Series Name: WPS
Paper No.: 048
Issue Date: 19 August 2015

Disgust, Shame and Soapy Water: Tests of Novel Interventions to Promote Safe Water and Hygiene

Raymond P. Guiteras, David I. Levine, Stephen P. Luby, Thomas H. Polley, Kaniz Khatun-e-Jannat, Leanne Unicomb

Disgust, Shame and Soapy Water:
Tests of Novel Interventions to Promote Safe Water and Hygiene

Raymond P. Guiteras[a,b], David I. Levine[c], Stephen P. Luby[d],

Thomas H. Polley[e], Kaniz Khatun-e-Jannat[f], Leanne Unicomb[f,†]

August 2015

Abstract

Lack of access to clean water is among the most pressing environmental problems in developing countries, where diarrheal disease kills nearly 700,000 children per year. While inexpensive and effective practices such as chlorination and handwashing with soap exist, efforts to motivate their use by emphasizing health benefits have seen only limited success. This paper measures the effect of messages appealing to negative emotions (disgust at consumption of human feces) and social pressure (shame at being seen consuming human feces) on handwashing behavior and use of and willingness to pay for water chlorination among residents of slum compounds in Dhaka, Bangladesh. Neither the traditional, health-based message nor the new disgust and shame message led to high levels of chlorination during a free trial, nor to high willingness to pay for the chlorine at the end of the free trial. Provision of low-cost handwashing facilities did increase handwashing, although the effect size is modest.

Keywords: environmental health; water, sanitation and hygiene; social pressure; disgust; Becker-Degroot-Marschak; sequential randomization

JEL Classifications: O12, O13, Q53, Q56, C93, I15

---

[a] Department of Economics, University of Maryland, College Park, MD 20742 USA.

[b] Corresponding author. Tel. +1 301 405 3480; fax +1 301 405 3582. guiteras@econ.umd.edu

[c] Haas School of Business, University of California, Berkeley, CA 94720 USA

[d] Woods Institute for the Environment, MC 4205, Stanford University, Stanford, CA 94305 USA.

[e] Department of Economics, Duke University, Durham, NC 27708 USA.

[f] Water, Sanitation and Hygiene Research Group, ICDDR,B, GPO Box 128, Dhaka 1000, Bangladesh.

† The authors gratefully acknowledge financial support from the Bill and Melinda Gates Foundation, the Clausen Center for International Business and Policy at U.C. Berkeley, and the International Initiative for Impact Evaluation, Inc. (3ie) through the Global Development Network (GDN). Valerie Curtis, Richard Stevenson and Peter Winch provided invaluable comments on an early draft protocol. We thank James Berry, Maureen Cropper, Sebastian Galiani, Jill Luoto, Ted O'Donoghue, Stephanie Rennane and participants at the 2012 UC Berkeley WASH Conference, the HELVETAS 2012 Workshop, and SEEDEC 2013 for useful comments. We learned a great deal from our collaboration with Michael Kremer on the Introduction of Cholera Vaccine in Bangladesh (ICVB) study, and we are especially grateful to him. We thank Rabiul Chowdhury, Sanam Ispahani, Rahin Khandker, Minhaj Mahmud, Mahfuzur Rahman and Bijoya Sarker for their contributions to the field work. The views expressed in this article are not necessarily those of 3ie or its members. The funders had no role in study design or execution, analysis, writing of the paper, or decision to submit for publication. Competing interests: S.L. is currently on the Board of Advisors of Evidence Action, a non-profit that promotes, among other interventions, chlorine dispensers. S.L. joined the board in May 2014, after the research reported in this paper was conducted. Evidence Action had no role in the study design or execution, analysis, writing of the paper, or decision to submit for publication. The other authors declare no competing interests.

1 Introduction

Lack of access to clean water is among the most pressing environmental problems in developing countries. Approximately 1.1 million people die each year from diarrheal disease, including 700,000 children under age 5 (Walker et al. 2013). Typhoid fever, which is commonly transmitted through contaminated water, claims an additional 190,000 lives per year (Lozano et al. 2012, Baker et al. 2011). In addition to this mortality impact, water-borne diseases can have long-lasting negative consequences for human capital development. Childhood episodes of diarrhea contribute to undernutrition, which in turn can lead to stunting, wasting and reduced cognitive development (Black et al. 2008). Furthermore, repeated exposure to gastrointestinal pathogens can permanently harm the body's ability to absorb nutrition, reducing the effectiveness of compensatory household behavior or nutritional interventions (Lin et al. 2013, Taniuchi et al. 2013).

From a biological perspective, the solution is known: safe sanitation prevents pathogens from entering the environment; while water treatment and handwashing prevents infection by pathogens that do reach the environment (Hunt 2006, Acharya and Paunio 2008). However, from the point of view of economics, the problem is more difficult. First, individuals often lack the relevant information and are either unaware of the link between untreated water and disease (Gupta et al. 2008) or hold beliefs that contradict good health practices (Bennett, Naqvi, and Schmidt 2015). In addition, present bias and the lack of salience of the link between prevention activities and health may reduce water treatment and handwashing (Kessler and Zhang 2015). As a result, private behavior is unlikely to lead to privately optimal outcomes. Second, behaviors to prevent water-borne disease carry large positive externalities, so even if individuals' behavior were privately optimal, prevention behavior would be less than the social optimum. Third, there are large economies of scale in providing clean water and sanitation, so a fully efficient water and sewer system requires large-scale investment, coordination and solving challenging issues of governance.

In today's rich countries, governments typically solve these problems by investing in the large-scale infrastructure needed to provide water and sanitation, raising bonds and levying taxes to pay for their construction, and mandating their use. These public interventions are generally credited with the historical reductions in water-borne disease in developed countries (Cutler and Miller 2006, Mackenbach 2007). To date, this strategy has not been successful in poor countries. For example, no major city in India provides treated water without interruption (McKenzie and Ray 2009).[1] High costs and weak institutions are the most commonly cited reasons for this deficit (Zwane and Kremer 2007).

There are several possible strategies for addressing this problem. In the long run, at least in urban environments, infrastructure-based solutions are likely to be most efficient. However, this requires waiting for sufficient economic growth and institutional development, discovering interventions to improve institutional capacity, or finding alternative institutions.[2] In the short term, though, it is important to understand the determinants of behaviors individuals and households undertake to protect themselves from environmental risks. Important dimensions of private behavior include propensity to use and willingness to pay for different health technologies. Willingness to pay is especially important when institutional or government budget constraints require private cost-sharing.

A neoclassical model of behavior in the face of environmental threats to health posits that individuals rationally weigh the perceived private costs and benefits of available mitigating actions, and choose the behavior that maximizes their expected utility.[3] The implication of this model is that a policymaker who wishes to increase a behavior should promote interventions that reduce costs (both financial and non-financial) and increase perceived health benefits. Such interventions include the subsidization or free provision of hardware, or educational programs to increase awareness of the relationship between untreated water and diarrheal disease.

However, water or hygiene interventions rarely achieve sustained large-scale behavior change and improved public health (Luby, Agboatwalla, et al. 2009, Olembo et al. 2004, Stockman et al. 2007). Thus, both handwashing with soap (HWWS) and treating drinking water remain uncommon among the world's poor (Curtis, Danquah, and Aunger 2009, Rosa and Clasen 2010), and willingness to pay for water treatment is low in many contexts (Ahuja, Kremer, and Zwane 2010, Luoto et al. 2011, Berry, Fischer, and Guiteras 2015).

The lack of success of large interventions based on increasing information and reducing financial costs suggests that the neoclassical economic model may have limitations in explaining health-related

[1] Continuous service is required to maintain positive water pressure. If service is interrupted, pipes are not pressurized, which allows contaminants from the environment to seep in and pollute the water (Kumpel and Nelson 2013).
[2] For an example of the latter, see Duflo, Greenstone, Guiteras and Clasen (2015).
[3] See, for example, Freeman III (2003) or, in a developing country context, Pattanayak and Pfaff (2009). The Health Belief Model of public health has similar assumptions and conclusions (Janz and Becker 1984).

behavior. Research in other fields suggests that other motivations, not typically considered in a neoclassical model, are potentially important.[4]

First, psychology and anthropology suggest that disgust can be a powerful motivator for avoiding environmental health risks (Oaten, Stevenson, and Case 2009, Curtis, Barra, and Aunger 2011). For example, a field experiment in Australia found that a disgust-based treatment to encourage handwashing was substantially more effective at increasing rates of handwashing than health education (Porzig-Drummond et al. 2009). A recent review of 11 developing country studies conducting interviews with over 4,000 primary caregivers of children found that subjects consistently reported disgust at the possibility of feces remaining on hands as the most powerful motivator of promoting handwashing with soap (HWWS) after defecation. In contrast, fear of disease was not generally reported as a motivating factor (Curtis, Danquah, and Aunger 2009). A study of a television and radio campaign in Ghana that emphasized disgust at contamination of hands found, in a pre-post design, a 13 percentage point (pp) increase in self-reported HWWS after defecation and a 41 pp increase in self-reported HWWS before eating (Scott et al. 2008). A study of a social marketing campaign in Burkina Faso that included disgust found, again in a pre-post design, a 16 pp increase in observed HWWS after defecation and an 18 pp increase in HWWS after handling a child's feces (Curtis et al. 2001).

Second, evidence suggests that people are often concerned that their failure to engage in socially normative behaviors will be observed by others, and change their behavior when they are observed. For example, until recently most medical doctors in developed-country hospitals did not wash their hands between patients (Boyce 1999). Presumably, neither knowledge nor monetary cost was an impediment, because doctors know how germs are spread, and soap and handwashing facilities are free and easily available in hospitals. However, Pittet et al. (2004) found that rates of handwashing with soap between patients increased from 44% to 61% when doctors in a Swiss hospital knew someone was watching them, while an education intervention had no effect. Similarly, Munger and Harris (1989) found that only 39% of visitors to a public restroom at a U.S. college washed their hands when they believed they were alone, but 77% did when they knew someone was watching. We label this motivation as "shame."

[4] In this Introduction, we provide a brief, intuitive framework for the economic ideas that motivated our intervention. We provide a simple formal model in Section 1 of the Online Appendix.

In economic terms, disgust and shame may increase the private disutility of failing to engage in a positive health behavior, and thereby increase the behavior. One useful feature is that both disgust and shame operate quickly, and so may be less subject to present bias, a likely source of under-investment in health (Kessler and Zhang 2015). Shame can be seen as a tool for countering the negative externality of failing to take action to prevent communicable disease – Pigovian taxation is usually not feasible in a local, informal community, but sanctions against those observed to violate norms can act as a tax (Habyarimana and Jack 2011).

This paper studies handwashing behavior and the use of and willingness to pay for chlorine-based water treatment among over 2,000 low-income households living in over 400 compounds (groups of 4-18 households sharing a water source and latrine) in slums of Dhaka, Bangladesh. We randomized two treatments. First, we varied the promotional message. One intervention focused on disgust at the consumption of human feces and shame at being seen consuming human feces by one's neighbors, while the other used a standard, high-quality message focused on health benefits. Second, we added a specific handwashing component to the messaging treatment and provided a "soapy water bottle," a sturdy plastic bottle in which packets of powdered soap (also provided) and tap water are mixed to provide convenient, inexpensive access to soap and water for handwashing.

We allocated compounds to treatment arms using the optimal sequential method of Atkinson (1982). To our knowledge, this is the first application of this method in any field (McEntegart 2003, Ciolino et al. 2011). To reduce courtesy bias, we analyze impacts on: (1) presence of soap and water near the latrine, an important proxy for handwashing with soap after defecation; (2) observed handwashing with soap using low-salience methods; (3) free residual chlorine in household drinking water. To estimate the impact on demand, we developed a variant of the Becker-Degroot-Marschak (BDM) method (Becker, Degroot, and Marschak 1964) to measure each compound's collective willingness to contribute to a shared chlorine dispenser.

Neither messaging treatment achieved high levels of water chlorination, nor substantial willingness to pay for water treatment hardware. Two months after our promotion, during an extended free trial, 11.8% of households in the disgust treatment had detectable levels of chlorine in their drinking water, versus 8.5% in the standard treatment (estimated difference 3.4 pp, $p<0.10$) and essentially zero in the non-study population. This difference between treatments vanished by the end of the four-month free trial, with chlorination rates of approximately 8% in both arms. Mean willingness to pay was not significantly higher in the disgust treatment than in the standard treatment (USD 0.58 vs USD 0.52 per compound per month, not statistically significant).

The handwashing intervention did increase rates of handwashing. Two months after the intervention began, the rate of handwashing with soap after toilet visits was 15-17% in compounds that received the handwashing treatment, versus 10-11% in compounds that did not (estimated difference 4.8-6.0 pp, $p<0.01$). While this effect size is modest, the intervention was extremely cheap, costing approximately USD 0.75 per household per year. The handwashing intervention also significantly increased the availability of soap and water: at the end of the free trial, soap and water were present at the latrine in 58.5% of compounds that received the handwashing treatment, versus 15.9% of compounds that did not (estimated difference 42.6 pp, $p<0.01$). Using quasi-random variation in whether compounds continued to receive free powdered soap after the free trial of the dispenser, we find ending this free supply eliminates roughly one-third of the gains in the availability of soap and water, even though the costs of powdered soap were extremely low – approximately USD 0.05 per household per month. There were no detectable differences in handwashing rates or availability of soap and water between the two messaging arms.

Follow-up surveys and qualitative interviews with compound residents identified several main barriers to water treatment. The first barrier was dislike for the taste and smell of chlorine. Second, many households stated that they disliked sharing hardware with other households in the compound. Third, the effectiveness of the shame component of the disgust and shame treatment was limited by the fact that there were not strong ties within these compounds. Although water treatment decisions were typically visible to neighbors, most residents placed little importance on their neighbors' opinions. Also, our messaging was not repeated often and was disseminated during the day, when most men were at work outside the compound. Thus, we estimate our messages reached fewer than 20% of adult men.

2 Experimental Design and Intervention

To test the relative effectiveness of a combined disgust and shame message versus a traditional health message at increasing handwashing and use of and willingness to pay for chlorine-based water treatment, we conducted a randomized trial among 434 compounds in slums of Dhaka, Bangladesh. In all compounds, we conducted a promotional meeting and provided the compound with a 4-month free trial of a chlorine dispenser, a device that allowed households to treat their drinking water in a convenient and safe manner (Jameel Poverty Action Lab 2012). We randomized the promotional message, allocating half of compounds to a standard health message and half to a message targeting disgust and shame, both emphasizing water treatment. We orthogonally randomized two-thirds of treatment compounds to receive additional messaging on handwashing, and provided a simple, inexpensive "soapy water bottle" to facilitate washing hands with soap (Hulland et al. 2013). Finally, compounds were randomized between a group auction, in which we asked the compound for its collective willingness to pay, and a "weakest link" auction intended to uncover household-level willingness to pay.

We describe the context and the selection of the sample in Section 2.1, the treatments in Section 2.2, and the treatment assignment procedure in Section 2.3. A timeline of the key intervention activities is provided in the Online Appendix. Full scripts for the promotional activities and the sales exercises are provided in the Supplementary Materials.

2.1 Context and Sample Selection

This study was conducted in compounds in slums of Dhaka, Bangladesh. Compounds are clusters of households, typically located around a small courtyard, sharing a common toilet, water source and cooking facilities. This setting was chosen for the following reasons:

- Poor water quality and high incidence of water-borne disease (diarrhea, cholera);
- Water collection (and treatment / non-treatment) and post-toilet handwashing were easily observed by other compound residents, making social norms a potential tool for behavior change;
- Experience piloting the chlorine dispenser in areas of Dhaka with similar demographics;
- The implementing organization, the International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b), was well-known and well-regarded in the area, primarily because of its free cholera hospital, facilitating access to compounds.

The setting had the following disadvantages:

- The population had little familiarity with chlorinated water, making distaste a potential barrier to adoption;

- Natural gas was priced at zero marginal cost in most areas, providing a heavy subsidy to boiling, the most common mode of water treatment.

- The population was highly mobile and most compounds contained migrants from a variety of rural areas, potentially social pressure.

We identified five communities as best suited for the study, on the basis of poor water quality, high levels of water-borne disease, and high population density: Mohammedpur, Mirpur, Badda, Khilgaon and Bashabo. Within these communities, we selected a sample of 434 compounds meeting the following criteria: 4-18 households per compound; shared toilet(s), water source and cooking facilities; adequate space to hold promotional activities; and no other water or hygiene interventions occurring at the time. Further details on the sample selection process are provided in Section 2 of the Online Appendix.

2.2 Treatments

2.2.1 Chlorine dispenser

We provided all compounds with a four-month free trial of a chlorine dispenser, an easy-to-use appliance that supplies a specific dose of dilute sodium hypochlorite. We mounted the dispenser near the compound's shared water source, usually a hand pump. Households in these communities typically collect drinking water from the common tap using their own vessels and then store the water in their private room.

This dispenser was developed in rural western Kenya (Jameel Poverty Action Lab 2012). In a randomized trial, over half of households in intervention communities were chlorinating their water six months after the intervention, with over 40% continuing to chlorinate 2.5-3 years after the intervention (Kremer et al. 2011). This intervention was estimated to cost approximately USD 0.50 per person per year. As part of a separate study in Dhaka, the Kenya design was piloted extensively. Icddr,b made several design changes to adapt to the local context (Arman et al. 2013).

In addition to the dispenser, the free trial included two 15 liter reservoirs that the compound could use to have treated water available at all times, and two plastic stools to help with dosing private vessels and the shared reservoirs. All compounds received detailed demonstrations on use of the dispenser. During the free trial, all compounds also had free visits for maintenance and to refill the chlorine.

After the free trial ended, compounds that purchased the subscription received monthly visits for maintenance, refilling and fee collection.

2.2.2 Messaging

We randomized compounds to one of two behavior change messages, one based on standard health messages and the other targeting feelings of disgust and shame. There were three promotional meetings at each compound, the first lasting approximately two hours and two for reinforcement and follow-up, each lasting approximately one hour.

The standard treatment meeting was modeled after existing water treatment interventions used by icddr,b. An icddr,b field intervention specialist (FIS) gave a presentation accompanied by flip charts. The content of the presentation included explaining how germs can enter our bodies via untreated water, how they can cause illness or death, especially among children, and how these risks can be reduced by water treatment. The messaging emphasized the role of externalities, explaining that all members of the compound should treat water not just for their own family's health but also to improve the health of others.

The disgust and shame meetings contained similar explanations and demonstrations of contamination mechanisms and risk, but emphasized the presence of fecal matter in or on contaminated objects, and neighbors' role in spreading fecal matter to their families. To provoke a strong reaction from participants, the FIS used the Bengali word *gu*, roughly equivalent to "shit," rather than the more polite, clinical *paykhana* ("feces"). They communicated that we sometimes unknowingly serve *gu* to our family by not washing hands with soap or by not treating our drinking water. They emphasized how *gu* can spread between people, especially neighbors, to encourage people to care about others' behaviors and what others think of their behaviors. To demonstrate that even clear water from the hand pump can be contaminated, the FIS used a custom "disgust box" to show how *gu* gets into drinking water through leaky pipes.[5] As in the standard messaging treatment, the FIS discussed externalities, but with a different emphasis: if your neighbors do not treat their water, they are feeding *gu* not just to themselves and their children, but making it more likely that you and your children will eat *gu*. Residents were encouraged to monitor each other's behavior and speak up when they observed dangerous, disgusting acts. The messaging also emphasized social interactions: if your neighbors see

[5] More information on the disgust box, including a photo, a description of its use and a link to a video of its use in a promotional meeting, is provided in Section 4 of the Online Appendix.

you failing to treat your water, they will know that you feed *gu* to yourself and your family, which is a shameful act.

The disgust and shame presentation was developed with extensive pilot testing. While the message was occasionally disturbing, only one person left a presentation because it was too upsetting. Qualitative evidence collected during the design stage suggested that the presentation was effective in conveying its message among those who attended, at least in the short run (Rahman et al. 2013).

2.2.3 Handwashing Intervention

The presence of soap at a latrine is strongly positively correlated with handwashing after defecation or contact with a child's feces (Luby, Halder, et al. 2009, Hoque 2003, Baker et al. 2014). However, Bangladeshi households often cite the inconvenience or expense of soap as a barrier to handwashing, and shared latrines make it especially challenging to keep soap near the latrine (*ibid.*).

To study the effect of providing inexpensive, convenient hardware on the presence of soap and on handwashing, we randomly assigned 2/3 of compounds to a handwashing treatment, in which we provided a sturdy plastic bottle and two small packets of detergent per household (approximately one month's supply). We demonstrated how to mix water and detergent in the soapy water bottle and how to use discarded plastic bottles to create additional or replacement units (Hulland et al. 2013). We resupplied detergent periodically during the free trial period.

In compounds assigned to the handwashing arm, we included handwashing messages in the behavior change activities. In the arm receiving the standard health messages, the messaging focused on transmission of germs by unwashed hands and how washing with soap would reduce risks of illness. In the arm receiving the disgust and shame messages, we emphasized that failure to wash hands led to eating *gu* and causing neighbors to eat *gu*. We used Glo Germ, a powder that requires soap to remove and is visible only under ultraviolet light, to demonstrate how invisible dirt or *gu* can remain on hands when not washed with soap, and to show how easily *gu* was spread in the compound.

2.2.4 Auction

At the end of the four-month free trial, we measured willingness to pay (WTP) for a year's subscription to the chlorine dispenser, including use of the hardware, maintenance, and monthly refills of chlorine. In the compounds assigned to the handwashing treatment, the subscription included monthly resupply of detergent packets. To measure the compound's collective WTP, we developed a variant of the Becker-Degroot-Marschak (BDM) mechanism.

The standard use of BDM is to measure individual WTP for a private good (Shogren 2005). The subject states her "bid", B, the maximum amount she is willing to pay for an item. She then draws a random number, in our context choosing one envelope from an unmarked set. This "draw", D, is then compared against the bid B. If the draw is greater than the subject's bid ($D > B$), the subject cannot purchase the item. If the bid is at least as high as the draw ($D \leq B$), then the subject purchases the item and pays D. For expected utility maximizers, the subject's optimal strategy is to bid her true maximum willingness to pay, i.e. $B^* = WTP$ (Becker, Degroot, and Marschak 1964, Horowitz 2006). While BDM has primarily been used in laboratory or laboratory-like environments, a few recent papers have used BDM in field settings in developing countries to measure households' preferences (Hoffmann 2009, Luoto et al. 2012, Cole, Gine, and Vickery 2013, Guiteras and Jack 2014, Berry, Fischer, and Guiteras 2015).

Compound-level WTP depends on two components: first, the individual WTPs of the constituent households; second, the compound's ability to coordinate and cooperate. As is typical with public goods, information asymmetries and desire for free-riding likely mean that the outcome of the bargaining process within the compound will result in a collective WTP that is less than the sum of the households' individual WTPs. Our intervention could alter compound-level WTP by increasing the WTP of individual households, by improving coordination, or both. In collective BDM, we seek only to measure the total effect of these two processes. That is, we assume the compound's bidding behavior in BDM truthfully reveals the compound's collective WTP that arises from its within-compound bargaining process.

Our tests of how our experimental interventions affect WTP require a weaker assumption: even if a compound's BDM response does not equal its compound-level collective WTP, such deviations are uncorrelated with our randomized treatments. That is, if bias is similar across treatment arms, comparing BDM responses still provides useful information about relative effects.

Approximately two weeks before the sale, the FIS convened compound residents for coaching. The FIS reminded the compound that the free trial was coming to an end, and that they would have an opportunity to purchase a year's subscription to use of the hardware, refills and maintenance, with an option to renew for a second year after the first year ended. The FIS explained the auction mechanism in detail, conducted several role-play demonstrations designed to show that truthful revelation was their best strategy, and held a real-money practice round for a token item (a shared bag of laundry detergent). The FIS told the compound that they would be responsible for deciding how much collectively they were willing to pay and how payment would be divided among households, and asked them to designate a spokesperson who would be the official speaker for the compound at the sale.

At the sales meeting, the FIS reviewed the auction procedure with the households in attendance. Compound members were informed that possible prices in the envelope were 25, 50, 75, 100, 150, 200 and 250 taka per month (between USD 0.30 and USD 2.10, or roughly USD 0.03 and USD 0.21 per household per month).[6] The FIS went over each possible price and asked the group whether they would subscribe if that price were drawn. At each price that the spokesperson indicated the compound was willing to pay, the FIS asked the households present for a general voice confirmation that the compound understood that they were committing to pay that amount each month for a year, and that they were prepared to meet that obligation. At each price that the spokesperson indicated that the compound would not pay, the FIS asked households to confirm that they did not wish to pay the indicated price, that they understood that if that price were drawn, they would not be permitted to change their decision, and that they would not regret this decision after the price was revealed.

We also developed a variant of BDM, which we call "weakest link BDM," as an attempt to measure individual households' willingness to contribute to the compound-level public good. This variant, assigned to 217 compounds, was not successful, in that it did not provide useful information on household-level WTP. We discuss this attempt in a separate note (Guiteras et al. 2015). We do not analyze WTP data from these compounds, but they are retained in other analyses.

2.3 Treatment Assignment

Our design was a 2-by-2-by-2 interaction of

- behavior change message: disgust and shame vs. standard health;
- handwashing message and soapy water bottle;

[6] The range of prices was chosen to cover the full range of non-zero WTP observed in piloting. The number of prices was chosen to balance resolution against simplicity of exposition and implementation. There is evidence indicating that the range and distribution of prices can influence WTP (Mazar, Koszegi, and Ariely 2010, Urbancic 2011). Given our sample size and implementation capacity, it was not feasible to test for these effects with any useful precision, so we chose the simplest, most transparent design that provided satisfactory detail on the distribution of WTP.

- BDM type: collective vs. individual.

This design created 8 cells. Because we were especially interested in the effect of the disgust and shame treatment on handwashing in the handwashing arm, we gave handwashing 2/3 weight. The other treatments were equally weighted. We stratified on compound size (number of households) and the presence of gas burners connected to the municipal supply. Compound size was chosen because social dynamics could depend on compound size. Gas was chosen because boiling with gas was a low-cost alternative to chlorination.[7] Operational constraints required us to collect stratification variables and assign treatments at the time of enrollment, so we used the D_A-optimal sequential allocation method first proposed in Atkinson (1982), which optimally balances treatments and stratification covariates given the constraint of not knowing *ex ante* the characteristics of the entire sample. To our knowledge, this study is the first to apply this method.[8]

We provide descriptive statistics for the sample and balancing tests in Tables A1-A4 in the Appendix. Table A1 presents descriptive statistics and p-values from a test of the joint significance of all 8 treatment categories, while Tables A2-A4 provide pairwise comparisons for the messaging, handwashing and auction type treatments, respectively.

3 Data

In this section, we describe the measures used for our key outcome variables: water chlorination, availability of soap and water at the latrine, and handwashing with soap after visiting the latrine. Data on willingness to pay were collected as part of the BDM auction during the sales meeting, as described in Section 2.2.4. A timeline is provided in the Online Appendix.

3.1 Water treatment

We measured water treatment by testing for chlorine in household drinking water. We made two unannounced visits to each compound during the free trial, approximately 2 months and 3.5 months after the promotional meeting. We requested a sample of household drinking water from each of 6

---

[7] This choice was based on our piloting, which was conducted in areas that were physically close to our study area and generally similar in terms of most characteristics. It turned out that gas coverage nearly universal in our study area, so *ex post* we learned that stratifying on gas was a poor choice. However, this did not lead to imbalance on the other stratification variable, nor on the other observables (see Tables A1-A4 in the Appendix).

[8] See two surveys of the clinical trials literature (McEntegart 2003, Ciolino et al. 2011); also confirmed via personal communication with J. Ciolino, Northwestern University, January 17, 2014. We provide further exposition of the method and details on field implementation in Section 3 of the Online Appendix and in Guiteras, Levine, and Polley (2015). Stata code is provided at http://www.econ.umd.edu/research/papers/617.

households in the compound.[9] We then tested the samples for the presence of free residual chlorine using the Hach color wheel.[10] We code a sample as positive if any residual chlorine was detected.

Free residual chlorine was our preferred measure of water treatment for several reasons. First, survey responses on water treatment practices are subject to courtesy or social desirability bias. Second, chlorine testing has almost no false positives: the test is almost never positive unless there is chlorine in the water and there was essentially zero chlorination of drinking water in this population other than in our study.[11] Third, direct measures of contamination are either prohibitively expensive (*E. coli* counts)[12], or have high rates of false positives (H2S testing). The primary disadvantage of chlorine testing is that chlorine residual declines over time and is non-detectable roughly 24 hours after treatment. This may lead to false negatives: although the household did treat their water, the chlorine test is negative because the water no longer contains detectable amounts of free residual chlorine.[13]

3.2 Availability of soap and water at compound latrine

At baseline (2 weeks before the first promotional meeting), midline (3.5 months after the first promotional meeting), and endline (7 months after), an enumerator checked whether soap and water were available at the compound's common latrine. These data were collected at the beginning of these unannounced visits to avoid bias due to households placing soap at the latrine because of the arrival of an observer. A compound was recorded as positive if (a) a soapy water bottle was present, contained water, and the enumerator could detect soap mixed into the water or (b) soap and water were otherwise available.

3.3 Handwashing

As with water treatment, survey methods for measuring handwashing are subject to both courtesy and social desirability bias (Stanton et al. 1987, Biran et al. 2008). To reduce these biases, we measured

---

[9] See Section 2 of the Online Appendix for details on the selection process.

[10] More precisely, the N,N-diethyl-p-phenylenediamine colorimetric method, Hach Company, Loveland, CO.

[11] Another icddr,b study in similar neighborhoods in Dhaka carried tested 1,264 homes' stored water for chlorine from June to December 2012. Exactly zero had detectable chlorine residual (personal communication with Dr. Shwapon Kumar Biswas, iccrd,b, January 23, 2013).

[12] In a study of similar neighborhoods in Dhaka, Luoto et al. (2011) find that the presence of chlorine residual is strongly negatively correlated with *E. coli* counts.

[13] We cannot quantify this error rate but we believe it is low: the most common time for households to fill their private vessels is in the early morning, so it is unlikely that we were taking samples of water that had been collected more than a few hours earlier. Furthermore, this type of false negative is not so false in practical terms: the protective efficacy of chlorine declines over time, so water that was chlorinated more than 24 hours earlier is more likely to be recontaminated; both messages emphasized the importance of drinking water within 24 hours of treatment, so treatment more than 24 hours prior is only weak adherence to the desired behavior.

handwashing using structured observation, in which enumerators would observe handwashing behaviors during a 5-hour session in each compound. The enumerator stayed in the compound from 7:00 am to 12:00 pm.

To reduce the influence of the presence of an observer (Ram et al. 2010), the enumerator stated she was there to observe daily household activities, without mentioning handwashing specifically, and visited at a busy time with many residents and non-residents coming and going. Using a pre-tested instrument (Luby et al. 2011), field workers noted handwashing behavior at key times: before preparing food, feeding a child or eating, after visiting the latrine, and after cleaning the anus of a child. They observed all available household members and noted whether they used water, whether they used soap and whether they washed both hands.

4 Results

In this section, we describe the effect of our interventions on: water treatment, as measured by detectable chlorine residual; availability of soap and water at the toilet; observed handwashing behavior; and willingness to pay for the continued use of the dispenser for one year. In all cases, standard errors of estimates using household-level data are clustered by compound. Unless otherwise noted, estimates are intention to treat: while some compounds dropped out of the intervention during the free trial, they were surveyed and remain in the sample.

4.1 Water treatment

Our preferred measure of water treatment, as discussed in Section 3.1, is detectable chlorine residual in household drinking water, sampled from six households per compound 2 months and 3.5 months after the promotional meeting. At both times, the free trial was ongoing.

We estimate the effect of the messaging treatment on water treatment using logit regression, estimating

$$P\left(y_{hct}=1 \middle| T_{ct}\right)=\Lambda\left(\beta_{0,t}+\beta_{1,t}T_{c}\right), \tag{1}$$

where y_{hct} indicates detectable chlorine in the sample taken from household *h* in compound *c* at time *t*, and T_c represents the treatment status of compound *c*. The positive health treatment is the excluded category, so $T_c=1$ indicates that compound *c* was assigned to the disgust and shame treatment.

Predicted probabilities and treatment effects (discrete differences) are provided in Table 1, and predicted probabilities with 95% confidence intervals are plotted in Figure 1. Chlorination rates are low in both arms, with between 8 and 12 percent of samples testing positive. As noted in Section 3.1, the baseline rate is approximately zero in this population. The disgust and shame treatment does not have a large differential impact relative to the standard health treatment: there is a small and marginally significant difference (3.4 pp, $p<0.10$) at the 2-month measurement, but this difference is not maintained at 3.5 months. Table S1 and Figure S1 in the Online Appendix show that short-run chlorination rates are higher overall in smaller compounds (8 or fewer households), but the short-run differential effect of the disgust and shame treatment is concentrated in larger compounds (more than 8 households). In both large and small compounds, chlorination rates fall between the 2-month and 3.5-month surveys, and there is no differential effect of the disgust treatment in either subgroup in the 3.5-month survey.

4.2 Availability of soap and water at compound latrine

We measure the effect of the handwashing treatment on the presence of soap and water at the communal latrine using logit regression, estimating

$$P\left(y_{ct}=1\middle|T_c\right)=\Lambda\left(\beta_{0,t}+\beta_{1,t}T_c\right), \tag{2}$$

where y_{ct} indicates the presence of soap and water at the latrine of compound *c* at time *t* and T_c represents the treatment status of compound *c*. The non-handwashing treatment is the excluded category so $T_c=1$ indicates that the compound was assigned to the handwashing treatment. Standard errors are clustered by compound.

Predicted probabilities and treatment effects (discrete differences and difference-in-differences) are provided in Table 2, and predicted probabilities with 95% confidence intervals are plotted in Figure 2. At midline (3.5 months), the handwashing treatment has a practically large and statistically significant effect on the availability of soap and water: 42.6 pp ($p<0.01$) in a raw comparison and 54.0 pp ($p<0.01$) using the difference in differences estimate.[14,15]

[14] As a robustness check, we estimated the model with a balanced panel, restricting the sample to compounds observed in all three survey rounds. The results, presented in Table S2 of the Online Appendix, are generally similar.

[15] We are unable to explain the statistically significant baseline difference (11.4 pp, $p<0.05$) between the two groups. The survey procedures were the same across rounds and neither the surveyors nor the compound residents knew the compound's assignment status at the time of the baseline. This baseline difference increases the difference-in-differences

Roughly 2 weeks after the midline survey, compounds participating in the BDM auction either won, in which case free detergent continued to be provided monthly, or lost, in which case no further detergent was provided, although the compound retained the soapy water bottle itself. This provides a quasi-experiment to estimate the importance of free provision of soap. While there is some randomness in BDM, it is not a pure randomization, because winning compounds on average had higher WTP than losing compounds. Therefore, we compare midline (3.5-month) and endline (7-month, after free soap ended) outcomes within the winning and losing groups in the handwashing arm, estimating

$$\begin{aligned} P\left(y_{ct}=1\middle|t,W_c\right) &= \Lambda\left(\beta_{0,\mathrm{W}}+\beta_{1,W}\mathrm{EndLine}\right) \\ P\left(y_{ct}=1\middle|t,L_c\right) &= \Lambda\left(\beta_{0,L}+\beta_{1,L}\mathrm{EndLine}\right) \end{aligned} \tag{3}$$

for winning compounds W_c and losing compounds L_c. The omitted category is the response at midline, so $\beta_{1,L}$ represents the change in availability between midline and endline for losing compounds. As shown in Table 3 and plotted in Figure 3, the probability that soap and water would be present at the latrine in compounds that lost the auction fell from midline to endline by 15.3 pp ($p<0.05$, 51.0% to 35.7%), as compared to a small and statistically insignificant reduction of 0.6 pp among winning compounds (63.1% to 62.5%). This decline represents approximately 1/3 of the 42.6-54.0 pp gain between baseline and midline, which suggests some habit formation but also strong sensitivity to a very low price.[16] The losing compounds retained the plastic bottle, and any discarded water bottle could serve as a replacement. Detergent packets are commonly available and extremely cheap: 2 packets costing approximately USD 0.05 total will last a household at least one month. Therefore, the difference between winner and loser compounds can probably be attributed to the very

---

estimate relative to a simple difference between treatment and control. However, the effect of the handwashing treatment is still large and statistically significant in the simple difference. As shown in Table S3 and Figure S2 of the Online Appendix, this baseline difference is concentrated in larger compounds (more than 8 households), where compounds assigned to the handwashing treatment were 21.6 pp less likely to have soap and water present at baseline, versus a difference of just 3.2 pp (with the opposite sign) among smaller compounds (8 or fewer compounds). The simple differences estimates are similar between small and large compounds, while the difference-in-differences adjustment leads to a large increase in the estimated effect among large compounds. In all cases (simple differences vs. difference-in-differences; small vs. large compounds), the estimated effect is statistically significant and practically important.

[16] This effect is generally similar between small (-13.9 pp) and large (-15.9 pp) compounds, as shown in Tables S4 and Figure S3 in the Online Appendix, although the smaller sample size within subgroups makes these effects no longer statistically significant.

small and inexpensive nudge of delivering detergent once per month or to the transaction costs of tracking and enforcing whose turn it was to contribute a tiny amount of soap.[17]

To measure the effect of the messaging treatment on the presence of soap and water at the latrine, we re-estimated Equation (2), but with T_c representing the messaging treatment ($T_c = 1$ indicates the disgust and shame treatment). The messaging treatment does not have a large or statistically significant effect, either in simple differences or in difference-in-differences (Table 4, col. 2 and 3).[18] The same null result is found when estimating the model separately for small and large compounds (Table S5.A-S5.B) and for compounds with and without the handwashing treatment (Table S6).

4.3 Handwashing behavior

To determine whether the interventions affected handwashing behavior, we collected data on handwashing behavior during structured observation sessions approximately 2 months after the promotional meeting and start of the free trial. The key outcome of interest is handwashing after toilet visits. We examine three variables, in increasing order of effective hygienic practice: whether the subject rinses her hands with water (whether or not she uses soap); whether she washes at least one hand with soap; and whether she washes both hands with soap. We estimate Equation (2), although we have only one cross-section so we cannot perform difference-in-differences estimation.

The handwashing treatment increased handwashing with soap by 5-6 pp, relative to control group levels of 10-11 percent (Table 5, columns (2) and (3)). There was no effect on the use of water (with or without soap, column (1)), suggesting that the mechanism is the result of easier access to soap for those who choose to wash their hands, rather than an increased propensity to wash hands among those who otherwise would not. Effects were somewhat larger in small compounds (+7-8 pp, $p < 0.01$) than in large compounds (+3-4 pp, marginally significant), as shown in Tables S7.A-S7.B in the Online Appendix. There was no detectable effect of the messaging treatment on handwashing (Table 6; Tables S8.A-S8.B).

Our survey results shed light on the modest effects of the disgust-and-shame method. First, although our qualitative results suggest the shame and disgust intervention increased self-reported feelings that

[17] There was no explicit handwashing promotion during delivery, although it is possible that the visit itself motivated handwashing in some way, perhaps by reminding the household of its importance.

[18] This is an intention-to-treat analysis (i.e, we do not condition on winning in the midline and endline data) because, as seen in Section 4.4.1, there was no detectable impact of the messaging treatment on WTP.

untreated water is disgusting in the short term, the effects were not sustained. For example, 78% of those with the disgust-and-shame message vs. 73% of those with the health messages self-reported feelings of disgust at untreated water after 3.5 months, converging to 70 vs. 71% at 7 months (neither difference statistically significant at the 5% level). In addition, these compounds had high mobility. Perhaps for that reason, in qualitative research, most respondents reported low concern for how their neighbors viewed them (Rahman et al. 2013). Thus, the preconditions for the effectiveness of the disgust-and-shame intervention were not satisfied in this setting.

4.4 Willingness to pay

4.4.1 WTP by messaging treatment

We can assess the impact of the messaging treatment on collective WTP visually, by comparing the inverse demand curves by treatment, and by regression. Figure 4 plots inverse demand curves by messaging treatment. Specifically, for each price $p \in \{25, 50, \ldots, 250\}$, we estimate

$$1\{WTP_c \geq p\} = \Lambda\left(\beta_0 + \beta_1 T_c\right), \tag{4}$$

where WTP_c is WTP for compound *c* and T_c is compound *c*'s treatment assignment. We also report results for WTP per household, dividing by the number of households in the compound at baseline. The standard health message is the omitted category, so β_1 represents the effect of the disgust and shame treatment relative to the standard health treatment. We then use the estimates to compute predicted shares.

The share of compounds willing to purchase at any given price is slightly higher among compounds assigned to the disgust and shame message than compounds assigned to the standard health message, but this difference is not statistically significant at any price. As an aggregate measure, we obtain the effect on mean WTP by estimating

$$WTP_c = \beta_0 + \beta_1 T + \varepsilon_c \tag{5}$$

by OLS. The results, presented in Table 7, are similar to those shown visually in Figure 4. Mean WTP is low in both treatment groups: no more than 5 BDT (USD 0.14) per household per month in either arm. WTP is slightly higher in the disgust arm (by 0.7 BDT, USD 0.01, per household per month), but this difference is neither economically nor statistically significant. Similar null effects are found in both large and small compounds (Tables S9.A-S9.B in the Online Appendix).

4.4.2 WTP by handwashing treatment

We did not have a strong *ex ante* belief that the handwashing treatment would increase WTP for the chlorine dispenser. However, during implementation the field staff reported that the households liked the soapy water bottle and that it increased interest in the intervention generally. As a result, we decided *ex post* to test whether there was any effect on WTP. As shown in Table 8 and Figure 5, there was no systematic increase in WTP. A slightly higher share of handwashing compounds were willing to participate in the auction (0.62 vs 0.58), but this increase was not statistically significant. Similar null effects are found in both large and small compounds (Tables S10.A-S10.B in the Online Appendix).

4.4.3 WTP and payment compliance

The extent to which BDM can successfully measure true willingness to pay remains an open question even in the case of a single agent (Berry, Fischer, and Guiteras 2015), and little is known about whether the compound's collective bid in a group BDM approximates a true underlying parameter. One useful indicator of the compound's sincerity in the bidding process is whether it completes the full year subscription. Using the sample of compounds that won the subscription in the BDM auction, we estimate

$$1\{\text{Comply}_c\} = \Lambda\left(\beta_0 + \beta_1 WTP_c + \beta_2 D_c\right), \tag{6}$$

where $1\{\text{Comply}_c\}$ indicates that compound c made all 12 payments to complete its yearly subscription, WTP_c is the compound's BDM bid, and D_c is the compound's price draw. We control for the price draw because winning compounds that bid more will, on average, also have a higher monthly payment. Table 9 presents logit coefficients and average marginal effects for Equation (6), as well as a specification interacting WTP_c and D_c.[19] While precision is limited by the small sample size, we find a small but positive relationship: a 10 BDT increase in the compound's bid is associated with a 1.2 – 1.4 pp increase in the probability of completing the year's subscription. This positive point estimate provides some evidence that BDM obtained useful information from these compounds.

5 Discussion and Conclusion

When people do not use an effective method to prevent illness, economists typically assume that price is a barrier or that consumers lack information on the problem and/or the effectiveness of this

[19] Estimated marginal effects from linear probability models, presented in Table S11 in the Online Appendix, are similar.

prevention. However, in the case of the chlorine dispenser, providing a free trial of an effective prevention while also teaching people about the presence of dangerous germs in untreated water only increased water treatment modestly.

Mean willingness to pay for the chlorine dispenser after the free trial was low: no more than USD 0.14 per household per month. This low WTP cannot be explained by lack of interest in water treatment: in subsequent research with similar households in the same communities, median household WTP for a ceramic water filter was approximately USD 17 (Guiteras and Jack 2014). This filter was marketed as lasting up to 2 years, implying median household WTP of approximately USD 0.70 per month. Thus, the low WTP for the chlorine dispenser relative to the filter is likely the result of dislike of the dispenser, dislike of chlorine, non-treatment value from the filter (e.g., safe storage or the filter's attractive appearance), and challenges of group coordination in agreeing on collective payments.

We hypothesized that a message combining motives based on disgust and shame might be effective in promoting water treatment. We find little evidence that this message performed better than a traditional health message. However, our disgust-and-shame messages were repeated only a few times and never reached many compound members. Importantly, we reached only perhaps 20% of men – who are typically much more socially powerful than women in Bangladeshi culture. Qualitative interviews with households suggested the presentation evoked feelings of disgust at untreated water, but these feelings did not last. Specifically, the disgust and shame intervention did not increase self-reported feelings of disgust at untreated water seven months later. In addition, residents reported low concern for how their neighbors viewed them, even though household mobility was somewhat lower than expected.[20]

The results from providing a soapy water bottle were modest but encouraging. Two months after the intervention began, people in compounds with the soapy bottle were washing their hands with soap after 15-17% of visits to the toilet, which was more often than the 10-11% share at compounds without that intervention. Because this intervention was so cheap (just USD 0.75 per household per year), even this modest impact is cost-effective by conventional standards.[21] The finding that free

[20] Among households surveyed at baseline, 92% were found at midline, which was typically 3-4 months after the intervention (q10-q90 range 85-112 days). This is an upper bound on household turnover, because some of these households may not have been present that day. Similarly, enumerators found 82% of baseline households at endline, which was typically 6-7 months after the intervention (q10-q90 range 186-220 days).

[21] See Section 4 of the Online Appendix. Our point estimates suggest the chlorine dispenser, in contrast, would not be "highly effective" using the World Bank's criteria.

provision of a simple tool for handwashing substantially increases availability of soap and water and rates of handwashing with soap, while messaging treatments have no differential effect, is consistent with a growing body of research finding that the effects of price and convenience are large relative to other treatments such as promotion or education (Dupas 2011). Furthermore, the fact that the availability of soap and water dropped sharply when we no longer provided free soap shows that even a small positive price can dramatically reduce use (Holla and Kremer 2009).

This paper's results indicate the need for further research in several key areas. First, the low rates of adoption we found contrast with those of Kremer et al. (2011) and other dispenser interventions, primarily in rural Africa, which report sustained chlorination rates of 40% or greater (Evidence Action 2014). It is important to understand why adoption rates of very similar technologies differ so dramatically. This may be the result of subtle differences in the context (in rural Africa, water is typically carried for 20 minutes or more after collection and treatment, which allows the smell of chlorine to dissipate, while in urban Dhaka, water might be consumed very soon after it is collected), or other cultural factors (although chlorination rates are low in rural Africa, households are usually familiar and comfortable with chlorine-based treatment, while there is much less familiarity with chlorine-based treatment in urban Dhaka). Second, it remains an open question whether a more intensive disgust-and-shame based intervention, conducted in communities with stronger social ties, could be more effective in establishing and enforcing new and more healthful norms. Third, the sharp decrease in the availability of soap after monthly delivery of inexpensive detergent packets ceased suggests that more research is needed to improve understanding of both small behavioral barriers and lack of coordination in providing local public goods.

6 References

Acharya, Anjali, and Mikko K. Paunio. 2008. Environmental health and child survival : epidemiology, economics, experiences. The World Bank.

Ahuja, Amrita, Michael Kremer, and Alix Peterson Zwane. 2010. "Providing Safe Water: Evidence from Randomized Evaluations." *Annual Review of Resource Economics* no. 2 (1):237-256. doi: 10.1146/annurev.resource.012809.103919.

Arman, Shaila, Shwapon K. Biswas, Leanne Unicomb, Elli Leontsini, Pavani K. Ram, Fazlul K. Chowdhury, Smriti Roy, Subas C. Biswas, Rouha Anamika Sarker, Farhana Akand, Farzana Begum, Peter J. Winch, and Stephen P. Luby. 2013. Factors affecting use of drinking water chlorination: Experience from a cluster randomized controlled trial in a low income urban Bangladesh.

Atkinson, Anthony C. 1982. "Optimum Biased Coin Designs for Sequential Clinical Trials with Prognostic factors." *Biometrika* no. 69 (1):61-67.

Baker, Kelly K., Fahmida Farzana, Farzana Ferdous, and others. 2014. "Association between Moderate-to-Severe Diarrhea in Young Children in the Global Enteric Multicenter Study (GEMS) and Types of Handwashing Materials Used by Caretakers in Mirzapur, Bangladesh." *The American Journal of Tropical Medicine and Hygiene*:13-0509. doi: 10.4269/ajtmh.13-0509.

Baker, Stephen, Kathryn E. Holt, Archie Clements, and others. 2011. "Combined high-resolution genotyping and geospatial analysis reveals modes of endemic urban typhoid fever transmission." *Open Biology* no. 1:110008. doi: 10.1098/rsob.110008.

Becker, Gordon M., Morris H. Degroot, and Jacob Marschak. 1964. "Measuring utility by a single-response sequential method." *Behavioral Science* no. 9 (3):226-232. doi: 10.1002/bs.3830090304.

Bennett, D., S. Naqvi, and W.P. Schmidt. 2015. Learning, Hygiene, and Traditional Medicine.

Berry, James, Gregory M. Fischer, and Raymond P. Guiteras. 2015. Eliciting and Utilizing Willingness to Pay: Evidence from Field Trials in Northern Ghana.

Biran, A., T. Rabie, W. Schmidt, S. Juvekar, S. Hirve, and V. Curtis. 2008. "Comparing the performance of indicators of hand-washing practices in rural Indian households." *Tropical Medicine & International Health* no. 13 (2):278-285. doi: 10.1111/j.1365-3156.2007.02001.x.

Black, R. E., L. H. Allen, Z. A. Bhutta, L. E. Caulfield, Onis M. de, M. Ezzati, C. Mathers, J. Rivera, and Maternal Child Undernutrition Study Group. 2008. "Maternal and child undernutrition: global and regional exposures and health consequences." *Lancet* no. 371:243-260. doi: 10.1016/S0140-6736(07)61690-0.

Boyce, John M. 1999. "It Is Time for Action: Improving Hand Hygiene in Hospitals." *Annals of Internal Medicine* no. 130 (2):153-155. doi: 10.7326/0003-4819-130-2-199901190-00011.

Ciolino, J., W. Zhao, R. Martin, and Y. Palesch. 2011. "Quantifying the cost in power of ignoring continuous covariate imbalances in clinical trial randomization." *Contemporary clinical trials* no. 32 (2):250-259.

Cole, Shawn, Xavier Gine, and James Vickery. 2013. How does risk management influence production decisions? evidence from a field experiment. The World Bank.

Curtis, V. A., L. O. Danquah, and R. V. Aunger. 2009. "Planned, motivated and habitual hygiene behaviour: an eleven country review." *Health Education Research* no. 24 (4):655-73. doi: 10.1093/her/cyp002.

Curtis, V., B. Kanki, S. Cousens, I. Diallo, A. Kpozehouen, M. Sangare, and M. Nikiema. 2001. "Evidence of behaviour change following a hygiene promotion programme in Burkina Faso." *Bulletin of the World Health Organization* no. 79 (6):518-527.

Curtis, Valerie, Mícheál de Barra, and Robert Aunger. 2011. "Disgust as an adaptive system for disease avoidance behaviour." *Philosophical Transactions of the Royal Society B: Biological Sciences* no. 366:389-401. doi: 10.1098/rstb.2010.0117.

Cutler, David M., and Grant Miller. 2006. "Water, Water Everywhere. Municipal Finance and Water Supply in American Cities." *NBER*:153-184.

Duflo, Esther, Michael Greenstone, Raymond P. Guiteras, and Thomas F. Clasen. 2015. The Short- and Medium-Term Impacts of Household Water Supply and Sanitation on Diarrhea in Rural India. In *Population Working Paper*: Maryland Population Research Center.

Dupas, Pascaline. 2011. "Health Behavior in Developing Countries." *Annual Review of Economics* no. 3 (1):425-449. doi: 10.1146/annurev-economics-111809-125029.

Evidence Action. *Chlorine Dispensers Work for Safer Water: A Review of the Evidence* 2014. Available from http://www.evidenceaction.org/blog-full/chlorineevidence.

Freeman III, A. Myrick. 2003. *The Measurement of Environmental and Resource Values*. 2 ed. Washington, DC: Resources for the Future.

Guiteras, Raymond P., and B. Kelsey Jack. 2014. Incentives, Selection and Productivity in Labor Markets: Evidence from Rural Malawi. National Bureau of Economic Research.

Guiteras, Raymond P., David I. Levine, Kristian Lopez-Vargas, and Thomas H. Polley. 2015. Weakest Link BDM: Measuring Willingness to Pay for Public Goods.

Guiteras, Raymond P., David I. Levine, and Thomas Polley. 2015. The Pursuit of Balance in Sequential Randomized Trials.

Gupta, S. K., M. S. Islam, R. Johnston, P. K. Ram, and Stephen P. Luby. 2008. "The chulli water purifier: acceptability and effectiveness of an innovative strategy for household water treatment in Bangladesh." *The American Journal of Tropical Medicine and Hygiene* no. 78:979-84.

Habyarimana, James, and William Jack. 2011. "Heckle and Chide: Results of a randomized road safety intervention in Kenya." *Journal of Public Economics* no. 95:1438-1446. doi: 10.1016/j.jpubeco.2011.06.008.

Hoffmann, Vivian. 2009. "Intrahousehold allocation of free and purchased mosquito nets." *The American Economic Review* no. 99 (2):236-241. doi: 10.1257/aer.99.2.236.

Holla, Alaka, and Michael Kremer. 2009. "Pricing and Access: Lessons from Randomized Evaluations in Education and Health." In *What Works in Development? Thinking Big and Thinking Small* edited by William Easterly and Jessica Cohen, 91-119. Brookings Institution Press.

Hoque, BA. 2003. "Handwashing practices and challenges in Bangladesh." *International Journal of Environmental Health Research* no. 13:S81-S87. doi: 10.1080/0960312031000102831.

Horowitz, John K. 2006. "The Becker-DeGroot-Marschak mechanism is not necessarily incentive compatible, even for non-random goods." *Economics Letters* no. 93 (1):6-11.

Hulland, Kristyna, Elli Leontsini, Robert Dreibelbis, Leanne Unicomb, Aasma Afroz, Notan Dutta, Fosiul Nizame, Stephen Luby, Pavani Ram, and Peter Winch. 2013. "Designing a handwashing station for infrastructure-restricted communities in Bangladesh using the integrated behavioural model for water, sanitation and hygiene interventions (IBM-WASH)." *BMC Public Health* no. 13 (1):877.

Hunt, Caroline. 2006. Sanitation and Human Development. Human Development Report Office (HDRO), United Nations Development Programme (UNDP).

Jameel Poverty Action Lab. 2013. *Chlorine Dispensers for Safe Water* 2012 [cited January 17 2013]. Available from Chlorine Dispensers for Safe Water.

Janz, Nancy K., and Marshall H. Becker. 1984. "The Health Belief Model: A Decade Later." *Health Education & Behavior* no. 11 (1):1-47. doi: 10.1177/109019818401100101.

Kessler, Judd B., and C. Yiwei Zhang. 2015. "Behavioural Economics and Health." In *Oxford Textbook of Public Health*, edited by Roger Detels, Martin Guillford, Quarraisha Abdool Karim and Chorh Chuan Tan. Oxford University Press.

Kremer, Michael, Edward Miguel, Sendhil Mullainathan, Clair Null, and Alix Zwane. 2011. Social Engineering: Evidence from a Suite of Take-up Experiments in Kenya.

Kumpel, Emily, and Kara L. Nelson. 2013. "Comparing microbial water quality in an intermittent and continuous piped water supply." *Water Research* no. 47:5176-5188. doi: 10.1016/j.watres.2013.05.058.

Lin, Audrie, Benjamin F. Arnold, Sadia Afreen, and others. 2013. "Household Environmental Conditions Are Associated with Enteropathy and Impaired Growth in Rural Bangladesh." *The American Journal of Tropical Medicine and Hygiene* no. 89:130-137. doi: 10.4269/ajtmh.12-0629.

Lozano, Rafael, Mohsen Naghavi, Kyle Foreman, and others. 2012. "Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010." *The Lancet* no. 380:2095-2128. doi: 10.1016/S0140-6736(12)61728-0.

Luby, S. P., M. Agboatwalla, A. Bowen, E. Kenah, Y. Sharker, and R. M. Hoekstra. 2009. "Difficulties in maintaining improved handwashing behavior, Karachi, Pakistan." *Am J Trop Med Hyg* no. 81 (1):140-5.

Luby, Stephen P., Amal K. Halder, Tarique Huda, Leanne Unicomb, and Richard B. Johnston. 2011. "The Effect of Handwashing at Recommended Times with Water Alone and With Soap on Child Diarrhea in Rural Bangladesh: An Observational Study." *PLoS Med* no. 8 (6). doi: 10.1371/journal.pmed.1001052.

Luby, Stephen P., Amal K. Halder, Carole Tronchet, Shamima Akhter, Abbas Bhuiya, and Richard B. Johnston. 2009. "Household Characteristics Associated with Handwashing with Soap in Rural Bangladesh." *The American Journal of Tropical Medicine and Hygiene* no. 81 (5):882-887. doi: 10.4269/ajtmh.2009.09-0031.

Luoto, Jill, Minhaj Mahmud, Jeff Albert, Stephen Luby, Nusrat Najnin, Leanne Unicomb, and David I. Levine. 2012. "Learning to Dislike Safe Water Products: Results from a Randomized Controlled Trial of the Effects of Direct and Peer Experience on Willingness to Pay." *Environmental Science & Technology* no. 46 (11):6244-6251. doi: 10.1021/es2027967.

Luoto, Jill, Nusrat Najnin, Minhaj Mahmud, Jeff Albert, M. Sirajul Islam, Stephen Luby, Leanne Unicomb, and David I. Levine. 2011. "What Point-of-Use Water Treatment Products Do Consumers Use? Evidence from a Randomized Controlled Trial among the Urban Poor in Bangladesh." *PLoS ONE* no. 6 (10). doi: 10.1371/journal.pone.0026132.

Mackenbach, Johan P. 2007. "Sanitation: pragmatism works." *BMJ* no. 334:s17-s17. doi: 10.1136/bmj.39044.508646.94.

Mazar, Nina, Botond Koszegi, and Dan Ariely. 2010. Price-Sensitive Preferences. In *SSRN eLibrary*.

McEntegart, Damian J. 2003. "The pursuit of balance using stratified and dynamic randomization techniques: an overview." *Drug Information Journal* no. 37 (3):293-308.

McKenzie, David, and Isha Ray. 2009. "Urban water supply in India: status, reform options and possible lessons." *Water Policy* no. 11:442-460. doi: 10.2166/wp.2009.056.

Munger, Kristen, and Shelby J. Harris. 1989. "Effects of an observer on hand washing in a public restroom." *Perceptual and Motor Skills* no. 69 (3):733-734. doi: 10.2466/pms.1989.69.3.733.

Oaten, Megan, Richard J. Stevenson, and Trevor I. Case. 2009. "Disgust as a disease-avoidance mechanism." *Psychological Bulletin* no. 135:303-321. doi: 10.1037/a0014823.

Olembo, L., F. Kaona, M. Tuba, and G. Burnham. 2004. Safe water systems: An evaluation of the Zambia CLORIN Program (Final Report) Washington, DC: U.S. Agency for International Development through the Environmental Health Project

Pattanayak, Subhrendu K., and Alexander Pfaff. 2009. "Behavior, Environment, and Health in Developing Countries: Evaluation and Valuation." *Annual Review of Resource Economics* no. 1:183-217. doi: 10.1146/annurev.resource.050708.144053.

Pittet, Didier, Anne Simon, Stéphane Hugonnet, Carmen Lúcia Pessoa-Silva, Valérie Sauvan, and Thomas V. Perneger. 2004. "Hand Hygiene among Physicians: Performance, Beliefs, and Perceptions." *Annals of Internal Medicine* no. 141 (1):1-8. doi: 10.7326/0003-4819-141-1-200407060-00008.

Porzig-Drummond, Renata, Richard Stevenson, Trevor Case, and Megan Oaten. 2009. "Can the emotion of disgust be harnessed to promote hand hygiene? Experimental and field-based tests." *Social Science & Medicine* no. 68:1006-1012. doi: 10.1016/j.socscimed.2009.01.013.

Rahman, Mahfuzur, David I. Levine, Kaniz Khatun-E-Jannat, Robiul A. Chowdhury, and Leanne Unicomb. 2013. Testing Disgust- and Shame-Based Safe Water and Handwashing Promotion in Urban Dhaka. ICDDR,B.

Ram, Pavani K., Amal K. Halder, Stewart P. Granger, and others. 2010. "Is Structured Observation a Valid Technique to Measure Handwashing Behavior? Use of Acceleration Sensors Embedded in Soap to Assess Reactivity to Structured Observation." *The American Journal of Tropical Medicine and Hygiene* no. 83 (5):1070-1076. doi: 10.4269/ajtmh.2010.09-0763.

Rosa, Ghislaine, and Thomas Clasen. 2010. "Estimating the scope of household water treatment in low-and medium-income countries." *The American journal of tropical medicine and hygiene* no. 82 (2):289.

Scott, Beth E., Wolf P. Schmidt, Robert Aunger, Nana Garbrah-Aidoo, and Rasaaque Animashaun. 2008. "Marketing hygiene behaviours: the impact of different communication channels on reported handwashing behaviour of women in Ghana." *Health Education Research* no. 23 (3):392-401.

Shogren, Jason F. 2005. "Experimental Methods and Valuation." In *Handbook of Environmental Economics*, edited by Karl-Gran Mler and Jeffrey R. Vincent, 969-1027. Elsevier.

Stanton, B. F., J. D. Clemens, K. M. A. Aziz, and M. Rahman. 1987. "Twenty-four-hour recall, knowledge-attitude-practice questionnaires, and direct observations of sanitary practices: a comparative study." *Bulletin of the World Health Organization* no. 65 (2):217-222.

Stockman, L. J., T. K. Fischer, M. Deming, B. Ngwira, C. Bowie, N. Cunliffe, J. Bresee, and R. E. Quick. 2007. "Point-of-use water treatment and use among mothers in Malawi." *Emerg Infect Dis* no. 13 (7):1077-80. doi: 10.3201/eid1307.060767.

Taniuchi, Mami, Shihab U. Sobuz, Sharmin Begun, and others. 2013. "Etiology of Diarrhea in Bangladeshi Infants in the First Year of Life Using Molecular Methods." *Journal of Infectious Diseases*:jit507. doi: 10.1093/infdis/jit507.

Urbancic, Michael B. 2011. Testing Distributional Dependence in the Becker-DeGroot-Marschak Mechanism.

Walker, C. L., I. Rudan, L. Liu, H. Nair, E. Theodoratou, Z. A. Bhutta, K. L. O'Brien, H. Campbell, and R. E. Black. 2013. "Global burden of childhood pneumonia and diarrhoea." *Lancet* no. 381 (9875):1405-16. doi: 10.1016/S0140-6736(13)60222-6.

Zwane, Alix Peterson, and Michael Kremer. 2007. "What Works in Fighting Diarrheal Diseases in Developing Countries? A Critical Review." *World Bank Research Observer* no. 22:1-24. doi: 10.1093/wbro/lkm002.

7 Tables

Table 1: Share of households with detectable chlorine residual, by motivational treatment

	(1) 2-month	(2) 3.5 month
Standard message	0.0846	0.0814
	(0.0137)	(0.0111)
Disgust message	0.118	0.0837
	(0.0146)	(0.0119)
Estimated difference (disgust - standard)	0.0337*	0.0023
Std. Err.	(0.0201)	(0.0163)
Number of compounds	417	413
Number of households	2259	2036

Note: this table shows the share of households, by treatment and survey wave, with detectable chlorine in their drinking water, as well as the estimated difference between the disgust and standard treatments. Estimation by logit regression. Estimated discrete differences presented with standard errors clustered at the compound level in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table 2: Effect of handwashing treatment on availability of soap and water

	(1) Baseline	(2) 3.5-month
No handwashing treatment	0.291	0.159
	(0.049)	(0.031)
Handwashing treatment	0.176	0.585
	(0.029)	(0.030)
Estimated difference	-0.114**	0.426***
Std. Err.	(0.057)	(0.043)
Difference in differences		0.540***
Std. Err.		(0.069)
Number of compounds	256	413

Note: this table shows the share of compounds, by handwashing treatment and survey wave, with soap and water available at the common latrine, as well as the estimated difference between treatments. Estimation by logit regression. Standard errors clustered at the compound level in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table 3: Effect of losing free soap delivery on soap and water availability

	(1) Winners	(2) Losers
Midline (3.5 mo.)	0.631	0.510
	(0.060)	(0.051)
Endline (7 mo.)	0.625	0.357
	(0.061)	(0.049)
Estimated difference	-0.006	-0.153**
Std. Err.	(0.085)	(0.070)
Number of compounds	65	98

Note: this table shows the share of compounds with soap and water available at the latrine by BDM outcome and survey wave. Column (1) shows levels for compounds that won the BDM auction at the 3.5-month midline survey, i.e. during the free trial and approximately two weeks before the BDM auction, and the 7-month endline survey, approximately 3 months after the BDM auction, as well as the estimated difference between the midline and endline. Column (2) shows the same for compounds that lost the BDM auction. Compounds that won kept the chlorine dispenser and the soapy water bottle, and continued to receive 2 packets of detergent per household per month for use in the soapy water bottle. Compounds that lost retained the soapy water bottle, but did not receive resupply of detergent. The sample consists of compounds in the handwashing arm and in which an auction was conducted. Estimation by logit regression. Standard errors clustered at the compound level in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table 4: Share of compounds with soap and water available at the latrine, by messaging treatment

	(1) Baseline	(2) 3.5-month	(3) 7-month
Standard	0.177	0.462	0.400
	(0.034)	(0.035)	(0.034)
Disgust	0.254	0.424	0.409
	(0.039)	(0.035)	(0.035)
Estimated difference	0.077	-0.037	0.009
Std. Err.	(0.051)	(0.049)	(0.049)
Difference in differences		-0.114	-0.068
Std. Err.		(0.070)	(0.073)
Number of compounds	256	413	408

Note: this table shows the share of compounds, by messaging treatment and survey wave, with soap and water available at the common latrine, as well as the estimated difference between treatments. Difference-in-difference estimates for the 3.5-month midline (column 2) and 7-month endline (column 3) surveys use differences at baseline for comparison. Estimation by logit regression. Standard errors clustered at the compound level in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table 5: Handwashing after visiting toilet, by handwashing treatment

	(1) Used water	(2) Used soap	(3) Used soap, both hands
No handwashing	0.557	0.114	0.106
	(0.0187)	(0.0116)	(0.0115)
Handwashing	0.549	0.173	0.154
	(0.0136)	(0.00960)	(0.00897)
Estimated difference (HW - no HW)	-0.008	0.060***	0.048***
Std. Err.	(0.023)	(0.015)	(0.015)
Number of compounds	417	417	417
Number of observations	5141	5151	5182

Note: this table shows the share of toilet events after which compound residents (1) rinsed their hands with water (with or without soap), (2) used soap to wash at least one hand, (3) used soap to wash both hands, by handwashing treatment, as well as the estimated difference between the handwashing and non-handwashing treatments. Data collected during structured observation at approximately month 2 of the free trial. Estimation by logit regression. Standard errors clustered at the compound level in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table 6: Handwashing after visiting toilet, by motivational treatment

	(1) Used water	(2) Used soap	(3) Used soap, both hands
Standard	0.555	0.146	0.134
	(0.0166)	(0.0106)	(0.0100)
Disgust	0.548	0.161	0.143
	(0.0144)	(0.0109)	(0.0103)
Estimated difference	-0.006	0.015	0.009
Std. Err.	(0.022)	(0.015)	(0.014)
Number of compounds	417	417	417
Number of observations	5141	5151	5182

Note: this table shows the share of toilet events after which compound residents (1) rinsed their hands with water (with or without soap), (2) used soap to wash at least one hand, (3) used soap to wash both hands, by motivational treatment, as well as the estimated difference between the disgust and standard arms. Data collected during structured observation at approximately month 2 of the free trial. Estimation by logit regression. Standard errors clustered at the compound level in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table 7: Willingness to pay by messaging treatment

	(1) Compound WTP	(2) WTP per HH
Standard	38.835	4.479
	(4.373)	(0.532)
Disgust	43.868	5.137
	(5.014)	(0.619)
Estimated difference	5.033	0.658
Std. Err.	(6.653)	(0.817)
Number of compounds	209	209

Note: this table shows mean willingness to pay (WTP) for a one-year subscription to the chlorine dispenser by messaging treatment, as well as estimated differences between treatments (disgust - standard). Column (1) reports total compound WTP, while column (2) reports WTP per household. WTP for compounds that dropped out before the sale is coded as zero. The sample is limited to compounds assigned to the group auction treatment. Units are Bangladesh Taka (BDT), approximately 75 BDT / 1 USD at the time of the sale. Estimation by OLS regression. Robust standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table 8: Willingness to pay by handwashing treatment

	(1) Compound WTP	(2) WTP per HH
No handwashing	43.116	5.203
	(6.176)	(0.759)
Handwashing	40.536	4.620
	(3.943)	(0.483)
Estimated difference	-2.580	-0.583
Std. Err.	(7.327)	(0.900)
Number of compounds	209	209

Note: this table shows mean willingness to pay (WTP) for a one-year subscription to the chlorine dispenser by handwashing treatment, as well as estimated differences between treatments (handwashing - no handwashing). Column (1) reports total compound WTP, while column (2) reports WTP per household. WTP for compounds that dropped out before the sale is coded as zero. The sample is limited to compounds assigned to the group auction treatment. Units are Bangladesh Taka (BDT), approximately 75 BDT / 1 USD at the time of the sale. Estimation by OLS regression. Robust standard errors in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

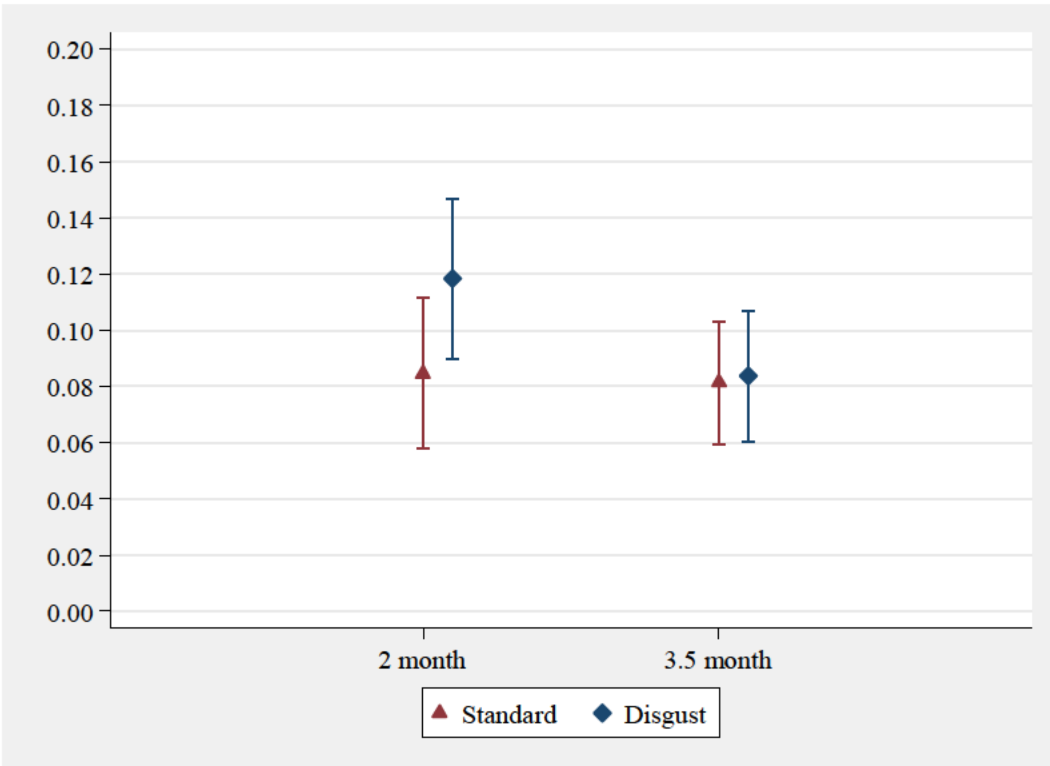
Table 9: WTP and Payment Compliance

	(1)	(2)
Compound bid	0.0065	0.0030
	(0.0079)	(0.0104)
Monthly payment	-0.0170	-0.0265
	(0.0113)	(0.0249)
Interaction of bid and payment		0.0001
		(0.0001)
Avg. marg. effect of WTP	0.0012	0.0014
Std. Err.	(0.0015)	(0.0015)
Number of compounds	52	52

Note: this table presents estimates from a logit model where the dependent variable is an indicator for whether the compound completed its yearly subscription, i.e. makes all 12 monthly payments, and the independent variables are the compound's WTP, i.e. its bid in BDM, the monthly subscription fee, i.e. the price drawn in BDM, and, in column (2), their interaction. The first three rows present logit coefficients, while the fourth row presents the average marginal effect of an increase in a compound's WTP. The sample consists of all compounds that participated in the group auction and won the subscription, i.e. the lottery price was less than or equal to the compound's bid.

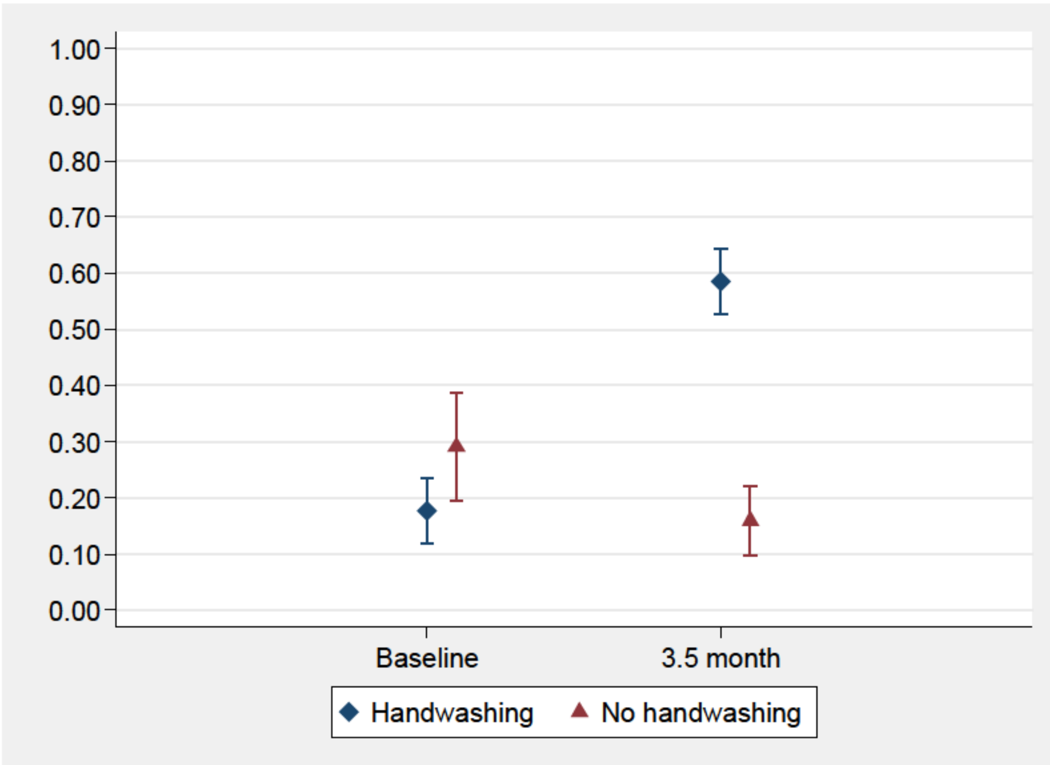
8 Figures

Figure 1: Rates of detectable chlorine in household drinking water, by treatment



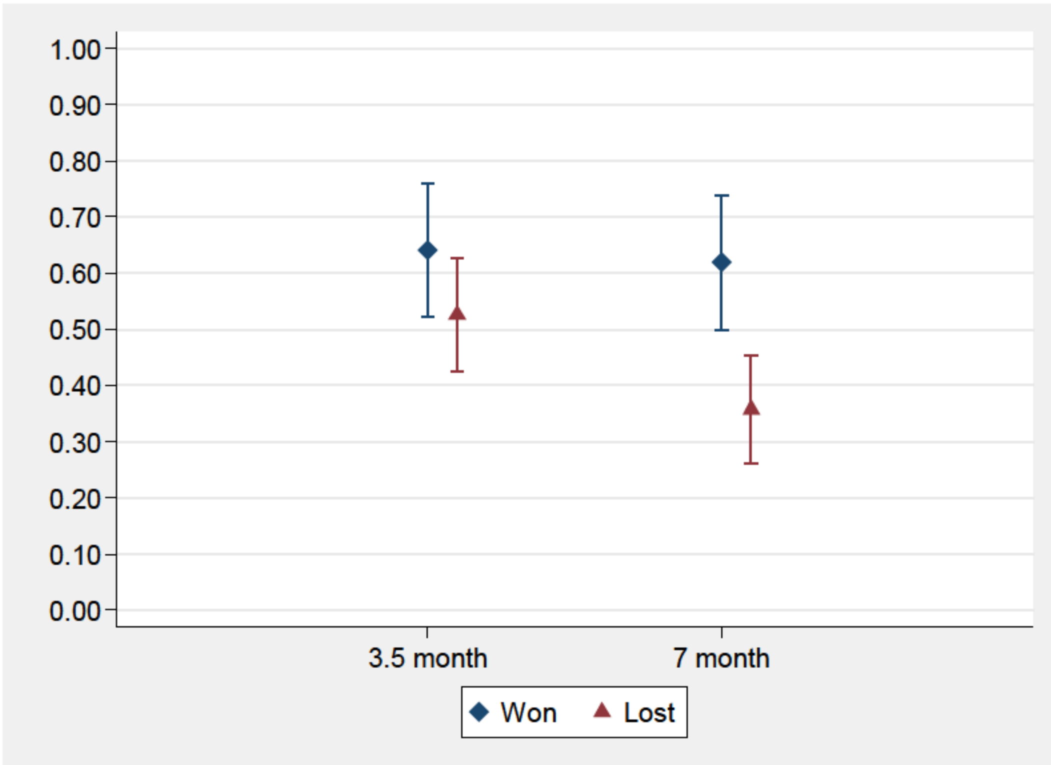
Note: this figure shows the share of households, by treatment and survey wave, with detectable chlorine in their drinking water. Point estimates and 95% confidence intervals estimated via logit regression. Standard errors clustered at the compound level. 2-month survey: 417 compounds, 2259 households. 3.5-month survey: 413 compounds, 2036 households.

Figure 2: Share of compounds with soap and water available at the common toilet, by treatment



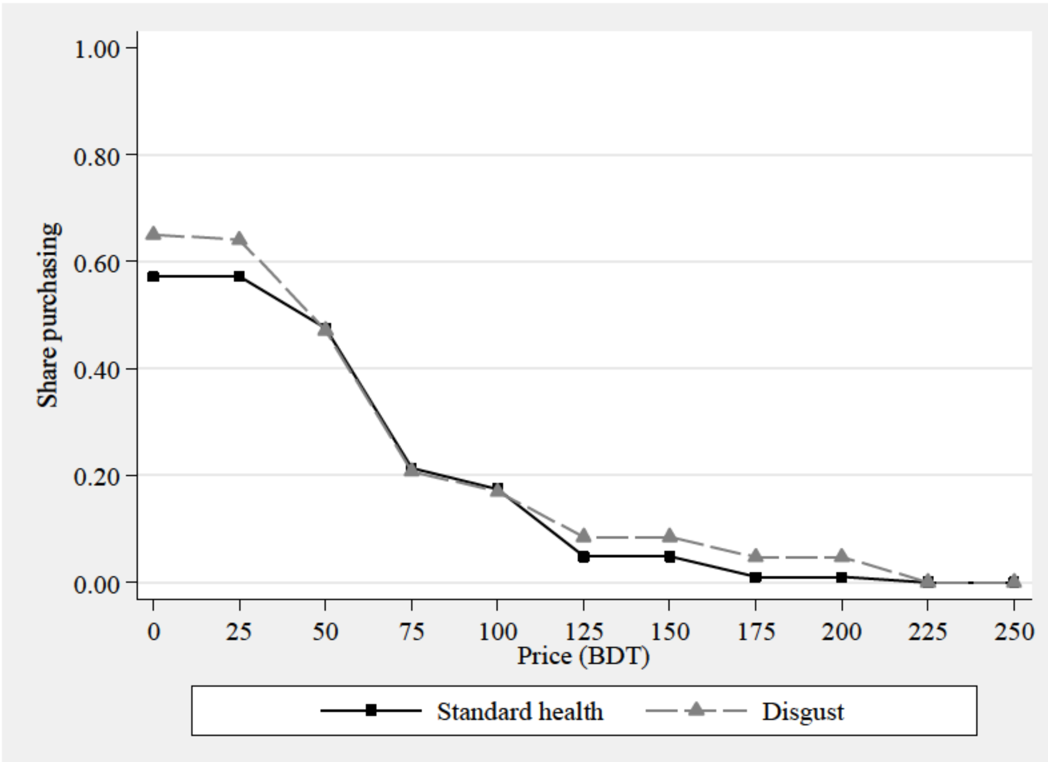
Note: this figure shows the share of compounds, by handwashing treatment and survey wave, with soap and water available at the common latrine. Point estimates and 95% confidence intervals estimated via logit regression. Standard errors clustered at the compound level. Baseline survey: 256 compounds. 3.5-month survey: 413 compounds.

Figure 3: Effect of losing free soap delivery on soap and water availability at the latrine



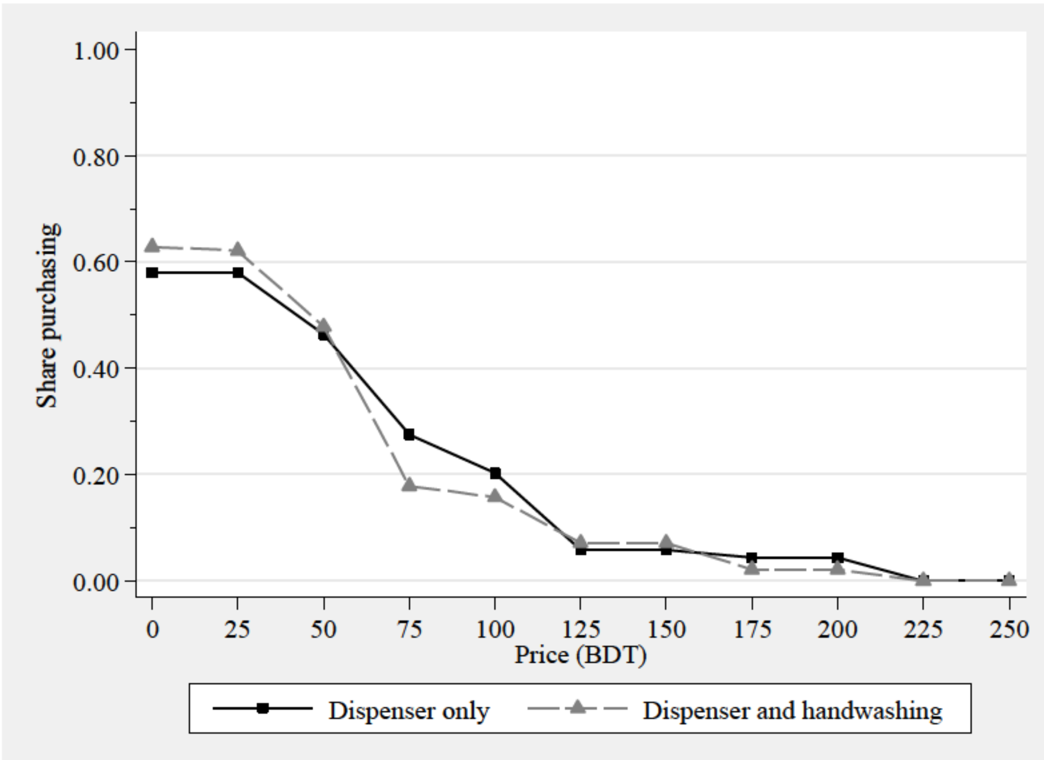
Note: this figure shows the share of compounds, by BDM outcome and survey wave, with soap and water available at the common latrine. Compounds that won kept the chlorine dispenser and the soapy water bottle, and continued to receive 2 packets of detergent per household per month for use in the soapy water bottle. Compounds that lost retained the soapy water bottle, but did not receive resupply of detergent. The sample consists of compounds in the handwashing arm and in which an auction was conducted. Point estimates and 95% confidence intervals estimated via logit regression. Standard errors clustered at the compound level. 3.5-month survey: 163 compounds. 7-month survey: 162 compounds.

Figure 4: Willingness to pay, by messaging treatment (group auction)



Note: this figure plots the share of compounds with total WTP greater than or equal to the indicated price, by messaging treatment. The sample is restricted to compounds in the group auction treatment.

Figure 5: Willingness to pay, by handwashing treatment (group auction)



Note: this figure plots the share of compounds with total WTP greater than or equal to the indicated price, by handwashing treatment. The sample is restricted to compounds in the group auction treatment.

Table A1: Descriptive Statistics and Balance Across All Treatments

	Mean (S.D.) (1)	p-value (2)
Compound characteristics:		
Number of households	9.21 (2.65)	1.000
Compound has gas	0.98 (0.13)	0.566
Soap and water available	0.215 (0.412)	0.047**
Household characteristics:		
Time lived in compound:		
Less than 6 months	0.165 (0.371)	0.829
6 months – 1 year	0.209 (0.407)	0.274
1–2 years	0.135 (0.342)	0.567
2 years or more	0.491 (0.500)	0.413
Household monthly income (BDT):		
Less than 8,000	0.285 (0.452)	0.385
8,000–12,000	0.389 (0.488)	0.180
12,000 or more	0.326 (0.469)	0.365
Household health behavior:		
Household consistently boils drinking water	0.629 (0.483)	0.988
Washed hands with soap after last latrine visit	0.681 (0.466)	0.975
Compound social environment:		
Compound residents take care of each other	0.876 (0.330)	0.116
Post-latrine handwashing behavior visible	0.584 (0.493)	0.145
Would feel shame if observed not washing hands	0.948 (0.222)	0.151

Notes: This table presents summary statistics collected at baseline and the p-value from an F-test of the hypothesis that the coefficients on all eight treatment categories (the interaction of standard health vs. disgust and shame, water treatment only vs. water treatment and handwashing hardware, collective auction vs. weakest-link auction) are equal to zero. Number of households and availability of gas collected at enrollment. Presence of soap and water at the latrine collected at the first rapid observation visit (see study timeline). Household characteristics and attitudes collected in a baseline survey, interviewing two households per compound. "Compound residents take care of each other," "Post-latrine handwashing behavior visible," and "Would feel shame if observed not washing hands" indicate the respondent responded "Agree" or "Strongly agree" to the statement "In this compound we take care of each other," "I can observe others washing or failing to wash hands with soap (or, if they carry soap) when they leave the latrine," and "I would feel ashamed if my neighbors saw me leave the latrine without carrying soap," respectively.

Table A2: Descriptive Statistics and Balance
Messaging: Standard Health vs. Disgust & Shame

	Standard Health (1)	Disgust & Shame (2)	Diff. (3)	Norm. diff. (4)
Number of compounds	219	215		
Compound characteristics:				
Number of households	9.23	9.18	−0.05	−0.01
	(2.70)	(2.61)	[0.25]	
Compound has gas	0.99	0.98	−0.00	−0.03
	(0.12)	(0.14)	[0.01]	
Soap and water available	0.177	0.254	0.077	0.133
	(0.383)	(0.437)	[0.051]	
Household characteristics:				
Time lived in compound:				
Less than 6 months	0.182	0.147	−0.035	−0.067
	(0.386)	(0.354)	[0.027]	
6 months – 1 year	0.223	0.194	−0.029	−0.051
	(0.417)	(0.396)	[0.029]	
1–2 years	0.124	0.147	0.023	0.047
	(0.330)	(0.354)	[0.025]	
2 years or more	0.471	0.512	0.042	0.059
	(0.500)	(0.500)	[0.036]	
Household monthly income (BDT):				
Less than 8,000	0.283	0.287	0.004	0.007
	(0.451)	(0.453)	[0.035]	
8,000–12,000	0.374	0.404	0.030	0.044
	(0.484)	(0.491)	[0.035]	
12,000 or more	0.343	0.309	−0.035	−0.053
	(0.475)	(0.462)	[0.036]	
Household health behavior:				
Household consistently boils drinking water	0.640	0.617	−0.023	−0.034
	(0.481)	(0.487)	[0.040]	
Washed hands with soap after last latrine visit	0.675	0.687	0.011	0.017
	(0.469)	(0.464)	[0.034]	
Compound social environment:				
Compound residents take care of each other	0.846	0.907	0.061**	0.131
	(0.362)	(0.291)	[0.025]	
Post-latrine handwashing behavior visible	0.561	0.608	0.047	0.067
	(0.497)	(0.489)	[0.039]	
Would feel shame if observed not washing hands	0.956	0.940	−0.015	−0.049
	(0.206)	(0.237)	[0.017]	

Notes: This table presents summary statistics of key baseline variables for compounds assigned to the standard health messaging treatment (Column 1) and those assigned to the disgust and shame messaging treatment (Column 2), as well as the estimated difference (Column 3) and the normalized difference (Column 4) $\Delta_x = \left(\overline{X}_1 - \overline{X}_0\right) / \sqrt{S_0^2 + S_1^2}$ (Imbens and Wooldridge 2009) between the two groups. Number of households and availability of gas collected at enrollment. Presence of soap and water at the latrine collected at the first rapid observation visit (see study timeline). Household characteristics and attitudes collected in a baseline survey, interviewing two households per compound. "Compound residents take care of each other," "Post-latrine handwashing behavior visible," and "Would feel shame if observed not washing hands" indicate the respondent responded "Agree" or "Strongly agree" to the statement "In this compound we take care of each other," "I can observe others washing or failing to wash hands with soap (or, if they carry soap) when they leave the latrine," and "I would feel ashamed if my neighbors saw me leave the latrine without carrying soap," respectively.

Table A3: Descriptive Statistics and Balance
Hardware: Chlorine Dispenser Only vs. Dispenser and Handwashing Station

	Dispenser Only (1)	Dispenser & Handwashing (2)	Diff. (3)	Norm. diff. (4)
Number of compounds	143	291		
Compound characteristics:				
Number of households	9.16	9.23	0.07	0.02
	(2.66)	(2.65)	[0.27]	
Compound has gas	1.00	0.98	−0.02***	−0.16
	(0.00)	(0.15)	[0.01]	
Soap and water available	0.291	0.176	−0.114**	−0.192
	(0.457)	(0.382)	[0.057]	
Household characteristics:				
Time lived in compound:				
Less than 6 months	0.154	0.170	0.016	0.031
	(0.361)	(0.376)	[0.027]	
6 months – 1 year	0.198	0.214	0.017	0.029
	(0.399)	(0.411)	[0.030]	
1–2 years	0.150	0.128	−0.023	−0.046
	(0.358)	(0.334)	[0.027]	
2 years or more	0.498	0.488	−0.010	−0.014
	(0.501)	(0.500)	[0.037]	
Household monthly income (BDT):				
Less than 8,000	0.284	0.285	0.001	0.002
	(0.452)	(0.452)	[0.037]	
8,000–12,000	0.424	0.371	−0.053	−0.077
	(0.495)	(0.484)	[0.037]	
12,000 or more	0.291	0.343	0.052	0.079
	(0.455)	(0.475)	[0.038]	
Household health behavior:				
Household consistently boils drinking water	0.644	0.621	−0.022	−0.033
	(0.480)	(0.485)	[0.042]	
Washed hands with soap after last latrine visit	0.687	0.678	−0.009	−0.014
	(0.465)	(0.468)	[0.037]	
Compound social environment:				
Compound residents take care of each other	0.896	0.866	−0.029	−0.064
	(0.306)	(0.341)	[0.026]	
Post-latrine handwashing behavior visible	0.626	0.563	−0.063	−0.090
	(0.485)	(0.496)	[0.042]	
Would feel shame if observed not washing hands	0.975	0.935	−0.040**	−0.137
	(0.157)	(0.247)	[0.016]	

Notes: This table presents summary statistics of key baseline variables for compounds assigned to the chlorine dispenser only treatment (Column 1) and those assigned to the chlorine dispenser and handwashing station treatment (Column 2), as well as the estimated difference (Column 3) and the normalized difference (Column 4) $\Delta_x = \left(\overline{X}_1 - \overline{X}_0\right) / \sqrt{S_0^2 + S_1^2}$ (Imbens and Wooldridge 2009) between the two groups. Number of households and availability of gas collected at enrollment. Presence of soap and water at the latrine collected at the first rapid observation visit (see study timeline). Household characteristics and attitudes collected in a baseline survey, interviewing two households per compound. "Compound residents take care of each other," "Post-latrine handwashing behavior visible," and "Would feel shame if observed not washing hands" indicate the respondent responded "Agree" or "Strongly agree" to the statement "In this compound we take care of each other," "I can observe others washing or failing to wash hands with soap (or, if they carry soap) when they leave the latrine," and "I would feel ashamed if my neighbors saw me leave the latrine without carrying soap," respectively.

Table A4: Descriptive Statistics and Balance
Auction: Collective Bid vs. Individual Bid

	Collective (1)	Individual (2)	Diff. (3)	Norm. diff. (4)
Number of compounds	217	217		
Compound characteristics:				
Number of households	9.20	9.21	0.01	0.00
	(2.74)	(2.57)	[0.25]	
Compound has gas	0.99	0.98	−0.00	−0.03
	(0.12)	(0.13)	[0.01]	
Soap and water available	0.195	0.234	0.039	0.067
	(0.398)	(0.425)	[0.051]	
Household characteristics:				
Time lived in compound:				
Less than 6 months	0.160	0.170	0.010	0.019
	(0.367)	(0.376)	[0.027]	
6 months – 1 year	0.206	0.211	0.005	0.009
	(0.405)	(0.409)	[0.029]	
1–2 years	0.147	0.123	−0.025	−0.051
	(0.355)	(0.329)	[0.025]	
2 years or more	0.486	0.496	0.010	0.014
	(0.500)	(0.501)	[0.036]	
Household monthly income (BDT):				
Less than 8,000	0.297	0.273	−0.025	−0.039
	(0.458)	(0.446)	[0.035]	
8,000–12,000	0.368	0.410	0.042	0.061
	(0.483)	(0.492)	[0.035]	
12,000 or more	0.335	0.318	−0.017	−0.026
	(0.473)	(0.466)	[0.036]	
Household health behavior:				
Household consistently boils drinking water	0.627	0.630	0.003	0.004
	(0.484)	(0.483)	[0.040]	
Washed hands with soap after last latrine visit	0.696	0.666	−0.030	−0.045
	(0.461)	(0.472)	[0.034]	
Compound social environment:				
Compound residents take care of each other	0.894	0.858	−0.036	−0.077
	(0.308)	(0.350)	[0.026]	
Post-latrine handwashing behavior visible	0.616	0.552	−0.063	−0.091
	(0.487)	(0.498)	[0.039]	
Would feel shame if observed not washing hands	0.953	0.943	−0.010	−0.031
	(0.212)	(0.232)	[0.017]	

Notes: This table presents summary statistics of key baseline variables for compounds assigned to the collective bid auction treatment (Column 1) and those assigned to the individual bid auction treatment (Column 2), as well as the estimated difference (Column 3) and the normalized difference (Column 4) $\Delta_x = \left(\overline{X}_1 - \overline{X}_0\right)/\sqrt{S_0^2 + S_1^2}$ (Imbens and Wooldridge 2009) between the two groups. Number of households and availability of gas collected at enrollment. Presence of soap and water at the latrine collected at the first rapid observation visit (see study timeline). Household characteristics and attitudes collected in a baseline survey, interviewing two households per compound. "Compound residents take care of each other," "Post-latrine handwashing behavior visible," and "Would feel shame if observed not washing hands" indicate the respondent responded "Agree" or "Strongly agree" to the statement "In this compound we take care of each other," "I can observe others washing or failing to wash hands with soap (or, if they carry soap) when they leave the

Disgust, Shame and Soapy Water:
Tests of Novel Interventions to Promote Safe Water and Hygiene
Online Appendix

Raymond P. Guiteras[a,b], David I. Levine[c], Stephen P. Luby[d],
Thomas H. Polley[e], Kaniz Khatun-e-Jannat[f], Leanne Unicomb[f]

August 2015

[a] Department of Economics, University of Maryland, College Park, MD 20742 USA.
[b] Corresponding author. Tel. +1 301 405 3480; fax +1 301 405 3582. guiteras@econ.umd.edu
[c] Haas School of Business, University of California, Berkeley, CA 94720 USA
[d] Woods Institute for the Environment, MC 4205, Stanford University, Stanford, CA 94305 USA.
[e] Department of Economics, Duke University, Durham, NC 27708 USA.
[f] Water, Sanitation and Hygiene Research Group, ICDDR,B, GPO Box 128, Dhaka 1000, Bangladesh.

Contents

1 Model 3

1.1 Baseline case 3

1.2 Present Bias 4

1.3 Disgust 5

1.4 Shame 5

2 Sample Selection 7

Selection of Compounds 7

Selection of Households 7

3 Sequential Randomization 8

4 Disgust Box 11

5 References 12

6 Tables 13

7 Figures 24

1 Model

In this section, we expand on the intuition provided in the main text by developing a simple formal model of health behavior. In the model, we describe how disgust and shame can motivate positive health behavior, and outline the circumstances under which they may be more or less effective than a standard positive health treatment. Our baseline case is a neoclassical agent, i.e an expected utility maximizer who discounts future utility exponentially. We then extend the model to allow for (a) present bias, which we model as classic beta-delta discounting (Laibson 1997) and (b) non-standard preferences, in particular the utility impact of disgust and feelings of shame. To focus on the key issues at hand, the model abstracts from potentially important issues such as habit formation or discordant preferences within the household.

1.1 Baseline case

Consider an agent considering whether or not to take a preventative action to mitigate an environmental health risk. For concreteness, we will refer to this action as treating water with chlorine, but other behaviors such as handwashing apply as well. The cost of treating water is c, which can include both financial and non-financial costs (e.g., inconvenience, distaste for chlorinated water). The cost is incurred in the current period t. The benefit of water treatment is a reduction in the probability of illness in the next period, $t+1$. The agent believes the probability of illness is π_0 if she does not treat her water and π_1 if she does treat her water, and believes the cost of illness to be h.[7] We assume her per-period utility is linear and separable,

$$U_t = -c_t - h_t, \tag{1}$$

where c_t is zero if not treating or c if treating, h_t is zero if not sick and h if sick, and her per-period discount factor is δ. The current-period expected utility gain from water treatment is the discounted value of the increased probability of remaining healthy, $\delta(\pi_0 - \pi_1)h$, and the current-period cost is c, so the agent will treat her water if and only if

$$\delta(\pi_0 - \pi_1)h > c. \tag{2}$$

[7] We do not model the formation of beliefs, but do allow that, in principle, an intervention could alter these beliefs.

The implications for interventions to increase water treatment are clear. First, interventions that reduce financial costs, such as providing subsidized or free chlorine, or non-financial costs, such as increasing convenience or ease of use, or reduce negative elements such as the taste or smell of chlorine, are likely to increase treatment. Second, because many Bangladeshi households either do not know about the link between untreated water and disease (often because they do not believe piped water is contaminated) or do not believe that treatment can reduce the likelihood of disease (Gupta et al. 2008), educational interventions that increase the agent's subjective belief that water treatment reduces illness ($\pi_0 - \pi_1$) are also likely to increase treatment (e.g., Jalan and Somanathan (2008)). Finally, the agent might not be aware of all the costs of water-borne disease, such as long-term effects on child development, so educational efforts might seek to increase perceived h.

1.2 Present Bias

We now augment the model to allow present bias. We use the standard formulation: utility in the current period (t) is not discounted, utility in period $t+1$ is discounted by $\beta\delta$, and utility in any subsequent period $t+s$ is discounted by $\beta\delta^s$ (for $\beta < 1$). We assume agents are not sophisticated, in the sense of O'Donoghue and Rabin (1999): they do not account for how their present bias in the future will affect their future decisions.

Now the agent's gains from water treatment are reduced by the factor β, because these gains are realized in period $t+1$, while the costs, incurred in the current period, are not affected, so she will treat her water in period t if and only if

$$\beta\delta\left(\pi_0 - \pi_1\right)h > c. \tag{3}$$

This inequality is more difficult to satisfy than Equation (2). The present bias term β is especially important if the relevant time horizon is short: daily discount factors are rarely below 0.995, while the present bias factor in poor nations has been estimated at approximately 0.70 (Bisin and Hyndman 2014, Tanaka, Camerer, and Nguyen 2010, Nguyen 2009). Note that the agent may display time-inconsistency: in period t, an agent deciding on her action in period $t+1$ would follow the decision rule given by Equation (2), not Equation (3), because β is applied to both costs and benefits.

1.3 Disgust

We now enrich our model with the emotion of disgust. We assume: (a) disgust, as a visceral, emotional reaction, carries an immediate[8] utility cost d; (b) while disgust at consuming human feces is an inherent trait, interventions can "increase" disgust, in the sense of making it more salient to individuals that failing to treat water or wash hands will lead to the consumption of human feces. In the context of the model, then, an intervention that successfully causes agents to feel disgust if they do not treat their water means that not treating water will cause an immediate utility loss of d.

A time-consistent agent will now treat her water if and only if

$$\delta(\pi_0 - \pi_1)h > c - d, \tag{4}$$

which is always easier to satisfy than Equation (2), because the benefit of avoided disgust offsets some of the cost of treatment. A present-biased agent will now treat her water if and only if

$$\beta\delta(\pi_0 - \pi_1)h > c - d. \tag{5}$$

Note that, for a present-biased agent, an intervention targeting d will be especially effective relative to an intervention that increases $(\pi_0 - \pi_1)h$ by an equal amount, because d is not discounted.

1.4 Shame

We model shame as a utility cost to being observed violating social norms by others. In this context, shame can be caused by being observed committing a disgusting act, e.g. failing to treat water or failing to wash hands after defecating.

We consider two forms of shame, internal and external. Internal shame consists of the inherent and immediate disutility of being observed failing to treat water, independent of any action the observer might take. Mobilizing the emotion of internal shame requires that at the agent's decision time, she believe observers know about the fecal contamination. This internal shame cost (S_I) enters the utility function in parallel with disgust, but multiplied by the subjective probability of being observed by a neighbor who knows there is fecal contamination without prevention, which we denote π_{Obs}. This

[8] There can be some delay if drinking occurs after the opportunity for treating and if disgust occurs only when drinking; however, most water in this setting is drunk soon after it is collected, almost always in the same day.

probability is itself the product of the probability of being observed and the probability that an observer will consider failing to treat water a disgusting act. That is, shame is more likely to be incurred in situations where (a) one is very likely to be observed and (b) there is a strong social norm that failing to treat water is disgusting. Condition (a) is plausible in our context, where many families share a water source, latrine and handwashing station, and activities in these common areas are easily observable. Condition (b) depends on the effectiveness of an intervention targeting feelings of disgust.

External shame costs (S_E) are the consequence of social sanctions. If someone breaks social norms within a cohesive group, he or she may fear loss of status, ostracism, ridicule, and other sanctions (Curtis, Danquah, and Aunger 2009). Mobilizing fear of sanctions requires that the agent believe observers recall the presence of fecal contamination and that the agent care about her standing within a social group that he or she shares with observers. Because any sanctions would occur in the future, they are discounted by $\beta\delta$.

Incorporating internal and external shame into the agent's utility function leads to the inequality

$$\beta\delta\left[(\pi_0 - \pi_1)h + \pi_{Obs}S_E\right] > c - d - \pi_{Obs}S_I \, , \tag{6}$$

where $\beta = 1$ for an exponential discounter. An effective intervention targeting disgust and shame will increase all of disgust (d), the perceived probability of being observed and sanctioned (π_{Obs}), the cost of the social sanction (S_E) and internalized shame costs (S_I).

2 Sample Selection

Selection of Compounds

Because these are informal settlements, it was not possible to construct a proper sampling frame. Instead, within the chosen field sites, enumerators were instructed to follow a basic, designated route through the chosen field sites. Upon identifying an eligible compound, the enumerator would contact the compound manager, the person who runs day-to-day affairs of the compound on behalf of landlords, and who typically but not always resides in the compound. The enumerator would tell the manager that the compound was eligible for an icddr,b promotion, and ask for written consent to participate in the study. If the manager declined consent, the enumerator moved on to the next eligible compound. If the manager gave consent, the enumerator collected basic stratification data. To reduce possible spillovers, compounds within 50 meters of an enrolled compound were not subsequently approached.

Selection of Households

Water testing: six households were selected randomly among those with an adult present at the time of the visit. In subsequent visits, these original six households were prioritized, with additional households selected at random until six samples were obtained.

Household survey: two households per compound were randomly selected at baseline. At the midline and endline surveys, these two households were prioritized, and replaced with a randomly selected household if not available.

3 Sequential Randomization

Here, we provide a brief summary of the sequential randomization method we employed. We provide detailed exposition on the method and field implementation in Guiteras, Levine, and Polley (2015).[9] Stata code is available from the authors upon request.

For intuition, consider a single, binary treatment that the researcher wishes to randomize, stratifying on a single, binary covariate, e.g. men and women. However, the researcher receives subjects passively, without knowing the share of men and women in the sample. Suppose that the next subject to arrive is a woman. If more women are currently allocated to treatment than control, then allocating this woman to control will reduce the variance of the estimated treatment effect by more than allocating her to treatment, so the optimal allocation for her is to control.[10] Similarly, if more women are currently allocated to control than treatment, she should be allocated to treatment. If there are equal numbers of women in treatment and control, the researcher should allocate her to the arm with fewer men to minimize the overall variance, or flip a coin if men are balanced.

Atkinson's D_A-optimal sequential allocation method (Atkinson 1982) generalizes this intuition to more complex designs with multiple treatments and multiple stratification variables. The researcher's objective function is a weighted average of the expected variances of the estimated treatments, where the researcher chooses the weights. As each unit arrives, the algorithm chooses the assignment that minimizes that minimizes this objective function, given that unit's stratification covariates and the allocation of previously enrolled units. Because this weighted average is proportional to the determinant of a quadratic form involving the sample design matrix (treatments and stratification

[9] Stata code is provided at http://www.econ.umd.edu/research/papers/617, and we encourage any interested researchers to contact us with questions on the code or implementation.

[10] The variance of the estimated treatment effect on women, $V\left[\hat{\beta}_F\right]$, is equal to the variance of the difference in the estimated means, $V\left[\overline{y}_{F,T}-\overline{y}_{F,C}\right]$. This is equal to the sum of the variances of the components, $V\left[\overline{y}_{F,T}\right]+V\left[\overline{y}_{F,C}\right]$ (the covariance is zero), or $\sigma_F^2/N_{F,T}+\sigma_F^2/N_{F,C}$, where for simplicity we assume homoscedasticity and independence. The allocation that minimizes variance, then, is $N_{F,T}=N_{F,C}$.
In this simple example, it is unlikely that there will be any important efficiency loss from considering the subpopulation treatment effects separately, since the maximum imbalance in either subpopulation at any stage is 1. However, in a more complex design, it is not necessarily optimal to consider only the precision of the subgroup to which the current subject belongs. It may be that the allocation within that subgroup is imbalanced in one direction but the overall allocation is imbalanced in the other direction, so assigning this subject so as to minimize variance within its own subgroup does not minimize the variance of the estimate of the average treatment effect.

covariates), the algorithm requires only simple matrix algebra operations that an inexpensive computer using standard software can perform in real time.

The critical requirement is that the unit's exact place in the sequence be uncorrelated with potential outcomes.[11] This could be violated if, for example, in a clinical setting an intake nurse knew the algorithm and manipulated the order in which subjects were processed to ensure that a particular subject received a particular treatment. This was not likely in our context. Enumerators did not know which of several covariates they collected would be used as stratification variables, so they could not have anticipated which assignment any given compound would receive. In contexts where this is a concern, a "biased coin" version of the sequential allocation algorithm allocates a subject probabilistically, putting highest weight on the option that would reduce the variance the most.

Inference can be conducted using the usual regression-based methods, as we do in this paper. Alternatively, the researcher could follow the "reasoned basis for inference" logic of Fisher (1935) and construct counterfactual distributions by reshuffling the order in which subjects arrive. See also Rosenbaum (2010).

---

[11] Because the algorithm seeks to maintain balance at each point in the sequence, it is robust to trends or fluctuations in potential outcomes. For example, neither a geographic pattern to enrollment nor a change in recruitment methods would cause bias, even if these were correlated with potential outcomes (e.g., moving from richer to poorer neighborhoods, or making a greater effort to recruit poor subjects).

4 Cost-Effectiveness

About 1 of every 1000 children ages 1-59 months dies of diarrhea in Bangladesh each year,[12,13] and about 40% of households in our sample have a child in this age range. Clasen et al. (2007) report that consistent water treatment can avert about 40% of diarrhea. Assume an averted child death is "worth" 25 DALYs, which we believe is a conservative assumption given that global burden of disease calculations have assumed 33 DALYs per child life saved (Mathers, Ezzati, and Lopez 2007). We adopt the WHO standard that an intervention is "very cost-effective" if it the cost of saving 1 DALY is less than or equal to 1 year's GDP.[14] Bangladesh's GDP per capita was about $950 in 2013.[15]

With these assumptions, providing chlorine dispensers are a "highly effective" intervention if it costs $9.75 or less per household per year that uses chlorine regularly. With an 8 percent usage rate after 7 months (Table 1), providing a chlorine dispenser and either marketing message to a compound with 8 households is a "highly effective" intervention if it costs up to $6.24 / year. However, we estimate that a small business or NGO running at scale could promote and distribute chlorine dispensers and visit monthly to replenish chlorine and collect fees at a break-even cost of 200 to 300 taka ($2.50 to $3.50) per compound per month, or a minimum of $30 / year, well above the cost-effectiveness threshold (authors' calculations).

Almost twice as many children 1-59 months die of pneumonia as diarrhea in Bangladesh each year.[16] Handwashing with soap averts very approximately a third of both diarrhea and pneumonia (Fewtrell et al. 2005, Rabie and Curtis 2006). Thus, soapy bottles are a "highly effective" intervention if it costs $23.55 or less per household per year that washes hands with soap regularly. With about 4.8 percentage points higher handwashing with soap when we provided a soapy bottle (Table 5), providing soapy bottles, refills on soap and either marketing message to a compound with 8 households is a "highly effective" intervention if it costs up to $9 per year. The approximate cost of the soapy bottle intervention is approximately $6 per compound per year, below this threshold (authors' calculations).

These estimated benefits do not include medical costs saved (for households, governments and NGOs), the utility value of lower morbidity, or the time savings of avoiding illness. However, an offsetting downward bias may result from our assumption that no chlorine users would have boiled.

---

12 http://www.unicef.org/bangladesh/media_7870.htm
13 http://www.who.int/maternal_child_adolescent/epidemiology/profiles/neonatal_child/bgd.pdf
14 http://www.who.int/choice/costs/CER_levels/en/
15 http://data.worldbank.org/indicator/NY.GDP.PCAP.CD
16 http://www.who.int/maternal_child_adolescent/epidemiology/profiles/neonatal_child/bgd.pdf

5 Disgust Box

As part of the disgust-and-shame presentation we used a custom presentation tool, the "disgust box." The key moment occurs when presenter pours clear water on the top of a box after the audience has learned that the horizontal pipe has holes in it and after the presenter has placed (fake) feces on top of the pipe. "This water sprinkling down is like rain from the sky," she explained.

She poured clean water through the pipe on the left. The water ran clear from the pipe on the right.

When she offered the water to the audience, they agreed it looked clear but was disgusting to drink. Coupled with photos of pipes running through untreated sewage, the presentation evoked feelings of disgust in most audience members.

The full disgust and shame presentation is provided in the Online Supplement. A video of the disgust box portion of the presentation is available at https://www.youtube.com/watch?v=pnEqblSbzq8.

6 References

Atkinson, Anthony C. 1982. "Optimum Biased Coin Designs for Sequential Clinical Trials with Prognostic factors." *Biometrika* no. 69 (1):61-67.

Bisin, Alberto, and Kyle Hyndman. 2014. Present-Bias, Procrastination and Deadlines in a Field Experiment. In *NBER Working Paper*.

Clasen, Thomas, Wolf-Peter Schmidt, Tamer Rabie, Ian Roberts, and Sandy Cairncross. 2007. "Interventions to improve water quality for preventing diarrhoea: systematic review and meta-analysis." *BMJ* no. 334 (7597):782. doi: 10.1136/bmj.39118.489931.BE.

Curtis, V. A., L. O. Danquah, and R. V. Aunger. 2009. "Planned, motivated and habitual hygiene behaviour: an eleven country review." *Health Education Research* no. 24 (4):655-73. doi: 10.1093/her/cyp002.

Fewtrell, Lorna, Rachel B. Kaufmann, David Kay, Wayne Enanoria, Laurence Haller, and John 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 Diseases* no. 5 (1):42-52. doi: 10.1016/S1473-3099(04)01253-8.

Fisher, Ronald A. 1935. *Design of Experiments*. New York: Hafner.

Guiteras, Raymond P., David I. Levine, and Thomas Polley. 2015. The Pursuit of Balance in Sequential Randomized Trials.

Gupta, S. K., M. S. Islam, R. Johnston, P. K. Ram, and Stephen P. Luby. 2008. "The chulli water purifier: acceptability and effectiveness of an innovative strategy for household water treatment in Bangladesh." *The American Journal of Tropical Medicine and Hygiene* no. 78:979-84.

Jalan, Jyotsna, and E. Somanathan. 2008. "The importance of being informed: Experimental evidence on demand for environmental quality." *Journal of Development Economics* no. 87 (1):14-28.

Laibson, David. 1997. "Golden Eggs and Hyperbolic Discounting." *The Quarterly Journal of Economics* no. 112 (2):443-478. doi: 10.1162/003355397555253.

Mathers, Colin D., Majid Ezzati, and Alan D. Lopez. 2007. "Measuring the Burden of Neglected Tropical Diseases: The Global Burden of Disease Framework." *PLoS Negl Trop Dis* no. 1:e114. doi: 10.1371/journal.pntd.0000114.

Nguyen, Quang. 2009. Present biased and Rosca participation: Evidence from field experiment and household survey data in Vietnam. Paper read at Workshop Lyon-Toulouse BEE (Behavioral and Experimental Economics), at Toulouse.

O'Donoghue, Ted, and Matthew Rabin. 1999. "Doing It Now or Later." *American Economic Review* no. 89:103-124. doi: 10.1257/aer.89.1.103.

Rabie, Tamer, and Valerie Curtis. 2006. "Handwashing and risk of respiratory infections: a quantitative systematic review." *Tropical Medicine & International Health* no. 11:258-267. doi: 10.1111/j.1365-3156.2006.01568.x.

Rosenbaum, Paul R. 2010. *Design of observational studies*, *Springer series in statistics*. New York: Springer.

Tanaka, Tomomi, Colin F. Camerer, and Quang Nguyen. 2010. "Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam." *The American Economic Review* no. 100 (1):557-571. doi: 10.1257/aer.100.1.557.

7 Tables

Table S1: Share of households with detectable chlorine residual, by motivational treatment

	(1) Small; 2-mo.	(2) Small; 3.5 mo.	(3) Large; 2-mo.	(4) Large; 3.5 mo.
Standard message	0.1076	0.0956	0.0664	0.0707
	(0.0238)	(0.0174)	(0.0155)	(0.0144)
Disgust message	0.1232	0.0635	0.1146	0.0988
	(0.0209)	(0.0150)	(0.0203)	(0.0173)
Est. diff. (disgust - standard)	0.0156	-0.0320	0.0482*	0.0281
Std. Err.	(0.0317)	(0.0230)	(0.0255)	(0.0225)
Num. compounds	195	195	222	218
Num. households	989	875	1270	1161

Note: this table shows the share of households, by treatment and survey wave, with detectable chlorine in their drinking water, as well as the estimated difference between the disgust and standard treatments. Columns (1) and (2) present estimates for compounds with 8 or fewer households at baseline; Columns (3) and (4) present estimates for compounds with more than 8 households.Estimation by logit regression. Estimated discrete differences presented with standard errors clustered at the compound level in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table S2: Effect of handwashing treatment on availability of soap and water (Balanced Panel)

	(1) Baseline	(2) 3.5-month
No handwashing treatment	0.289	0.108
	(0.050)	(0.034)
Handwashing treatment	0.179	0.631
	(0.030)	(0.037)
Estimated difference	-0.111*	0.523***
Std. Err.	(0.058)	(0.051)
Difference in differences		0.633***
Std. Err.		(0.072)
Number of compounds	251	251

Note: this table shows the share of compounds, by handwashing treatment and survey wave, with soap and water available at the common latrine, as well as the estimated difference between treatments and, for survey wave 2 (3.5-month midline), a difference-in-differences estimate using the difference at baseline for comparison. Estimation by logit regression. Sample is restricted to a balanced panel, i.e. compounds surveyed in all 3 rounds. Standard errors clustered at the compound level in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table S3: Effect of handwashing treatment on availability of soap and water

	(1) Small; Baseline	(2) Small; 3.5-mo.	(3) Large; Baseline	(4) Large; 3.5-mo.
No handwashing treatment	0.111	0.111	0.420	0.200
	(0.053)	(0.040)	(0.070)	(0.046)
Handwashing treatment	0.143	0.568	0.204	0.601
	(0.040)	(0.043)	(0.042)	(0.041)
Estimated difference	0.032	0.457***	-0.216***	0.401***
Std. Err.	(0.066)	(0.059)	(0.082)	(0.062)
Difference in differences		0.425***		0.617***
Std. Err.		(0.086)		(0.099)
Number of compounds	113	195	143	218

Note: this table shows the share of compounds, by handwashing treatment and survey wave, with soap and water available at the common latrine, as well as the estimated difference between treatments and, for survey wave 2 (3.5-month midline), a difference-in-difference estimate using differences at baseline for comparison. Columns (1) and (2) present estimates for compounds with 8 or fewer households at baseline; Columns (3) and (4) present estimates for compounds with more than 8 households. Estimation by logit regression. Standard errors clustered at the compound level in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table S4: Effect of losing free soap delivery on soap and water availability, by compound size

	(1) Small; Won	(2) Small; Lost	(3) Large; Won	(4) Large; Lost
Midline (3.5 mo.)	0.676	0.520	0.581	0.500
	(0.081)	(0.071)	(0.090)	(0.073)
Endline (7 mo.)	0.576	0.280	0.677	0.438
	(0.087)	(0.064)	(0.085)	(0.072)
Estimated difference	-0.101	-0.240**	0.097	-0.062
Std. Err.	(0.105)	(0.101)	(0.134)	(0.096)
Diff-in-diffs		-0.139		-0.159
Std. Err.		(0.146)		(0.165)
Number of compounds	34	50	31	48

Note: this table shows the share of compounds with soap and water available at the latrine by BDM outcome and survey wave, by compound size (8 households or fewer vs. more than 8 households). Columns (1) and (3) show levels for compounds that won the BDM auction at the 3.5-month midline survey, i.e. during the free trial and approximately two weeks before the BDM auction, and the 7-month endline survey, approximately 3 months after the BDM auction, as well as the estimated difference between the midline and endline. Columns (2) and (4) show the same for compounds that lost the BDM auction. Columns (2) and (4) also provide difference-in-differences estimates comparing changes from midline to endline between compounds that lost the in the auction vs. those that won. Compounds that won kept the chlorine dispenser and the soapy water bottle, and continued to receive 2 packets of detergent per household per month for use in the soapy water bottle. Compounds that lost retained the soapy water bottle, but did not receive resupply of detergent. The sample consists of compounds in the handwashing arm and in which an auction was conducted. Estimation by logit regression. Standard errors clustered at the compound level in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table S5.A: Share of compounds with soap and water available at the latrine, by messaging treatment; Compounds with 8 or fewer households

	(1) Baseline	(2) 3.5-month	(3) 7-month
Standard	0.129	0.475	0.381
	(0.043)	(0.050)	(0.049)
Disgust	0.137	0.365	0.358
	(0.048)	(0.049)	(0.049)
Estimated difference	0.008	-0.110	-0.024
Std. Err.	(0.065)	(0.070)	(0.070)
Difference in differences		-0.118	-0.032
Std. Err.		(0.089)	(0.102)
Number of compounds	113	195	192

Table S5.B: Share of compounds with soap and water available at the latrine, by messaging treatment; Compounds with more than 8 households

	(1) Baseline	(2) 3.5-month	(3) 7-month
Standard	0.221	0.450	0.417
	(0.050)	(0.048)	(0.048)
Disgust	0.333	0.477	0.454
	(0.055)	(0.048)	(0.048)
Estimated difference	0.113	0.028	0.037
Std. Err.	(0.074)	(0.068)	(0.068)
Difference in differences		-0.085	-0.076
Std. Err.		(0.101)	(0.101)
Number of compounds	143	218	216

Note: this table shows the share of compounds, by messaging treatment and survey wave, with soap and water available at the common latrine, as well as the estimated difference between treatments and, for survey waves 2 (3.5-month midline) and 3 (7-month endline), difference-in-difference estimates using differences at baseline for comparison. Estimation by logit regression. Standard errors clustered at the compound level in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table S6: Share of compounds with soap and water available at the latrine, by messaging treatment

	Handwashing			Non-handwashing		
	(1) Baseline	(2) 3.5-month	(3) 7-month	(4) Baseline	(5) 3.5-month	(6) 7-month
Standard	0.161	0.607	0.457	0.209	0.162	0.284
	(0.040)	(0.041)	(0.042)	(0.062)	(0.045)	(0.055)
Disgust	0.193	0.563	0.433	0.372	0.157	0.362
	(0.043)	(0.043)	(0.043)	(0.074)	(0.044)	(0.058)
Estimated difference	0.032	-0.044	-0.024	0.163*	-0.005	0.079
Std. Err.	(0.059)	(0.059)	(0.060)	(0.097)	(0.063)	(0.080)
Difference in differences		-0.076	-0.056		-0.167	-0.084
Std. Err.		(0.079)	(0.086)		(0.112)	(0.131)
Number of compounds	170	275	272	86	138	136

Note: this table shows the share of compounds, by messaging treatment and survey wave, with soap and water available at the common latrine, as well as the estimated difference between treatments and, for survey waves 2 (3.5-month midline) and 3 (7-month endline), difference-in-difference estimates using differences at baseline for comparison. Columns (1) - (3) restrict the sample to compounds assigned to the Handwashing treatment; Columns (4) - (6) restrict the sample to compounds assigned to the Non-handwashing treatment. Estimation by logit regression. Standard errors clustered at the compound level in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table S7.A: Handwashing after visiting toilet, by handwashing treatment
Compounds with 8 or fewer households

	(1) Used water	(2) Used soap	(3) Used soap, both hands
No handwashing	0.528	0.108	0.0982
	(0.0251)	(0.0154)	(0.0146)
Handwashing	0.556	0.191	0.170
	(0.0182)	(0.0140)	(0.0131)
Estimated difference (HW - no HW)	0.029	0.083***	0.072***
Std. Err.	(0.031)	(0.021)	(0.020)
Number of compounds	195	195	195
Number of observations	2196	2200	2210

Table S7.B: Handwashing after visiting toilet, by handwashing treatment
Compounds with more than 8 households

	(1) Used water	(2) Used soap	(3) Used soap, both hands
No handwashing	0.579	0.118	0.112
	(0.0270)	(0.0167)	(0.0168)
Handwashing	0.544	0.159	0.143
	(0.0195)	(0.0129)	(0.0121)
Estimated difference (HW - no HW)	-0.035	0.042**	0.030
Std. Err.	(0.033)	(0.021)	(0.021)
Number of compounds	222	222	222
Number of observations	2945	2951	2972

Note: this table shows the share of toilet events after which compound residents (1) rinsed their hands with water (with or without soap), (2) used soap to wash at least one hand, (3) used soap to wash both hands, by handwashing treatment, as well as the estimated difference between the handwashing and non-handwashing treatments. Data collected during structured observation at approximately month 2 of the free trial. Estimation by logit regression. Standard errors clustered at the compound level in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table S8.A: Handwashing after visiting toilet, by motivational treatment
Compounds with 8 or fewer households

	(1) Used water	(2) Used soap	(3) Used soap, both hands
Standard	0.557	0.165	0.150
	(0.0217)	(0.0165)	(0.0155)
Disgust	0.537	0.164	0.143
	(0.0199)	(0.0148)	(0.0137)
Estimated difference	-0.020	-0.001	-0.007
Std. Err.	(0.030)	(0.022)	(0.021)
Number of compounds	195	195	195
Number of observations	2196	2200	2210

Table S8.B: Handwashing after visiting toilet, by motivational treatment
Compounds with more than 8 households

	(1) Used water	(2) Used soap	(3) Used soap, both hands
Standard	0.553	0.133	0.123
	(0.0239)	(0.0134)	(0.0130)
Disgust	0.558	0.160	0.143
	(0.0205)	(0.0156)	(0.0148)
Estimated difference	0.005	0.027	0.020
Std. Err.	(0.031)	(0.021)	(0.020)
Number of compounds	222	222	222
Number of observations	2945	2951	2972

Note: this table shows the share of toilet events after which compound residents (1) rinsed their hands with water (with or without soap), (2) used soap to wash at least one hand, (3) used soap to wash both hands, by motivational treatment, as well as the estimated difference between the disgust and standard arms. Data collected during structured observation at approximately month 2 of the free trial. Estimation by logit regression. Standard errors clustered at the compound level in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table S9.A: Willingness to pay by messaging treatment
Compounds with 8 or fewer households

	(1) Compound WTP	(2) WTP per HH
Standard	35.500	5.309
	(6.146)	(0.937)
Disgust	42.453	6.128
	(7.428)	(1.049)
Estimated difference	6.953	0.819
Std. Err.	(9.641)	(1.406)
Number of compounds	103	103

Table S9.B: Willingness to pay by messaging treatment
Compounds with more than 8 households

	(1) Compound WTP	(2) WTP per HH
Standard	41.981	3.696
	(6.237)	(0.526)
Disgust	45.283	4.146
	(6.804)	(0.641)
Estimated difference	3.302	0.450
Std. Err.	(9.230)	(0.829)
Number of compounds	106	106

Note: this table shows mean willingness to pay (WTP) for a one-year subscription to the chlorine dispenser by messaging treatment, as well as estimated differences between treatments (disgust - standard). Column (1) reports total compound WTP, while column (2) reports WTP per household. WTP for compounds that dropped out before the sale is coded as zero. The sample is limited to compounds assigned to the group auction treatment. Units are Bangladesh Taka (BDT), approximately 75 BDT / 1 USD at the time of the sale. Estimation by OLS regression. Robust standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table S10.A: Willingness to pay by handwashing treatment
Compounds with 8 or fewer households

	(1) Compound WTP	(2) WTP per HH
No handwashing	41.129	6.139
	(9.181)	(1.337)
Handwashing	38.194	5.555
	(5.718)	(0.830)
Estimated difference	-2.935	-0.584
Std. Err.	(10.816)	(1.574)
Number of compounds	103	103

Table S10.B: Willingness to pay by handwashing treatment
Compounds with more than 8 households

	(1) Compound WTP	(2) WTP per HH
No handwashing	44.737	4.440
	(8.410)	(0.832)
Handwashing	43.015	3.631
	(5.451)	(0.446)
Estimated difference	-1.722	-0.809
Std. Err.	(10.022)	(0.944)
Number of compounds	106	106

Note: this table shows mean willingness to pay (WTP) for a one-year subscription to the chlorine dispenser by handwashing treatment, as well as estimated differences between treatments (handwashing - no handwashing). Column (1) reports total compound WTP, while column (2) reports WTP per household. WTP for compounds that dropped out before the sale is coded as zero. The sample is limited to compounds assigned to the group auction treatment. Units are Bangladesh Taka (BDT), approximately 75 BDT / 1 USD at the time of the sale. Estimation by OLS regression. Robust standard errors in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

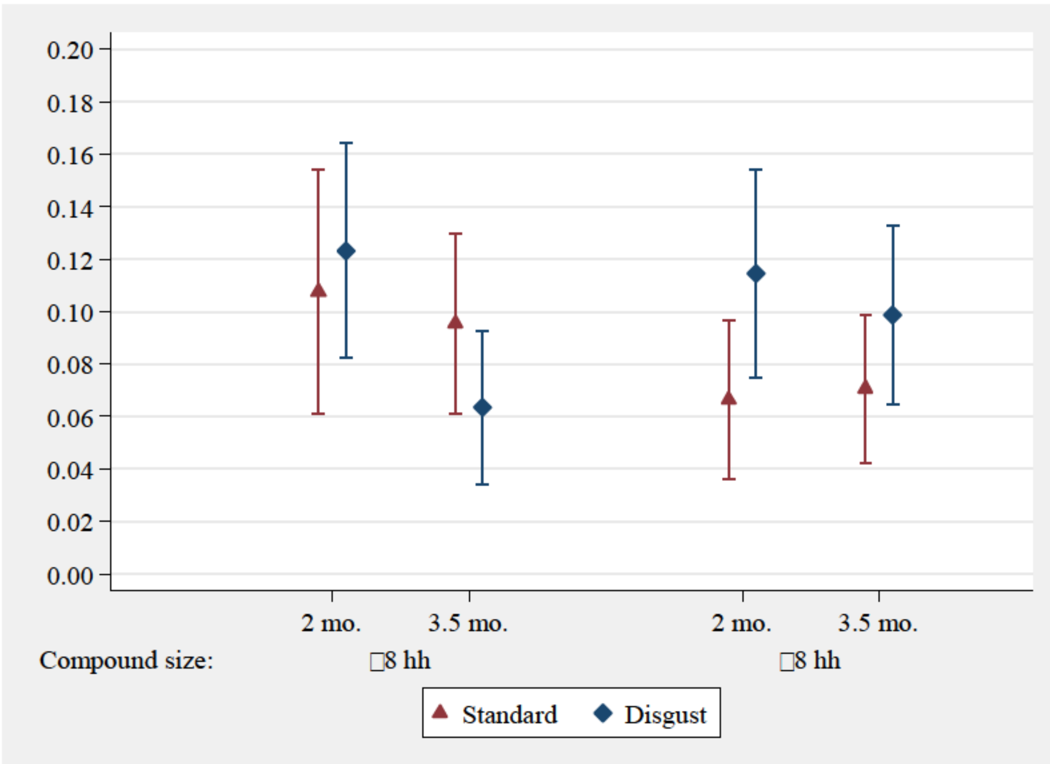
Table S11: WTP and Payment Compliance
Linear Probability Model

	(1)	(2)
Compound bid (BDT)	0.0011	0.0005
	(0.0012)	(0.0019)
Monthly payment (BDT)	-0.0034	-0.0051
	(0.0022)	(0.0054)
Interaction of bid and payment		0.0000
		(0.0000)
Avg. marg. effect of WTP	0.0011	0.0012
Std. Err.	(0.0012)	(0.0013)
Number of compounds	52	52

Note: this table presents estimates from a linear probability model where the dependent variable is an indicator for whether the compound completed its yearly subscription, i.e. makes all 12 monthly payments, and the independent variables are the compound's WTP, i.e. its bid in BDM, the monthly subscription fee, i.e. the price drawn in BDM, and, in column (2), their interaction. The first three rows present regression coefficients, while the fourth row presents the average marginal effect of a 1 BDT increase in a compound's WTP. The sample consists of all compounds that participated in the group auction and won the subscription, i.e. the lottery price was less than or equal to the compound's bid. Robust standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

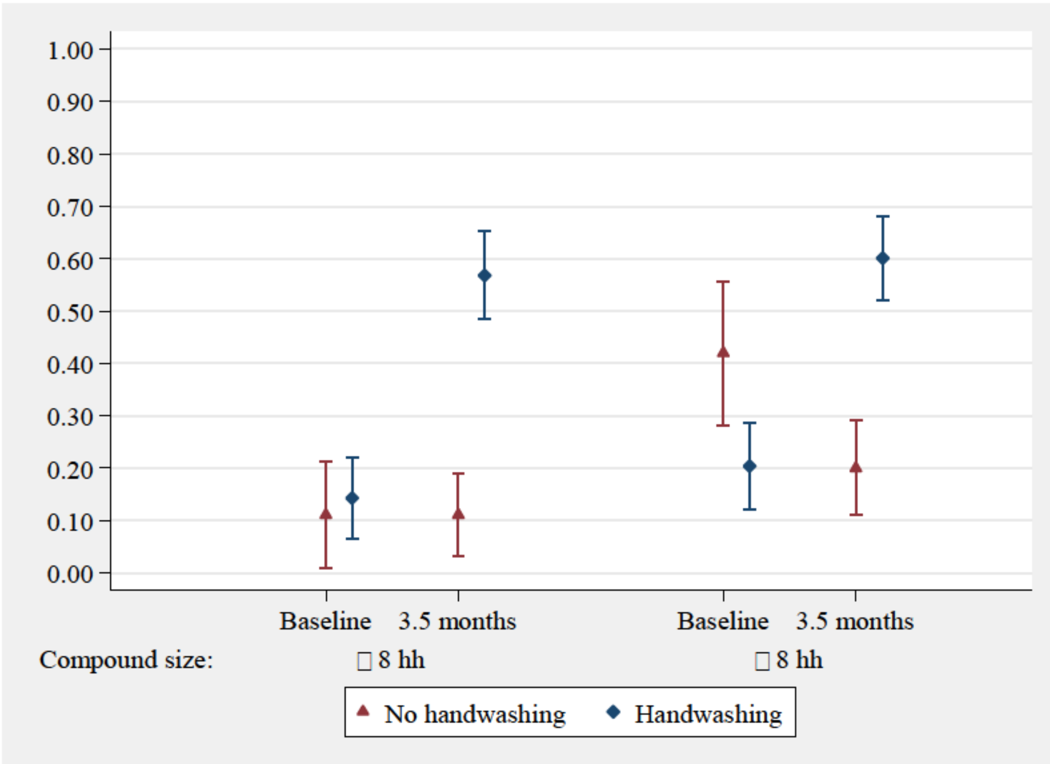
8 Figures

Figure S1: Rates of detectable chlorine in household drinking water, by treatment



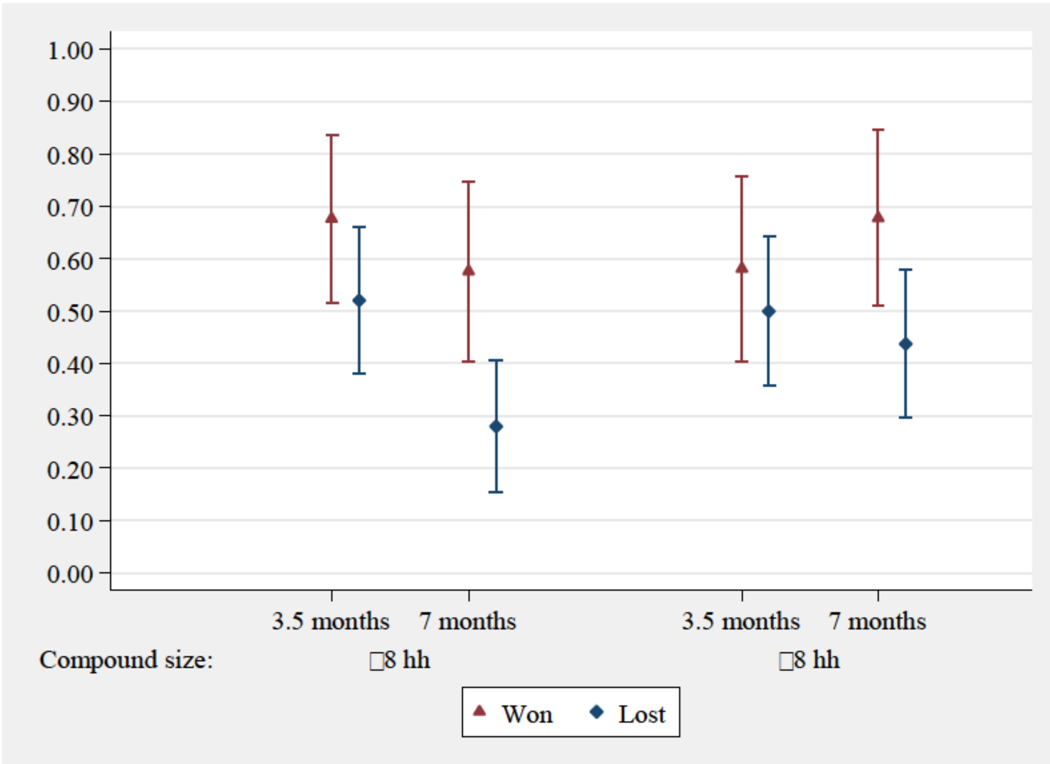
Note: this figure shows the share of households, by treatment and survey wave, with detectable chlorine in their drinking water. Estimates are presented, from left to right, for compounds with 8 or fewer households and compounds with more than 8 households. Point estimates and 95% confidence intervals estimated via logit regression. Standard errors clustered at the compound level.

Figure S2: Share of compounds with soap and water available at the common toilet, by treatment



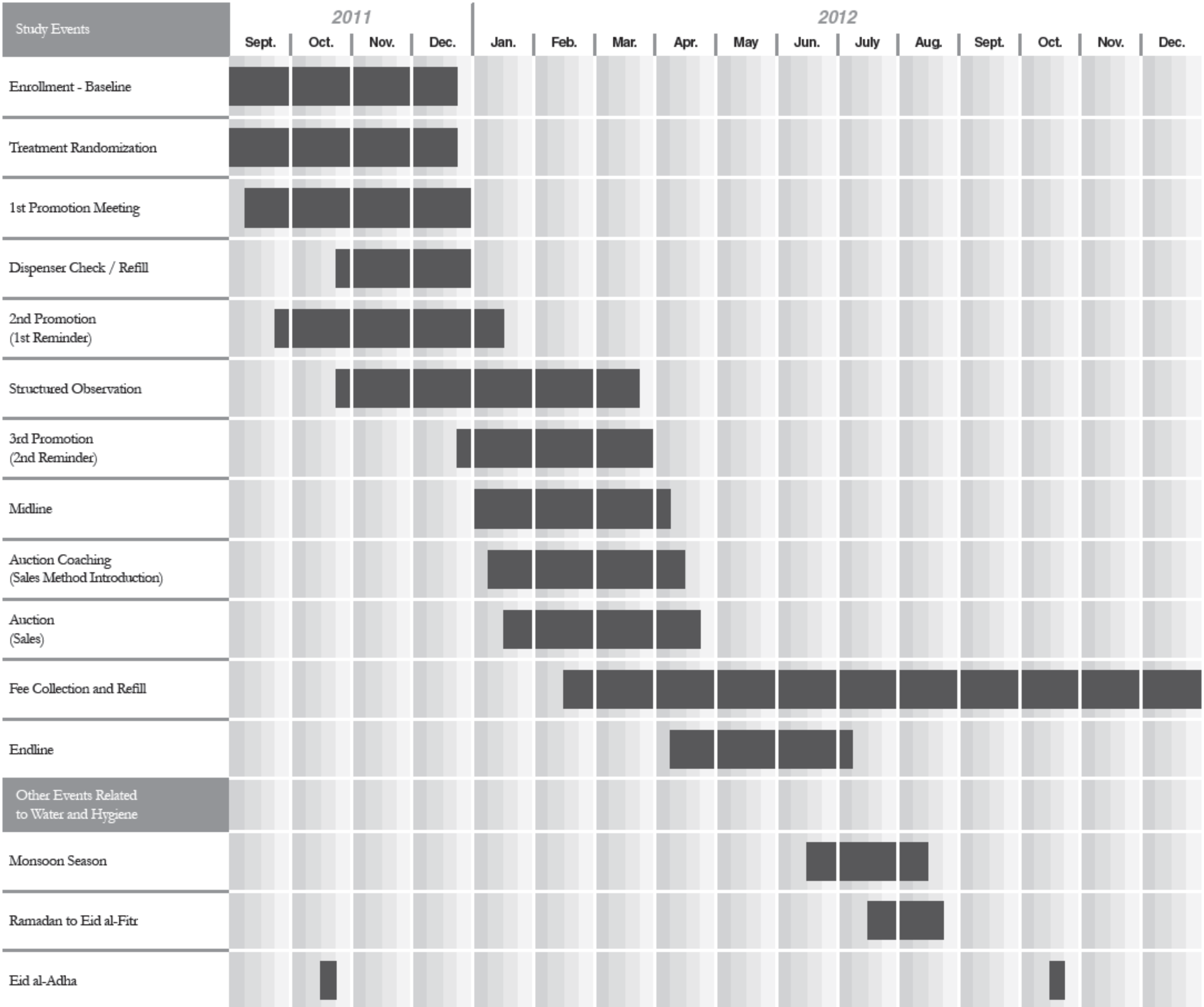
Note: this figure shows the share of compounds, by handwashing treatment and survey wave, with soap and water available at the common latrine. Point estimates and 95% confidence intervals estimated via logit regression. Standard errors clustered at the compound level. Estimates are presented, from left to right, for compounds with 8 or fewer households (baseline N=113; midline N=195) and compounds with more than 8 households (baseline N=143; midline N=218).

Figure S3: Effect of losing free soap delivery on soap and water availability at the latrine



Note: this figure shows the share of compounds, by BDM outcome and survey wave, with soap and water available at the common latrine. Compounds that won kept the chlorine dispenser and the soapy water bottle, and continued to receive 2 packets of detergent per household per month for use in the soapy water bottle. Compounds that lost retained the soapy water bottle, but did not receive resupply of detergent. The sample consists of compounds in the handwashing arm and in which an auction was conducted. Estimates are presented, from left to right, for compounds with 8 or fewer households (N=84, of which 34 won and 50 lost) and compounds with more than 8 households (N=79, of which 31 won and 48 lost).Point estimates and 95% confidence intervals estimated via logit regression. Standard errors clustered at the compound level.

Time Relative to Start of Free Trial	Visit Name	Data Collection	Behavior Change Campaign Activities
-1 week	Enrollment Baseline	Household Questionaire (2 hhs) Physical Observation (6 hhs) Rapid Observation (Spot Check)	None
-0.5 week	(Not a Visit) Treatment Randomization	None	None
Start of Free Trial	1st Promotion meeting	None	3 hour Campaign Meeting
2 weeks	2nd Promotion (1st Reminder)	Rapid Observation (Spot Check)	1.5 hour Campaign Meeting
As Needed	Dispenser Check / Refill	None	None
2 months	Structured Observation	Structured Observation (Handwashing, Water Treatment) Rapid Observation (Spot Check) Chlorine Residual Test	None
+3.5 months	3rd Promotion (2nd Reminder)	Rapid Observation (Spot Check)	1 hour Campaign Meeting
+3.5 months	Midline	Household Questionaire (2 hhs) Physical Observation (6 hhs) Rapid Observation (Spot Check)	None
+3.75 months	Auction Coaching (Sales Method Introduction)	Sales Practice BDM Understanding	10 minute Campaign Refresher
+4 months	Auction (Sales)	Dispenser WTP Rapid Observation (Spot Check)	10 minute Campaign Refresher
Monthly From +5 to +29 months (For Subscribers Only)	Fee Collection and Refill	Rapid Observation (Spot Check) Used Volume of Chlorine Payment	None
+7 months	Endline	Household Questionaire (2 hhs) Physical Observation (6 hhs) Rapid Observation (Spot Check)	None



Note: fee collection and refill continued for as long as 2 years, if compounds chose to renew their subscription after one year.