Feng, K. M. Ahmed, I. Choudhury, P. Barnwal, and A. van Geen. 2019. "Effectiveness of Different Approaches to Arsenic Mitigation over 18 years in Araihazar, Bangladesh: Implications for National Policy." *Environmental Science & Technology* 53 (10): 5596–5604.

Kremer, M., J. Leino, E. Miguel, and A. P. Zwane. 2011. "Spring Cleaning: Rural Water Impacts, Valuation, and Property Rights Institutions." *Quarterly Journal of Economics* 126 (1): 145–205.

Lin, W. 2013. "Agnostic Notes on Regression Adjustments to Experimental Data: Reexamining Freedman's Critique." *Annals of Applied Statistics* 7 (1): 295–318.

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

Madajewicz, M., A. Tompsett, and M. A. Habib. 2021. "How Does Delegating Decisions to Communities Affect the Provision and Use of a Public Service? Evidence from a Field Experiment in Bangladesh." *Journal of Development Economics* 150: 102609.

Pitt, M., M. R. Rosenzweig, and N. Hassan. 2021. "Identifying the Hidden Costs of a Public Health Success: Arsenic Well Water Contamination and Productivity in Bangladesh." *Review of Economic Studies* forthcoming.

Rahman, M. M., J. C. Ng, and R. Naidu. 2009. "Chronic Exposure of Arsenic via Drinking Water and Its Adverse Health Impacts on Humans." *Environmental Geochemistry and Health* 31 (1): 189–200.

Ranacher, P., R. Brunauer, W. Trutschnig, S. Van der Spek, and S. Reich. 2016. "Why GPS Makes Distances Bigger than They Are." *International Journal of Geographical Information Science* 30 (2): 316–33.

Roberts, L. C., S. J. Hug, T. Ruettimann, M. M. Billah, A. W. Khan, and M. T. Rahman. 2004. "Arsenic Removal with Iron (II) and Iron (III) in Waters with High Silicate and Phosphate Concentrations." *Environmental Science & Technology* 38 (1): 307–15.

Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022

Shields, K. F., R. E. Bain, R. Cronk, J. A. Wright, and J. Bartram. 2015. “Association of Supply Type with Fecal Contamination of Source Water and Household Stored Drinking Water in Developing Countries: A Bivariate Meta-Analysis.” *Environmental Health Perspectives* 123 (12): 1222–31.

Smith, A. H., E. O. Lingas, and M. Rahman. 2000. “Contamination of Drinking-Water by Arsenic in Bangladesh: A Public Health Emergency.” *Bulletin World Health Organization* 78 (9): 1093–1103.

Steinmaus, C. M., C. M. George, D. A. Kalman, and A. H. Smith. 2006. “Evaluation of Two New Arsenic Field Test Kits Capable of Detecting Arsenic Water Concentrations Close to 10 μg/L.” *Environmental Science & Technology* 40 (10): 3362–66.

Sutherland, D., M. O. Kabir, and N. A. Chowdhury. 2001. “Rapid Assessment of Technologies for Arsenic Removal at the Household Level.” In *Proceedings of BUET-UNU International Workshop on Technologies for Arsenic Removal from Drinking Water, Dhaka*. BUET-UNU.

Tarozzi, A., R. Maertens, K. M. Ahmed, and A. van Geen. 2020. “Demand for Information on Environmental Health Risk, Mode of Delivery, and Behavioral Change: Evidence from Sonargaon, Bangladesh.” *World Bank Economic Review* forthcoming.

UNICEF, New York. 2015. “Committing to Child Survival: A Promise Renewed Progress Report 2015.” http://www.unicef.org/publications/files/APR_2015_9_Sep_15.pdf.

UNICEF/WHO. 2011. “Drinking Water Equity, Safety and Sustainability: Thematic Report on Drinking Water 2011.” Available at: https://www.zaragoza.es/contenidos/medioambiente/onu/1194-eng.pdf.

van Geen, A., K. M. Ahmed, E. B. Ahmed, I. Choudhury, M. R. Mozumder, B. C. Bostick, and B. J. Mailloux. 2016. “Inequitable Allocation of Deep Community Wells for Reducing Arsenic Exposure in Bangladesh.” *Journal of Water, Sanitation and Hygiene for Development* 6 (1): 142–50.

van Geen, A., K. M. Ahmed, A. A. Seddique, and M. Shamsudduha. 2003. “Community Wells to Mitigate the Arsenic Crisis in Bangladesh.” *Bulletin of the World Health Organization* 81 (9): 632–38.

van Geen, A., Z. Cheng, A. A. Seddique, M. A. Hoque, A. Gelman, J. H. Graziano, and H. Ahsan et al. 2005. “Reliability of a Commercial Kit to Test Groundwater for Arsenic in Bangladesh.” *Environmental Science & Technology* 39 (1): 299–303.

Wright, J., S. Gundry, and R. Conroy. 2004. “Household Drinking Water in Developing Countries: A Systematic Review of Microbiological Contamination between Source and Point-of-Use.” *Tropical Medicine & International Health* 9 (1): 106–17.

Wu, J., A. van Geen, K. M. Ahmed, Y. A. J. Alam, P. J. Culligan, V. Escamilla, and J. Feighery et al. 2011. “Increase in Diarrheal Disease Associated with Arsenic Mitigation in Bangladesh.” *PLoS One* 6 (12): e29593.

Downloaded from https://academic.oup.com/wber/article/35/4/969/6291932 by guest on 20 February 2022