Improving the Adoption of Household Health Products: A Sales Experiment with Chlorine Tablets

Camille Boudot-Reddy Anita Mukherjee*

December 1, 2020

We test a door-to-door marketing intervention aimed to increase use of a targeted health product among poor households. Specifically, we examine three treatments in which this good–chlorine tablets for drinking water purification–is: (1) sold alone, (2) sold alongside a familiar and cheaper side good that is priced at its retail value, and (3) sold alongside the same side good that is priced on a promotional offer. The side good when sold at retail price is intended to be an "opt-out" good to reduce the marketing pressure, which should in turn reduce the amount of products sold that go unused. When the side good is sold on promotion, however, we hypothesize that it reintroduces marketing pressure due to the "gift" aspect of the promotion. Consistent with this hypothesis, we find that chlorine use is nearly double in the second condition compared to the other two conditions. Our results suggest that household valuation of a new product is shaped by both the presence and the price of a side good due to marketing pressure.

Keywords: health, technology adoption, chlorination, door-to-door marketing
JEL Codes: C93, D12, I11, M31, O12

*We thank Liang Bai, Marc Bellemare, Daniel Gottlieb, Jason Kerwin, Judd Kessler, Ty Leverty, Aprajit Mahajan, Emily Oster, Evan Polman, Laura Schechter, Jesse Shapiro, Justin Sydnor, C. Yiwei Zhang; seminar participants at the University of Edinburgh, University of Minnesota, and University of Wisconsin-Madison; and conference participants at the 2018 Royal Economic Society Conference, 2017 Annual International Conference of the Research Group on Development Economics, 2017 APPAM Fall Research Conference, and the 2017 Midwest International Economic Development Conference for helpful comments and conversations. Rubina Nafees Fatima, the CEO and founder of our partnering organization Safa, provided necessary support and guidance at each stage of this project. Anisha Singh and Alexis Reddy provided excellent research assistance in the field. We gratefully acknowledge financial support from Innovations for Poverty Action and the Wisconsin Alumni Research Fund. Boudot-Reddy was based at the Institute for Financial Management Research during part of the study period and acknowledges support from the organization. E-mails: camille.boudot@sms.ed.ac.uk and anita.mukherjee@wisc.edu.

1 Introduction

Many health products such as mosquito bednets and water purification tablets are widely considered to be both undervalued and underconsumed by poor households. Increasing consumption of these products has been the topic of much prior research, and there are no easy solutions. Informational campaigns can be costly and ineffective (Dupas 2009; Meredith et al. 2013), and interventions that work typically require large subsidies or free distribution (Cohen and Dupas 2010; Hoffmann, Barrett and Just 2009). To deal with these challenges, many agencies devote substantial resources towards door-to-door selling of products deemed to be beneficial or underconsumed.

Door-to-door marketing to poor households is not without drawbacks, however. A key challenge is that individuals may purchase items without using them—in fact, a number of prior studies using door-to-door marketing document large gaps between purchase and use even when the product is familiar. For example, Dupas (2009) finds that only 57% of households were using a bednet they had purchased upon a random follow-up in Kenya; Hanna, Duflo and Greenstone (2016) finds that purchases of clean cooking stoves went mostly unused in India; and Ashraf, Berry and Shapiro (2010) finds that nearly half of households did not use chlorine tablets upon buying them in Zambia. This is puzzling: why would poor households buy products that they won't use? The answer matters for two key reasons. First, since the goal of marketing efforts related to preventive health is product use, not just purchase, understanding the reasons for non-use matter. Second, if poor households are parting with valuable resources to buy goods they don't use, some marketing efforts may unintentionally make them worse off.

In this paper, we present results from a field experiment designed to evaluate whether simple changes to the door-to-door marketing setup could increase total use of a particular health product. Three conceptual questions drove our experimental design. First, we sought to understand whether the gap between purchase and use might be driven by behavioral factors such as households buying a product which they know they won't use out of guilt or reciprocity. After all, the agents who sell products door-to-door in poor areas are frequently volunteers who are well-respected in their communities, and households may feel uncomfortable saying "no" to either them or the organizations they represent. NGOs generally have played a large role in developing country communities as discussed in Brass et al. (2018), so it is plausible that they exert marketing pressure on households. (As a preview, we provide an opt-out for this feeling of guilt in our experiment and do not find evidence that it was helpful in our setting.)

Second, because the products being sold are frequently new or unfamiliar, we were interested to test whether co-selling a familiar good might help "frame" the unfamiliar product positively. As a thought experiment, consider someone selling a new brand of allergy medicine; might we be more receptive to purchasing and using this product if it is sold alongside something familiar like the branded painkiller Advil? Third, we investigated whether the pricing on this side good mattered. To stay with our thought experiment, would our perceived value of the allergy medicine be affected by whether the Advil was sold at market price versus on a promotional offer? For instance while the market price may serve as an "opt-out" good, the promotion may (re)introduce marketing pressure through a potential "gift" channel. Together, our conceptual questions attempt to shed light on how households make decisions about the purchase and use of (unfamiliar) preventive health products.

We take these conceptual questions to the field using an experiment centered on selling chlorine tablets (the unfamiliar target good) alongside a mosquito repellent product (the familiar side good). A key advantage of chlorine tablets is that we can collect information on actual usage via water quality tests, instead of having to rely on self-reported usage. The setting of our experiment is in the urban slums surrounding Hyderabad, India, where concerns about both water quality and mosquito-borne diseases loom large. We partnered with a well-known NGO in these localities, Safa, to market these products door-to-door. The first experimental treatment marketed the chlorine tablets alone. This serves as a control group to parallel the marketing norm of focusing on one good in door-to-door sales efforts. The second and third treatments sold these tablets alongside a second good, a mosquito repellent. The difference is that in the second treatment the mosquito repellent was sold at its market price, whereas in the third treatment it was sold on a "buy one get one free" promotion. (We considered a subsidy instead, but wanted to avoid concerns about liquidity constraints in our analysis.)

The appeal of these experimental variants is that they are low-cost and feasibly replicable in door-to-door and other marketing contexts. The experimental design also helps alleviate concerns about self-selection into preventive health products, which is otherwise prevalent; for example, Lee (2005) demonstrates meaningful self-selection into the demand for child immunizations in rural India. The present study is designed to test the conceptual questions of interest. For the first question about behavioral factors, we compare outcomes when the chlorine products are sold alone versus alongside a side (second) good, with the hypothesis being that the side good may provide an "opt-out" option to individuals seeking to purchase at least something from the marketer. When we do this, we find that less than 4% of those offered the two products buy the side good alone, suggesting that the screening effect we hypothesized is not present in our setting. This may be because our study setting was urban; in such an environment, individuals may be more comfortable saying "no" than in the rural areas which were the focus of prior related studies.

For the second and third questions, we compare the outcomes across all treatments to separate the impact of a side good into components based on the side good's effective price. Here, we find large effects. Compared to the other two treatments, when the side good (the mosquito repellent) is sold without any promotion, the take-up of the chlorine tablets stays the same but its total use nearly doubles. We also find that individuals are approximately 10% percent more likely to buy the chlorine tablets when the side good is on a promotional offer, though this does not translate to higher total use rates. We take these results as evidence that the pricing of the side good matters in determining focal product use. We hypothesize that the effect is through variation in marketing pressure induced by the different prices, and offer a conceptual model consistent with these findings in Appendix B. In this model, marketing efforts (including subsidies) reduce product use among purchasers by lowering the perceived quality or value of the product.

We note that our study has certain limitations. Due to the complex nature of the different theories we were interested in—i.e., effects of framing versus pressure or reciprocity are difficult to separate experimentally—we focused on simplistic marketing interventions that target how door-to-door marketing efforts can be improved. Since we set out to test for differences in the key outcomes only, our sample sizes are smaller than would be ideal for further analysis of the conceptual mechanisms that underlie our findings. Specifically, the marketing treatments did not yield enough variation in take-up to identify screening effects. Yet, we find our results compelling and hope that they can help further research on the use of health products in poor areas.

The rest of this article proceeds as follows. We begin by providing detailed information on the experimental products and providing a literature review in Section 2. In Section 3, we describe our marketing protocols and survey methods; Section 4 contains our results, and we offer concluding remarks in Section 5.

2 Background and Related Literature

Access to clean drinking water is a concern for nearly 800 million people worldwide, and diarrhea from unsafe water kills over half a million children under five years old each year (CDC 2017; Cutler and Miller 2005). Improved sanitation is a key way to reduce this risk, as shown in Kumar and Vollmer (2013). Water purification tablets can also help because they are effective, inexpensive, and easy to use. Their efficacy depends on use, however, so distributional campaigns must aim for behavioral changes along with making products widely available.

Most urban households in India receive water from government sources, which are not properly or regularly chlorinated and hence at risk of waterborne diseases. Thus, in our baseline survey, 85% of our sample obtained drinking water from public pipes and 78% of these households reported using some method of water purification. The most common technique, used by over 80% of this sub-sample, was straining. Though this process removes particulate matter, it does not clean the water from harmful pathogenic micro-organisms. Boiling and purifying filters—methods which would provide these households better protection—were used by only 7% and 3% of households, respectively. The main reason for low use of these methods is that they are both expensive and time-intensive. Chlorination could therefore provide these households with a cheaper and less time-consuming process to purify water.

The target good in our experiment is a bottle of chlorine tablets sold under the brand SafDrink; these tablets are intended for drinking water purification. While SafDrink is available in large pharmacies and sourced by the Government of India for mass-produced survival kits, it is not available or known in poor slum areas (we verified this to be the case through focus groups and visits to local pharmacies and retail shops). A bottle of SafDrink retails for Rs. 30 (about USD 0.50) and contains 100 tablets; a picture of the product is provided in Panel A of Figure 1. Each tablet purifies a liter of water. To use SafDrink correctly, the person treating the water must add a measured amount of water to a container, add the corresponding number of tablets, and allow the container to stand for 30 minutes before drinking.

An advantage of chlorine tablets is that product use can be measured objectively using chlorine strip tests of drinking water samples, alleviating the need to rely on household self-reports that could be inflated or inaccurate. The average amount of water consumed at home each day in our sample is one liter per person. An average household of six members (as in our sample) therefore needs approximately 42 liters of purified water per week. We therefore anticipated that if used regularly, one bottle of SafDrink would be consumed within approximately two weeks. To provide additional leeway between the sales and follow-up survey to ensure we could capture product use, we marketed SafDrink using a promotional offer "buy one get one free" under each treatment condition.

The side good was selected to be a mosquito repellent product: Good Knight Fast Card. The product is a packet of 10 small pieces of folding material that resembles paper which repels mosquitoes for about four hours once lit, shown in Panel B of Figure 1. One packet of the Good Knight Fast Card retails for Rs. 10 (about USD 0.15), about one-third the price of the bottle of chlorine tablets. This feature was essential in our hypothesis on "pressured" purchases. In order to help serve as an "opt out" option for individuals seeking to purchase at least something so as to avoid the disutility of saying "no", the side good had to be significantly cheaper than the target good.

An additional important feature of this side good was that it is not a direct substitute for the target good, though both products address preventive health needs. Year-round anxieties about contracting malaria, dengue, and other mosquito-borne diseases mean this product was widely available in local shops. As such, the majority of households in our study were aware of its existence, price, usage guidelines, and preventive health benefits. We hypothesized that the mosquito repellent could therefore serve as a positive "frame" for the chlorine tablets, to the extent that individuals increase their perceived value of a new unfamiliar product when sold alongside an already familiar, easy to use, and liked product. The choice of using the mosquito repellent product as our side good was thereby driven by its price, preventive health aspect, and potential ability to positively frame the SafDrink product given its popularity.

The present study is connected to a large body of research in economic development on ways to increase the purchase and use of preventive health products. For example, Dupas (2009) shows that subsidies increase take-up of mosquito bednets, though informational treatments regarding the health or financial returns to adoption have no impact. Mohapatra and Simon (2017) finds that women have particular influence over such products, and demonstrate the importance of intra-household bargaining in the context of clean cooking stoves. Other papers show the benefits of both one-time subsidies and free distribution in promoting product use (Fischer et al. 2014; Cohen and Dupas 2010; Dupas 2014). There is also evidence that liquidity constraints play an important role in the adoption of health products (e.g., Tarozzi et al. 2014). Together, this literature has determined that information campaigns do not reliably work in promoting purchase or adoption of specific goods, though thoughtfully designed price-discounts can be effective (even if expensive for the policymaker). The gap that our paper attempts to fill is on evidence regarding the impact of novel non-price marketing mechanisms, a topic on which there has been substantially less research.

Another set of related work has been conducted directly on chlorine tablets and other point-of-use water purification systems, the subject of our experiment. Ashraf, Jack and Kamenica (2013) examine a marketing strategy wherein newer and older versions of a water purification technology are sold together, finding that this contrast can make subsidies on the newer version more effective in encouraging purchase. Specifically related to product use, prior research finds that willingness to pay for chlorine tablets—not the actual price paid—is strongly linked with product use (Ashraf, Berry and Shapiro 2010). Günther and Schipper (2013) find that distributing improved water transport and storage containers to rural households is a cost effective way to increase water quality. Other research has tried to minimize the distribution of unused products by introducing "micro-ordeals" such as vouchers for chlorine tablets (Dupas et al. 2016), or health-oriented marketing messages (Luoto et al. 2014), both of which proved to be effective. The advantage of these methods is that they use non-price mechanisms to screen in potential users, which can be important in developing country settings.

Yet additional research examines non-price interventions in other preventive health technologies such as menstrual cups (Oster and Thornton 2012 studies peer effects) and condoms (Ashraf, Bandiera and Jack 2014 studies agent incentives), showing that these mechanisms hold promise. Among the recent research that does use product price incentives, behavioral factors have also been studied. For example, Lipscomb and Schechter (2018) examine the effect of subsidies, minimum investments, and access to other features of a mobile payment system on the demand for household sanitation services and document weak support for mental accounting-based policies.

Our research is also related to research examining the extent to which "pressure" may impact behaviors, particularly in door-to-door settings. For example, DellaVigna, List and Malmendier (2012) contribute evidence on the impact of social pressure in charitable contributions via a door-to-door marketing experiment that motivated the present study. In that study, the authors manipulated pressure in the door-to-door interaction by giving some households a pre-announcement that allowed them to opt-out by checking a "Do Not Disturb" option. The authors find that households donating under pressure are most likely to give small donations, which can be inferred as the cost of "saying no" to a charity seeking donations. The parallel cost of "saying no" in our study is purchase of a good without intention of use. Directly relevant to this paper, Das, Friedman and Kandpal (2018) find that implementation efforts matter a great deal in health interventions; they find that the quality and effort of the partner organization impacted the effectiveness of a malaria intervention in India. To the extent that these characteristics are correlated with the "pressure" households face when making purchase and usage decisions, that paper helps motivate the present study.

3 Experimental Setting and Survey Design

In this section, we describe the sampling procedures and methods for conducting the baseline and follow-up surveys. We also detail the door-to-door marketing experiment protocols and describe our measures for SafDrink use.[1]

3.1 Sample Selection and Baseline Survey

Our study setting is neighboring central-urban blocks in Hyderabad, a large city in southern India. These blocks are registered as 'notified slums' by the 2011 Indian Census, meaning they are "compact areas of at least 300 population or about 60-70 households of poorly built congested tenements, in an unhygienic environment, usually with inadequate infrastructure and lacking in proper sanitary and drinking water facilities." The problems with water quality motivated our choice of chlorine tablets as the target good.

The field experiment was conducted in partnership with Safa, an NGO that operates a range of programs aimed at income generation and female empowerment in our study location.[2] Safa was an ideal partner because of its strong presence and positive reputation among the study community. Our surveys indicated that 79% of Safa members reported being "very satisfied" with the NGO and 90% would recommend participating in Safa programs to others. Additionally, 60% of non-Safa members had heard of the NGO and 86% considered Safa to positively contribute to their community. Since these questions were asked by the surveyors not part of Safa, the respondents presumably provided an unbiased view of the NGO's reputation in the study areas.

The sample included Safa members (people with prior participation in a Safa program) and randomly selected non-members in the study areas. The Safa members were

[1] Note that all experimental scripts and surveys are provided in Appendices C, D, and E.

[2] Safa was founded as a social venture in 2006 with the mission of empowering women from low-income households through education and livelihoods. At the time of the study in 2016, the NGO had four centers in the locality and reached 5,000 households through 67 full-time staff. Safa primarily operates a large sponsorship program which enrolls girls in formal primary and secondary education. Additionally, the NGO conducts a range of training in skills such as stitching and mehendi (hand art) to provide women with home-based work opportunities.

selected from a complete list of members provided to us by the organization.[3] The non-members were selected using a standard methodology of sampling every n^{th} residence in a pre-specified direction from a beginning address. Specifically, we sampled every third household using a "right hand rule" from the home of each Safa member until we reached our target sample size. This process was designed to maximize the comparability of members and non-members.

The baseline survey was administered to a subset of the sample to collect pre-experiment information on knowledge of the target good, current preventive health practices, and perceptions about water quality. The survey was not administered to all sample households because of potential survey effects by which the questions would impact product purchase and use; for example, Zwane et al. (2011) provides evidence that surveys about water quality can increase product use specifically in the context of chlorine tablets. In our setting, we expected the effect of the baseline survey to be weaker for two reasons: (i) it was conducted by the research team, and not the NGO that later marketed the product, and (ii) it was conducted long before (on average eight weeks) collection of the follow-up measures that we focus on in our study.

Out of the 980 households in the study, 712 were administered the baseline survey. Upon arrival to each home, surveyors introduced themselves and asked to speak to the female head of the household. This person was home in almost every instance, since women tend to stay home during the day in these areas.[4] Our interviews targeted the female head of the household for two reasons. First, our focus group discussions prior to the experiment (and not in our study areas) suggested that generally women, not men, were responsible for household water treatment. Second, Safa predominantly targets

[3] The list included 350 members; some contact information was missing, and some members resided outside our study areas, reducing the sample size to 258.

[4] If the female head of household was unavailable, the surveyor attempted to return again the same day after inquiring about her schedule. If she could not be reached upon a second visit, a third appointment was fixed, after which if still unavailable the house was skipped and replaced by the next appropriate third home.

women in their door-to-door sales efforts as the products they promote are centered on women and child welfare.

The baseline survey took approximately 15 minutes to complete and covered four sections. The first section collected household demographic measures including age, marital status, education, along with household-composition, dwelling structure, and ownership of assets. The second section asked questions specific to prevalence of diarrhea in the household and preventive care practices including soap use and water treatment for drinking and washing. The third section focused on details about chlorine use if the household reported any use; this turned out not to be needed because no surveyed household reported using chlorine tablets. The fourth section asked about the individual's information sources regarding water purification.

Table 1 shows the sample characteristics by treatment category. On average, respondents were married, in their mid-30s, and cared for a household of about six people. No one reported using chlorine tablets prior to our intervention, though the vast majority—86%—indicated that they treated drinking water differently from washing water. Generally, drinking water was treated by straining with a fine cloth. On a scale of one to five regarding agreement level with a particular statement (with five indicating strong agreement), households assigned a level of about four to the statement that "Diarrhea is a major health issue in their locality." Approximately 35% of all households also reported a member having suffered from diarrhea in the past two weeks. A third of surveyed respondents indicated prior participation in a Safa program.

3.2 Door-to-Door Marketing Experiment

The marketing experiment was conducted among 980 households during the months of September and October 2016. For the 712 households that received a baseline survey, the marketing experiment therefore occurred approximately six weeks afterward.[5] Using a systematic tracking process and repeated visits, 92% of our baseline respondents participated in the marketing experiment.[6] Additionally 268 new households[7], not visited at the time of the baseline survey, were sampled among the same areas to also participate in the marketing experiment. These were sampled similarly to the baseline group, using the "right hand rule" approach starting from a random location within the slum.[8]

Each household in the experiment was randomly assigned to receive one of three offers as aforementioned, and summarized below:

Treatment	**Target Good (SafDrink chlorine tablets)**	**Side Good (Good Knight mosquito repellent)**
T1	Rs. 30 (retail price) & buy one, get one free	(not sold)
T2	Rs. 30 (retail price) & buy one, get one free	Rs. 10 (retail price)
T3	Rs. 30 (retail price) & buy one, get one free	Rs. 10 (retail price) & buy one, get one free

[5]While there was some variation in the precise number of days between the baseline survey and experiment due to staff and logistical constraints, 95% of these households were reached within 60 days.

[6]In cases wherein the main survey respondent was not at home when the marketer first visited, the marketer would return to the home at least three times on at least two different days to contact the original respondent. If the respondent was out of town, the surveyor would make an appointment to come back at a better time. If the household had relocated, the marketer obtained information on their new address from neighbors to approach them at the new location. Relocation or travel outside the locality was the principal reason for attrition.

[7]We required about 270 addresses to obtain a sample of 268 households, so there is again minimal selection into the experiment.

[8]In case the household selected during this process had already been sampled at the time of the baseline survey, the house was skipped and replaced by the next third home.

To be clear, the target and side goods were never bundled: in other words, participants could choose to buy the chlorine tablets only, or the mosquito repellent only.

The randomization was stratified by Safa membership and whether the household received a baseline survey to ensure the sample was balanced across these variables (see Figure 2 for a breakdown of each treatment group by strata). Column (4) of Table 1, which reports the p-value for an F-test of joint significance of the sales treatments, suggests that our randomization is balanced across all critical observables such as respondent status, household size, preventive health habits and beliefs. This test is only available for those households surveyed at baseline. Non-surveyed households were added intentionally at the time of the marketing experiment to control for the effect of having previously participated in a baseline survey, as mentioned in Section 3.1.[9]

Nine Safa staff members were trained to conduct the marketing experiment by staff from the Institute for Financial Management and Research according to the experimental procedures.[10] A list of households to be reached, including the name of the respondent, address, and offer category, were provided to the marketers daily. After locating and making contact with the female head of household, marketers were trained to follow a specific sales script. Upon meeting with the participant, marketers first introduced themselves as staff from Safa, conducting door-to-door sales as part of the NGO's promotion of socially beneficial products in the community. Marketers were also provided with leaflets on Safa describing the NGO and its main programs in the area to hand out to participants in case of interest. Each marketer carried the items to be sold in a closed jute bag with the NGO's logo. Depending on the sales treatment to be offered, the marketer took out only the SafDrink product, or both products, to show the respondent.

[9]Due to potential concerns about randomization balance of the additional households sampled during the marketing process, we also conduct our analysis among the baseline group alone as a robustness test. Results from this analysis are presented in column (2) of Table A1.

[10]Safa members were selected from Safa's Kiran Group, a group of women trained and employed by the NGO to promote and sell sanitary care products via door-to-door sales in the community.

To keep information constant across the conditions, the sales pitch surrounding SafDrink was always the same. The marketers provided a short pitch on chlorination as a disinfectant against water-borne diseases and went over usage guidelines for SafDrink, including the product's precautions and limitations. When the household was also randomized to be offered the Good Knight product, marketers similarly introduced the mosquito repellent fast card. Additionally, households in the condition with no side good were informed that Safa sells a suite of preventive health products including mosquito repellent. Participants were then told the price of the items and the presence of any promotional offer. If a sale was made, marketers filled out a receipt including the number of each product sold, amount received, and the signature of the participant. At the end of each day, this information was handed over to the field staff and verified for accuracy; these checks identified 14 households which were missold the allocated treatment and are therefore excluded from our final analysis (see Table 2 for a summary of the final sample size at each survey stage used in our analysis).[11] The marketers were paid a fixed rate per day worked, as well as the commission normally provided by Safa for each product sold. Note that commissions were not provided for the "free" promotional products, so as to maintain the same incentives.[12] We have included the experimental scripts and all surveys in Appendices B through D.

3.3 Follow-Up Survey

The follow-up surveys were conducted by the same team of surveyors used for the baseline surveys. Households were interviewed between two to three weeks following the marketing intervention. This was to ensure that the survey took place during the time

[11]Specifically, four were meant to receive T1 but were offered T2, and 10 were meant to receive T1 but were offered T3. These appeared to be random mistakes as they were spread across five agents on 10 different days. As an additional test, we report the main results on the sample including these 14 households in their assigned treatments in column (1) of Table A1.

[12]Per-product commissions were Rs. 5 for SafDrink and Rs. 2 for Good Knight (1 USD ≈ Rs. 60). We acknowledge the possibility of data integrity concerns as raised in Christensen et al. (2017), but the variance in sales patterns by marketer do not suggest any problems.

households could have still been using SafDrink if they had purchased the product.[13] Though there is some discrepancy in the precise number of days between the experiment and the follow-up survey, 90% of the households which responded to the survey were reached within 19 days.

Of the 980 households in the marketing experiment, 686 were successfully reached for the follow-up survey. Of those households we could not reach, over 80% were repeatedly unavailable for the interview. Respondent availability was largely affected by India's demonetization period in November 2016, which caused severe cash constraints and interrupted the daily schedules of many low-income households.[14] Safa members were slightly more likely to be re-surveyed (significant at the 10% level), and we control for this in our analysis. Note that these tests can only be performed on the non-attrition baseline group, as we did not collect time-invariant data during the follow-up survey; we therefore present results pertaining to this group as a robustness test.[15] We did, however, follow-up with more respondents who purchased SafDrink, and we believe this was because they were likely more willing to engage in door-to-door interactions. In terms of our analysis, this selection at follow-up suggests our results on purchase and usage are likely to be an upper-bound estimate.

The follow-up survey took approximately 20 minutes to cover several topics. First, checks were placed to verify the identity of the respondents. This was followed by a detailed section on water treatment and use of SafDrink, as well as Good Knight for those

[13]Each bottle of SafDrink contained 100 tablets and since the product was sold as a "buy one get one free" promotion, any purchasing household received at least 200 chlorine tablets. If each household member drinks a liter of water at home each day, an average household of six—as in our baseline sample—needs about 42 liters of purified water per week. Each SafDrink tablet purifies a liter of water, so 200 tablets should have lasted about five weeks.

[14]To reduce attrition, 4% of the interviews were conducted with women in the household other than the original primary respondent. Table A2 confirms that this attrition was not specific to any critical household covariate, providing comfort that the follow-up sample was equally balanced.

[15]These are reported in column (2) of Table A1. Importantly, Table A2 also demonstrates that attrition was not specific to any treatment category (also see Table 2 for a summary of the sample size by treatment showing the proportions did not change at each survey stage.)

in the group having received both products. A section on Safa membership, including years of membership, program participation, opinion on the NGO, and expected future involvement were aimed at capturing the importance of the relationship to the respondent. We also included a short section on numeracy[16] and health literacy to obtain information on cognitive abilities including the ability to follow usage guidelines. These factors can be important; for example, Barham et al. (2018) show that cognitive ability predicts early adoption of genetically modified seeds among U.S. farmers.

3.4 Measuring SafDrink Use

We use objective measures of chlorine presence in drinking water to capture SafDrink use. We obtain these measures using drinking water samples collected during each follow-up survey; the surveyors requested the respondent to fill a small container with the stored drinking water available. The household's drinking water was usually stored in an earthen pot containing between 20 and 40 liters of water. The container we used to collect samples was large enough to completely immerse a chemical test strip, which then changes color in different regions based on present chlorine and other chemical concentrations. The test strips indicated levels of free chlorine and total chlorine separately, along with a few other water quality indicators.

The test strips we use identify five possible concentrations of free chlorine: 0, 0.5, 1, 3, 5, and 10 parts per million. Scientifically, when chlorine is added to water, some of it first reacts with organic compounds to form chloramines. The remaining chlorine in water is known as total chlorine and includes both chloramines—the resulting by-product of the first reaction which is not effective in eradicating pathogens—and free chlorine, which kills disease-causing micro-organisms. To be consistent with the literature, we

[16]This variable is an index of 1 to 6 based on the number of numeracy questions answered correctly. There were separate questions for addition, division and multiplication, as well as three sequential questions on subtraction (Q1: "What does 100 minus 7 equal?", Q2: "Minus 7 from that amount?", Q3: "Minus 7 from that amount?").

focus our results on free chlorine but also provide results on other measures; including a binary indicator for free chlorine content, total chlorine level, and the household's survey response to SafDrink use, as robustness tests.[17] Ashraf, Berry and Shapiro (2010), one of the prior papers we follow in this method, highlight the reason for focusing on the measurement of free chlorine in their paper: "We focus on free chlorine, because our own experimentation, as well as conversations with the manufacturer [of the test strips], suggest that the free chlorine measurement is more reliable and less sensitive to variation in test conditions (such as light and heat) than measurement of total chlorine."

The chemical strip test is an imperfect measure of SafDrink use if households receive chlorinated water directly through municipal efforts. Another concern in measuring use via a chemical test is that the pre-treated water contamination level affects the level of free chlorine present after chlorination. Specifically, highly contaminated water may fail to test positive for free chlorine even after being treated with SafDrink. To mitigate these concerns, we conducted detailed water quality tests among the source of water for a randomly selected sub-sample of 50 households across the experimental conditions during baseline. Results from these tests showed the water to be largely unchlorinated and of average contamination from fecal content.[18] Also, since our sample is relatively small and reached a geographically similar area, it is likely that all households had access to similar water sources.

4 Results and Interpretation

We provide a review and discussion of the experimental results in this section. We explore potential theoretical motivations for these results in Appendix B; the motivations we explore center on the extent to which marketing pressure and framing effects may have played a role in the purchase and use decisions.

[17] These are reported in Table A3.

[18] Specifically, we found that about half of the sampled water tested positive for fecal content and a quarter showed some minimal levels of chlorine.

4.1 SafDrink Purchase

We begin by analyzing SafDrink purchase. Figure 3 shows the purchase rate by experimental condition; we observe that across the full sample, about 45% of households purchased the chlorine tablets. Purchase rates are not statistically different in the first two conditions, in which SafDrink is sold alone ($T1$) or alongside the side good at retail price ($T2$). About 42% and 40% buy SafDrink in these conditions, respectively. In the third condition ($T3$), in which SafDrink is sold alongside the side good on a promotion, the purchase rate of SafDrink was 49% (which we find to be statistically different from $T1$ and $T2$).

To capture the effects of the experimental variants on SafDrink purchase, we use estimates from the following regression:

$$\textit{Purchased SafDrink} = \gamma + \beta_1 T2 + \beta_2 T3 + \delta X + \epsilon, \tag{1}$$

where $T2$ and $T3$ represent their associated experimental treatments. The constant term, γ, captures the purchase rate of the reference category (condition $T1$). We follow the suggested method in Bruhn and McKenzie (2009) and include experimental controls for the variables on which we stratified our marketing randomization: whether the household received a baseline survey or had a member previously participate in a Safa program. Additionally, we control for whether it rained on the day of sales, as sales were lower on these days due to logistical challenges faced by the marketers. These variables are all included in the control vector X.

The regression results on SafDrink purchase are presented in Panel A of Table 3. Column (1) includes data from the 966 households having correctly received the allocated sales experiment, and column (2) includes the 686 households that were also available in the follow-up survey; we show results pertaining to this sample for consistency when considering the usage results (only available for those who participated in the follow-up survey).[19] In both samples, SafDrink purchase is not statistically different between conditions $T1$ (the reference category) and $T2$, as expected from the descriptive statistics. In condition $T3$, however, SafDrink purchase increases by about 10.1% in the sales sample and 9.5% in the follow-up sample. Recall that the price of SafDrink was the same across all conditions, so the higher purchase rate in condition $T3$ can only be explained by responses to the promotion on the side product.

Figure 4 shows the composition of purchases by experimental treatment. In condition $T2$, we find that the majority of households bought SafDrink alone rather then with Good Knight, with 31% versus 12% respectively. We find the opposite trend in $T3$, as most households now choose to purchase SafDrink along with Good Knight. To estimate this relationship formally, we analyze the treatment effects on additional measures from our sales experiment; including the probability of purchasing Good Knight and the total amount spent (see Table 4).[20] We find that households in $T3$ were 26.9% more likely to purchase Good Knight. Given that households in $T3$ bought both the promoted mosquito repellent and the chlorine tablets together (instead of using the mosquito repellent as an opt-out), we infer that the promotion on the side good led to the additional purchases of SafDrink.

Part of our motivation for selling the side good alongside the target good was to test whether it served as an "opt-out" option for households not wishing to buy SafDrink but perhaps feeling too guilty to forego any purchase. In Figure 4, we observe that only 2% (in $T2$) and 4% (in $T3$) of households bought the side good, Good Knight, alone. We take this as evidence that the side good did not serve as an "opt-out" product as we had

[19] As a robustness test, we also report these results on the complete sample including the 14 households missold the allocated treatment, as well as the baseline sample alone in columns (1) and (2) of Table A1 respectively.

[20] The estimates are from linear probability models with the same specification as Equation (1), but with different outcome variables.

hypothesized might be the case. More formally, we observe in column (1) of Table 4 that the "opt-out" condition of buying Good Knight alone was not used differentially in $T3$ versus $T2$.

There are a number of reasons for which households could have chosen not to "opt-out". The first is that households did not perceive the pressure to purchase any product from the door-to-door sales agent. This may be because of the urban nature of our experiment; in such an environment, individuals may be more comfortable saying "no" than in rural areas. Another reason could be that the opt-out good did not sufficiently relieve the obligation or pressure to buy the focus good; in the model we provide in Appendix B, this would occur if the price of the opt-out good was higher then the disutility of saying "no". Alternatively, given that we selected our study to take place in an area with high concerns about water quality, there may have been genuine demand for at least testing the chlorine product for preventive health reasons.

Our experimental design allows us to address the validity of alternate theories such as mental accounting and incentives from commissions as reasons for the higher SafDrink purchase rate in $T3$. A mental accounting theory would suggest that households in condition $T3$ would be more likely to buy SafDrink if they bought Good Knight and applied part or all of the promotional amount (Rs. 10) as discount on SafDrink. However, we do not find a difference in the proportion of households buying SafDrink conditional on buying Good Knight.[21] One may also be concerned that the additional commission received when co-selling Good Knight (Rs. 2) may have increased sales for SafDrink in these experimental conditions if the marketers exert more effort when the total amount of possible commission is higher. However, we find no difference in purchase rates between conditions $T1$ and $T2$, showing that this concern is not likely operative in our setting (note again that the marketers were not given commissions on the promotional Good

[21] Our results show that in both $T2$ and $T3$, approximately 85% of buyers who purchase Good Knight also purchase SafDrink.

Knight packets in condition $T3$.)

4.2 SafDrink Use

We next produce and analyze an analogous set of results on SafDrink use. This analysis has the same specification as that for purchase, but the dependent variable is the free chlorine level of the sampled water:

$$\textit{Free Chlorine Level} = \gamma + \beta_1 T2 + \beta_2 T3 + \delta X + \epsilon, \tag{2}$$

where all the covariates are the same as before. The only change is that we add whether it rained on the day of the follow-up survey as an additional experimental control, as rain may have impacted chlorine use for drinking water treatment (this control does not affect our results in a meaningful way, however).

Panel B of Figure 3 shows that the mean level of free chlorine (our main measure of use) was 0.19 across all households. The levels varied by condition, however, and was highest in condition $T2$—a result that persists in the formal analysis that follows. We observe that the mean levels of free chlorine in conditions $T1$, $T2$, and $T3$ were 0.16, 0.36, and 0.10, respectively. The confidence intervals of this measure for conditions $T1$ and $T3$ overlap, so we can only infer that product use was highest in the second condition. The reason for this may be because the households that purchased SafDrink in this condition perceived its value to be higher as it was sold alongside a familiar side good at its known retail price; we explore this hypothesis more fully in Appendix B.

The patterns in Panel B of Figure 3 are consistent with the regression estimates shown in Panel B of Table 3. This analysis captures total use, the distribution of free chlorine unconditional on purchase.[22] We only have data for the sample of households reached for follow-up, so the estimates are presented in column (2). We observe that the free

[22] We also present results on the use of SafDrink conditional on purchase in Table A3. Likely due to household self-selection (and perhaps the smaller sample size, since the coefficients are of similar magnitudes), the main results do not persist among this group.

chlorine level in condition $T2$ was 0.155 higher than in the reference category, an effect size of nearly 100%.[23] In the latter table, the binary measures of chlorine content likely lose too much variation to reliably detect statistical significance, though the coefficients are of consistent direction. Comparing the first and second treatments in both tables, our results show that having a familiar side good present during the sales process strongly impacts product use.

Comparing the second and third treatments, we observe again a marked difference in the level of chlorine tablet usage. Since the only difference between these treatments is the effective price of the mosquito repellent side good—in $T2$ it was sold at full price and in $T3$ it was sold on a "buy one, get one free" promotion, we conclude that it was this promotional offer which changed the selection of households into chlorine purchase and thus use. Conceptually, the promotional offer on the side good, may have impacted the household's valuation of the target chlorine product such that more people were willing to buy it even without intended use. Thus, the increase in purchase but not use in $T3$ is not too surprising.

Another potential explanation for the decreased use in the third condition is that the promotion on the mosquito repellent made the chlorine tablets seem less important by comparison. Such behavior would support our hypothesis that organizations should carefully set prices and promotions when marketing products door-to-door, especially when one of the products is new and unfamiliar. There were also more purchasers of the chlorine tablets in the third condition, so it could be that the marginal buyers were prompted by feelings of reciprocity or guilty due to the promotion on the mosquito repellent. If households engaged in this behavior, they may not have valued the chlorine tablets as much when sold alongside the promoted side good. Unfortunately our experiment was not designed to further identify the underlying mechanisms, but these differences in chlorine

[23] Note that these results are robust to using only the baseline group and excluding outliers on our free chlorine measure, presented in Table A1, as well as a measure of total chlorine reported in Table A3.

tablet use across conditions are statistically significant and practically meaningful.

More generally, we highlight that we find an large proportion of zero free chlorine levels across households having purchased SafDrink across all treatments; 78% in T1, 72% in T2, and 88% in T3 (see Figure 5); this gap between product purchase and use is not surprising in the context of prior literature, particularly Ashraf, Berry and Shapiro (2010). One possibility is that households may have been saving the tablets for later use, or learned that they dislike the product (e.g., as was true for chlorine tablet taste or mosquito net comfort in prior studies Ashraf, Berry and Shapiro 2010 and Dupas 2009), but still the gap between the purchase and use of preventive health products deserves further study.

4.3 Heterogeneity Analysis

In this section, we examine whether identifiable household characteristics can explain the differences in SafDrink purchase and use across the experimental conditions. To conduct this analysis, we include interaction terms between key observables and each treatment category to the regression specification in Equation (1).[24] The observables we test are the following: whether the household has a Safa member, whether the household received a baseline survey, household size, an index of numeracy,[25] whether the household already treats drinking water in some manner, and whether any household member had experienced diarrhea in the past two weeks. As our experiment was not designed to uncover such heterogeneity, our data are unable to offer much insight through this analysis. In particular, this is the case because our sample size is reduced when we test for certain observables as not all households received the baseline survey.

The results of this analysis with respect to SafDrink purchase and usage are presented

[24] Specifically, the model we estimate for testing heterogeneity is: $Outcome = \gamma + \beta_1 T2 + \beta_2 (T2 \times Observable) + \beta_3 T3 + \beta_4 (T3 \times Observable) + \beta_5 Observable + \delta X + \epsilon$

[25] Numeracy was measured on a six-point index based on the number of correctly answered questions; these covered basic addition, division and multiplication, as well as three sequential questions on subtraction.

in Tables 5 and 6, respectively. Focusing on the interaction effects in our interpretation, we observe few statistically significant effects so refrain from interpreting most findings. We observe that households with more members were more likely to be using SafDrink, perhaps because water quality is of greater concern for children. We also document statistically significant effects with respect to the incidence of diarrhea among household members; suggesting that the household may only treat their drinking water with chlorine in response to a health shock related to water quality.

5 Conclusion

The goal of our experiment was to test whether simple and inexpensive changes to the door-to-door sales process can increase adoption—purchase and use—of an unfamiliar preventive health product. Our research shows that marketing a target good alongside a second side good helps increase the total use of the target good. This increase in usage only occurs if the second good is sold at its retail price, however, as promotions or discounts on this good may reintroduce other unintended behaviors. The target good we studied was SafDrink, a brand of water chlorination tablets, and the side good was Good Knight, a brand of mosquito repellent product. This intervention nearly doubled use of the target good without impacting the amount of purchases, relative to when the target good was sold alone.

One hypothesis that is interesting but we cannot test is whether the use of Good Knight crowded out the use of SafDrink. If this were true, it could help explain some of the difference in SafDrink use between T2 and T3. (Though T1 might be expected to induce much greater chlorine use than observed under this theory.) We had specifically chosen a mosquito repellent as the side good to the water purification chlorine tablets so that they are not substitutes, but they could be viewed as such if participants held a budget for preventive health behaviors generally. Future work could explore this question.

There are many other avenues for future work. The concept of marketing pressure is inextricably linked to implementer identity, which has been shown to matter in scaling up efforts (Cameron, Olivia and Shah 2019). Potentially, varying the implementer identity would yield different levels of marketing pressure that then impact product purchase and use. Future work could also examine the effect of different side goods, building on the work by Ashraf, Jack and Kamenica (2013) in which the authors increase take-up of a chlorine tablets by leveraging the interaction of price and information effects. Improving the purchase and use of socially beneficial goods in developing countries is an important effort that requires continued refinement, and we hope that the present analysis helps push forward the design in door-to-door marketing for these goods.

References

Ashraf, Nava, James Berry and Jesse M. Shapiro. 2010. "Can higher prices stimulate product use? Evidence from a field experiment in Zambia." *American Economic Review* 100:2383–2413.

Ashraf, Nava, Kelsey B. Jack and Emir Kamenica. 2013. "Information and subsidies: Complements or substitutes?" *Journal of Economic Behavior & Organization* 88:133–139.

Ashraf, Nava, Oriana Bandiera and Kelsey B. Jack. 2014. "No margin, no mission? A field experiment on incentives for public service delivery." *Journal of Public Economics* 120:1–17.

Barham, Bradford L., Jean-Paul Chavas, Dylan Fitz and Laura Schechter. 2018. "Receptiveness to Advice, Cognitive Ability, and Technology Adoption." *Journal of Economic Behavior & Organization* 149(May 2018):239–268.

Brass, Jennifer N., Wesley Longhofer, Rachel S. Robinson and Allison Schnable. 2018. "NGOs and international development: A review of thirty-five years of scholarship." *World Development* 112:136–149.

Bruhn, Miriam and David McKenzie. 2009. "In pursuit of balance: Randomization in practice in development field experiments." *American Economic Journal: Applied Economics* 1(4):200–232.

Cameron, Lisa, Susan Olivia and Manisha Shah. 2019. "Scaling up sanitation: Evidence from an RCT in Indonesia." *Journal of Development Economics* 138:1–16.

CDC. 2017. "Center for Disease Control and Prevention, Global WASH Fast Facts." https://www.cdc.gov/healthywater/global/wash_statistics.html. Accessed: 2017-03-13.

Christensen, Lisa Jones, Enno Siemsen, Oana Branzei and Madhu Viswanathan. 2017. "Response pattern analysis: Assuring data integrity in extreme research settings." *Strategic Management Journal* 38(2):471–482.

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

Cutler, David and Grant Miller. 2005. "The role of public health improvements in health advances: The twentieth-century United States." *Demography* 42(1):1–22.

Das, Ashis, Jed Friedman and Eeshani Kandpal. 2018. "Does involvement of local NGOs enhance public service delivery? Cautionary evidence from a malaria-prevention program in India." *Health Economics* 27(1):172–188.

DellaVigna, Stefano, John A. List and Ulrike Malmendier. 2012. "Testing for altruism and social pressure in charitable giving." *Quarterly Journal of Economics* 127(1):1–56.

Dupas, Pascaline. 2009. "What matters (and what does not) in households' decision to invest in malaria prevention?" *American Economic Review* 99(2):224–230.

Dupas, Pascaline. 2014. "Short-run subsidies and long-run adoption of new health products: Evidence from a field experiment." *Econometrica* 82(1):197–228.

Dupas, Pascaline, Vivian Hoffmann, Michael Kremer and Alix Peterson Zwane. 2016. "Targeting health subsidies through a nonprice mechanism: A randomized controlled trial in Kenya." *Science* 353(6302):889–895.

Fischer, Greg, Dean S. Karlan, Margaret McConnell and Pia Raffler. 2014. "To charge or not to charge: Evidence from a health products experiment in Uganda." *NBER Working Paper No. w20170* .

Günther, Isabel and Youdi Schipper. 2013. "Pumps, germs and storage: the impact of improved water containers on water quality and health." *Health Economics* 22(7):757–774.

Hanna, Rema, Esther Duflo and Michael Greenstone. 2016. "Up in smoke: the influence of household behavior on the long-run impact of improved cooking stoves." *American Economic Journal: Economic Policy* 8(1):80–114.

Hoffmann, Vivian, Christopher B. Barrett and David R. Just. 2009. "Do free goods stick to poor households? Experimental evidence on insecticide treated bednets." *World Development* 37(3):607–617.

Kumar, Santosh and Sebastian Vollmer. 2013. "Does access to improved sanitation reduce childhood diarrhea in rural India?" *Health Economics* 22(4):410–427.

Lee, Sang-Hyop. 2005. "Demand for immunization, parental selection, and child survival: Evidence from rural India." *Review of Economics of the Household* 3(2):171–196.

Lipscomb, Molly and Laura Schechter. 2018. "Subsidies versus mental accounting nudges: Harnessing mobile payment systems to improve sanitation." *Working paper* .

Luoto, Jill, David Levine, Jeff Albert and Stephen Luby. 2014. "Nudging to use: Achieving safe water behaviors in Kenya and Bangladesh." *Journal of Development Economics* 110:13–21.

Meredith, Jennifer, Jonathan Robinson, Sarah Walker and Bruce Wydick. 2013. "Keeping the doctor away: experimental evidence on investment in preventative health products." *Journal of Development Economics* 105:196–210.

Mohapatra, Sandeep and Leo Simon. 2017. "Intra-household bargaining over household technology adoption." *Review of Economics of the Household* 15(4):1263–1290.

Oster, Emily and Rebecca Thornton. 2012. "Determinants of technology adoption: Peer effects in menstrual cup take-up." *Journal of the European Economic Association* 10(6):1263–1293.

Tarozzi, Alessandro, Aprajit Mahajan, Brian Blackburn, Dan Kopf, Lakshmi Krishnan and Joanne Yoong. 2014. "Micro-loans, insecticide-treated bednets, and malaria: Evidence from a randomized controlled trial in Orissa, India." *American Economic Review* 104(7):1909–1941.

Zwane, Alix Peterson, Jonathan Zinman, Eric Van Dusen, William Pariente, Clair Null, Edward Miguel, Michael Kremer, Dean S. Karlan, Richard Hornbeck, Xavier Giné and others. 2011. "Being surveyed can change later behavior and related parameter estimates." *Proceedings of the National Academy of Sciences* 108(5):1821–1826.

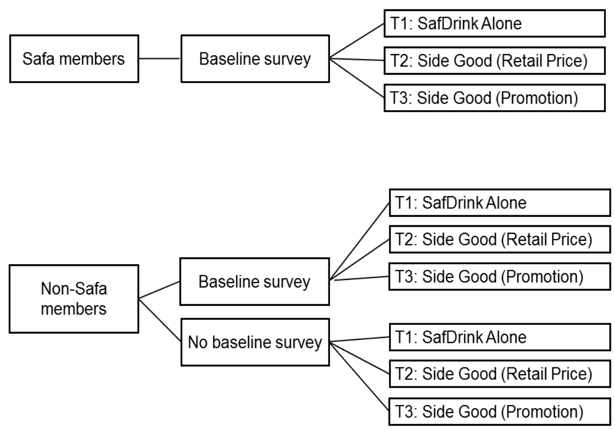
Figure 1: Experimental Goods

Panel A: Target Good—SafDrink

Panel B: Side Good—Good Knight

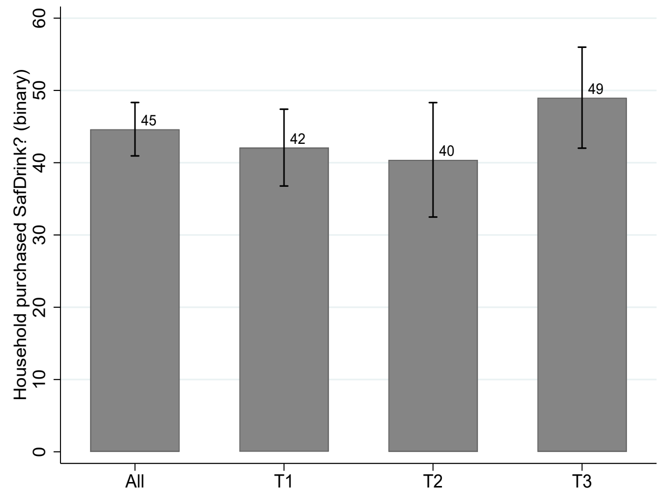
Notes: Panel A shows a bottle of SafDrink containing 100 tablets; this was the unit sold in the experiment, always on a "buy one get one free" promotion. The tablets are meant to purify drinking water; each bottle's price was Rs. 30 (about USD 0.50). Panel B shows a packet of Good Knight fast cards, and each packet sold in the experiment contained 10 cards. The cards are meant to be burnt to repel mosquitoes; each packet's price was Rs. 10 (about USD 0.20).

Figure 2: Experimental Design

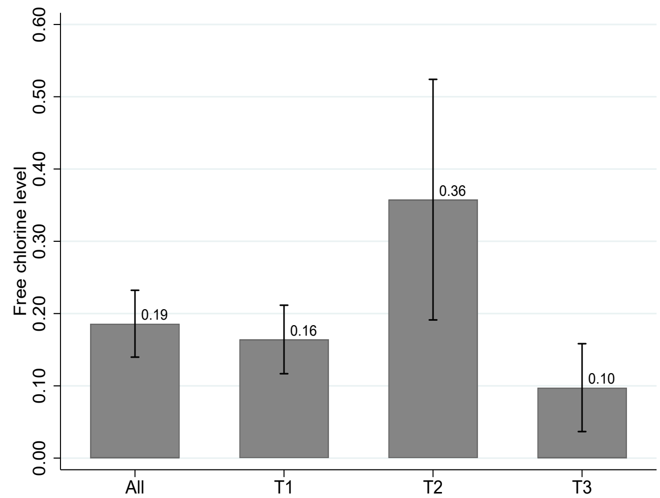


Notes: Households were randomized on prior participation in a Safa program, then on whether they received a baseline survey, and finally on the experimental treatment. All households with prior participation in a Safa program received the baseline survey.

Figure 3: SafDrink Purchase and Use by Experimental Treatment at Follow-Up



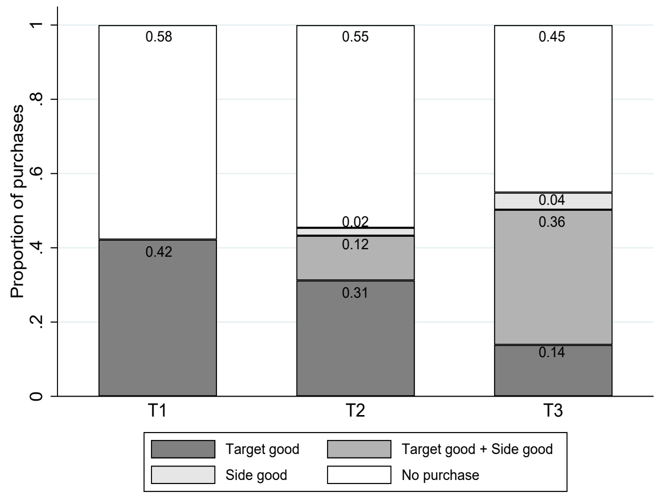
Panel A: Purchase of SafDrink



Panel B: Use of SafDrink

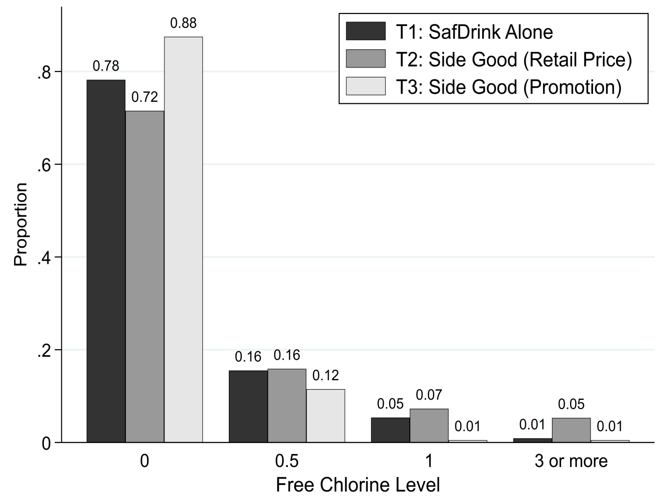
Notes: Bars show mean proportions with 95% confidence interval bands. $T1$: SafDrink Alone; $T2$: Side Good (Retail Price); and $T3$: Side Good (Promotion).The sample is restricted to respondents reached for the follow-up survey ($N = 686$).

Figure 4: Composition of Purchases by Experimental Treatment at Follow-Up

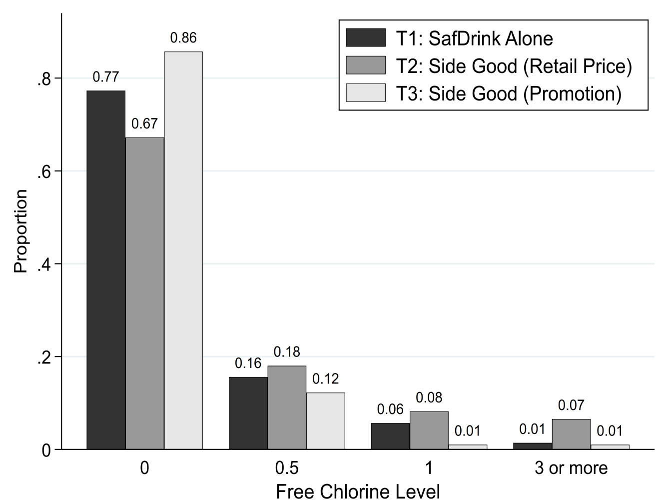


Notes: Bars show proportion of each purchase category. $T1$: Saf-Drink Alone; $T2$: Side Good (Retail Price); and $T3$: Side Good (Promotion). The sample is restricted to respondents reached for the follow-up survey($N = 686$).

Figure 5: Free Chlorine Distribution by Experimental Treatment at Follow-Up



Panel A : Unconditional on SafDrink Purchase



Panel B : Conditional on SafDrink Purchase

Notes: Columns show proportion of treatment samples across free chlorine levels tested at follow-up. The test strips identified five possible concentrations of free chlorine: 0, 0.5, 1, 3, 5, and 10 parts per million. Levels above 3 were rare: 13 in Panel A and 8 in Panel B. Panel A contains all respondents reached for the follow-up survey ($N = 686$); Panel B restricts to that sample conditional on SafDrink purchase ($N = 300$).

Table 1: Summary Statistics and Covariate Balance across Experimental Treatment at Baseline

	$T1$: SafDrink Alone (1)	$T2$: Side Good (Retail Price) (2)	$T3$: Side Good (Promotion) (3)	N	p-value (4)
Preventive health habits					
Use of soap before handling food[a]	2.54	2.52	2.64	698	0.40
	(1.03)	(1.01)	(0.91)		
Separate drinking water	0.87	0.84	0.87	698	0.55
	(0.33)	(0.37)	(0.34)		
Water treated with chlorine	0.00	0.00	0.00	698	—
	(0.00)	(0.00)	(0.00)		
Respondent demographics					
Age	35.34	36.13	35.81	697	0.73
	(11.06)	(10.85)	(11.43)		
Years of completed schooling	6.50	5.75	6.51	697	0.17
	(4.36)	(4.83)	(4.17)		
Currently married	0.80	0.84	0.81	698	0.66
	(0.40)	(0.37)	(0.39)		
Currently employed	0.28	0.31	0.21	698	0.11
	(0.45)	(0.47)	(0.41)		
Safa member	0.37	0.42	0.32	698	0.17
	(0.48)	(0.50)	(0.47)		
Respondent beliefs/attitudes					
Diarrhea is a major health issue[b]	3.86	3.75	4.07	689	0.20
	(1.68)	(1.71)	(1.57)		
Outside water is safe[c]	1.16	1.08	1.14	695	0.11
	(0.44)	(0.32)	(0.41)		
Trust in NGOs[d]	3.91	3.82	3.94	690	0.31
	(1.44)	(1.55)	(1.41)		
Household characteristics					
Number of household members	5.99	5.96	6.49	698	0.14
	(2.59)	(3.39)	(3.32)		
Diarrhea in past 2 weeks[e]	0.36	0.36	0.34	698	0.81
	(0.48)	(0.48)	(0.48)		
Concrete house	0.53	0.54	0.55	698	0.88
	(0.50)	(0.50)	(0.49)		
House has direct water pipe	0.85	0.88	0.83	698	0.43
	(0.36)	(0.33)	(0.37)		

Notes: Columns (1), (2), and (3) report variable means; standard errors are in parentheses. Column (4) reports the p-values associated with the F-test that all regressors equal zero from individual linear regressions. Each regression is a linear regression of the covariate on dummies for the sales treatments, controlling for whether a household member is associated with Safa [$Covariate = \gamma + \beta_1 T2 + \beta_2 T3 + \delta Safa + \epsilon$]. The sample is restricted to respondents reached for the baseline survey and who received the allocated experimental condition ($N = 698$).

[a]Index of use: 1(never) to 4(always).

[b]Index of belief: 1(strongly disagree) to 5(strongly agree).

[c]Index of safety: 1(always) to 3(never).

[d]Index of trust: 1(no trust) to 5(great deal of trust).

[e]Binary variable indicating whether the household reported a member experiencing diarrhea in the past two weeks.

Table 2: Sample Size by Experimental Treatment and Phase

	T1: SafDrink Alone (1)	T2: Side Good (Retail Price) (2)	T3: Side Good (Promotion) (3)	All (4)
Baseline	342 (49%)	140 (20%)	216 (31%)	698 (100%)
Sales	476 (49%)	195 (20%)	295 (31%)	966 (100%)
Follow-up	335 (49%)	151 (22%)	200 (29%)	686 (100%)
Baseline and Follow-up	275 (49%)	112 (20%)	170 (31%)	557 (100%)

Notes: Sample sizes presented here (and used in the analysis) exclude the 14 households that did not receive the allocated experimental condition, though we include these households in a robustness check in Table A1. At each phase, the sample is proportional by treatment: 0.5 (T1: SafDrink Alone), 0.2 (T2: Side Good (Retail Price)), and 0.3 (T3: Side Good (Promotion)).

Table 3: Impact of Experimental Treatment on SafDrink Purchase and Use

All households participating in:	Sales Experiment (1)	Follow-up Survey (2)
Panel A: SafDrink Purchase		
Dependent variable: Household purchased SafDrink? (binary)		
$T2$: Side Good (Retail Price)	-0.002	-0.048
	(0.042)	(0.049)
$T3$: Side Good (Promotion)	0.101***	0.095**
	(0.037)	(0.045)
Dep. var. mean ($T1$: SafDrink Alone)	0.351	0.421
Panel B: SafDrink Usage		
Dependent variable: Free chlorine level in sampled drinking water		
$T2$: Side Good (Retail Price)		0.155**
		(0.063)
$T3$: Side Good (Promotion)		-0.015
		(0.060)
Dep. var. mean ($T1$: SafDrink Alone)		0.164
Number of observations	966	686
Other experimental controls	Yes	Yes

Notes: Estimates are from linear models; standard errors in parentheses. Each regression is a linear regression of the dependent variable on the sales treatment dummies and experimental controls; specifically Panel A refers to the model outlined in Equation 1 and Panel B to Equation 2. The reference category is the treatment with no side good ($T1$). "Other experimental controls" include whether a household member had prior participation in a Safa program, whether that household received a baseline survey, whether it rained on the day of sales, and whether it rained on the day of the follow-up survey (this last control is only added to the usage regressions).

Table 4: Impact of Experimental Treatment on Composition of Purchases

Dependent variable:	Bought Good Knight alone (binary) (1)	Bought Good ? Knight (binary) (2)	Bought any product (binary) (3)
Panel A: Sales sample			
$T2$: Side Good (Retail Price)			0.019
			(0.042)
$T3$: Side Good (Promotion)	0.029	0.269***	0.161***
	(0.022)	(0.047)	(0.037)
Dep. var. mean (reference group)	0.015	0.126	0.351
Number of observations	490	490	966
Other experimental controls?	Yes	Yes	Yes
Panel B: Follow-up sample			
$T2$: Side Good (Retail Price)			-0.026
			(0.049)
$T3$: Side Good (Promotion)	0.011	0.272***	0.135***
	(0.022)	(0.054)	(0.045)
Dep. var. mean (reference group)	0.020	0.142	0.421
Number of observations	351	351	686
Other experimental controls?	Yes	Yes	Yes

Notes: Estimates are from linear models; standard errors in parentheses. Each regression is a linear regression of the dependent variable on the sales treatment dummies and experimental controls; specifically the analysis refers to the model outlined in Equation 1. The reference category for columns (1) and (2) is $T2$ whereas for column (3) the reference category is $T1$. "Other experimental controls" include whether a household member had prior participation in a Safa program, whether that household received a baseline survey, and whether it rained on the day of sales.

Table 5: Heterogeneity in SafDrink Purchase at Follow-Up by Experimental Treatment

Observable:	Safa member (binary) (1)	Baseline participation (binary) (2)	Number of children (continuous) (3)	Numeracy (index, 1-6)[a] (4)	Drinking water is treated (binary)[b] (5)	Diarrhea in past 2 weeks (binary)[c] (6)
Dependent variable: Household purchased SafDrink? (binary)						
$T2$: Side Good (Retail Price)	-0.079	-0.115	0.000	0.049	0.064	-0.021
	(0.059)	(0.100)	(0.065)	(0.193)	(0.115)	(0.070)
$T2$: Side Good (Retail Price)×Observable	0.090	0.091	-0.048	-0.022	-0.124	-0.019
	(0.103)	(0.114)	(0.055)	(0.044)	(0.131)	(0.114)
$T3$: Side Good (Promotion)	0.105*	0.320***	0.090	-0.197	-0.018	0.060
	(0.054)	(0.112)	(0.057)	(0.143)	(0.104)	(0.059)
$T3$: Side Good (Promotion)×Observable	-0.027	-0.262**	-0.042	0.065**	0.099	0.010
	(0.095)	(0.120)	(0.043)	(0.033)	(0.118)	(0.101)
Observable	0.022	-0.182**	0.017	-0.032	0.048	0.004
	(0.060)	(0.071)	(0.011)	(0.021)	(0.069)	(0.062)
Observable mean (All)	0.309	0.815	0.637	4.140	0.780	0.353
Sample	Follow-up	Follow-up	Follow-up	Follow-up	Follow-up	Follow-up
Number of observations	686	686	557	604	556	557
Other experimental controls?	Yes	Yes	Yes	Yes	Yes	Yes

Notes:Estimates are from linear models; standard errors in parentheses. The regression equation corresponding to the estimates is: $[Purchased\ SafDrink = \gamma + \beta_1 T2 + \beta_2(T2 \times Observable) + \beta_3 T3 + \beta_4(T3 \times Observable) + \beta_5 Observable + \delta X + \epsilon]$. The reference category is the treatment with no side good ($T1$). "Other experimental controls" include whether whether a household member had prior participation in a Safa program, whether that household received a baseline survey, and whether it rained on the day of sales. Sample for columns (3), (5), and (6) is restricted to respondents at follow-up having participated in the baseline survey ($N = 557$).
[a]Numeracy is an index of 1 to 6 based on the number of numeracy questions answered correctly. There was a separate question for addition, division and multiplication, as well as 3 sequential questions on subtraction.
[b]Binary variable indicating whether the household treats their drinking water.
[c]Binary variable indicating whether the household reported a member experiencing diarrhea in the past two weeks.

Table 6: Heterogeneity in SafDrink Use at Follow-Up by Experimental Treatment

Observable:	Safa member (binary) (1)	Baseline participation (binary) (2)	Number of children (continuous) (3)	Numeracy (index, 1-6)[a] (4)	Drinking water is treated (binary)[b] (5)	Diarrhea in past 2 weeks (binary)[c] (6)
Dependent variable: Free chlorine level in sampled drinking water						
$T2$: Side Good (Retail Price)	0.165**	0.036	0.271***	0.483*	-0.107	0.293***
	(0.076)	(0.127)	(0.084)	(0.249)	(0.149)	(0.091)
$T2$: Side Good (Retail Price)×Observable	-0.030	0.156	-0.121*	-0.070	0.393**	-0.252*
	(0.131)	(0.145)	(0.070)	(0.056)	(0.168)	(0.145)
$T3$: Side Good (Promotion)	-0.017	-0.061	0.019	-0.022	-0.130	0.019
	(0.071)	(0.145)	(0.076)	(0.185)	(0.135)	(0.078)
$T3$: Side Good (Promotion)×Observable	0.005	0.054	-0.042	0.006	0.150	-0.094
	(0.120)	(0.153)	(0.055)	(0.043)	(0.151)	(0.129)
Observable	-0.076	-0.154*	0.006	0.022	-0.109	0.045
	(0.076)	(0.091)	(0.037)	(0.028)	(0.089)	(0.079)
Observable mean (All)	0.309	0.815	0.637	4.140	0.780	0.353
Sample	Follow-up	Follow-up	Follow-up	Follow-up	Follow-up	Follow-up
Number of observations	686	686	557	604	556	557
Other experimental controls?	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Estimates are from linear models; standard errors in parentheses. The regression equation corresponding to the estimates is: [$\textit{Free Chlorine Level} = \gamma + \beta_1 T2 + \beta_2(T2 \times \textit{Observable}) + \beta_3 T3 + \beta_4(T3 \times \textit{Observable}) + \beta_5 \textit{Observable} + \delta X + \epsilon$]. The reference category is the treatment with no side good ($T1$). "Other experimental controls" include whether whether a household member had prior participation in a Safa program, whether that household received a baseline survey, whether it rained on the day of sales, and whether it rained on the day of the follow-up survey. Sample for columns (3), (5), and (6) is restricted to respondents at follow-up having participated in the baseline survey ($N = 557$).

[a]Numeracy is an index of 1 to 6 based on the number of numeracy questions answered correctly. There was a separate question for addition, division and multiplication, as well as 3 sequential questions on subtraction.

[b]Binary variable indicating whether the household treats their drinking water.

[c]Binary variable indicating whether the household reported a member experiencing diarrhea in the past two weeks.