What is the effect of short-term subsidies on demand for potable water in rural Bihar, India?
A randomized controlled trial

Pre-Analysis Plan

27 June 2019

Drew B. Cameron, MA
PhD Candidate, Health Policy
Department of Health Policy and Management
UC Berkeley School of Public Health

William H. Dow, PhD, M.Phil, MA
Henry J. Kaiser Professor of Health Economics,
Department of Health Policy and Management
UC Berkeley School of Public Health

1. Introduction

The following study is designed to test the impact of short-term subsidies on demand for potable water delivery among the rural poor in the Supaul district of Bihar. Specifically, it follows previous research on providing short-term subsidies for health prevention products in low- and middle-income countries in seeking to determine the mechanisms through which subsidy provision operates in influencing future purchase decisions (Dupas 2014; Fischer et al. 2018). Short term subsidies may induce *anchoring effects* – in which recipients reference the low price they previously faced, undervaluing the product on future offer, or *learning effects* – wherein recipients experience the positive benefits of the product, leading to higher future valuation and use. This study seeks to determine which of these mechanisms dominate.

To test for these potential mechanisms, this study explores the promotion of a novel health product – potable water delivery to 20-liter jugs – recently introduced in the Supaul region of rural Bihar by a local NGO and growing in popularity throughout states in north central India. Herein we describe a modified approach to the current delivery schemes in which families must first purchase the hardware required to receive water deliveries. This randomized controlled trial proposes to enroll families who are not current water delivery subscribers, in a price subsidy experiment in which beneficiary households will be randomly sorted into one of two arms for a month-long intervention: a) a high subsidy arm (rough 50% subsidy on the regular price of water for four weeks) or, b) a no subsidy arm (100% of the regular delivery cost for one month). Both arms will receive initial group-based health messaging. To further examine the impact of subsidies on water consumption, we examine a secondary research question to determine whether households who receive the product experience positive or negative learning about the product that influences their subsequent demand for the product after the subsidy period has ended.

This pre-analysis plan is organized into the following sections. Section 1 outlines our main research questions and hypotheses. Section 2 describes some of the early findings from our pilot study. Section 3 outlines the preparation work that will take place before the study can be implemented including survey mapping, listing and enumerator training. Section 4 describes the baseline survey, social marketing intervention, and random price offer which are all conducted as part of the baseline survey. Section 5 describes the auditing of households to verify household water deliveries are taking place and capture validated consumption information. Section 6 describes the follow-up survey and product offer. Section 7 outlines the power required to achieve our desired minimum detectable effect size and the analytical strategy we plan to employ to examine the effects of subsidies on demand for potable water.

2. Project aims, research questions and hypotheses

The primary aim of this randomized controlled trial is to determine the impact of short-term subsidies on demand for water delivery after subsidies have ended. Herein, we employ a simple strategy of providing a random offer price (either subsidized or not) to eligible households in several communities in the Supaul district of rural Bihar, India. The design is outlined in Figure 1, below beginning with the mapping of 516 eligible households, followed by a baseline survey. At the end of the baseline survey a team of social marketers from the local community will explain the water delivery product being sold as well as some of the health benefits of purchasing filtered water. Households will then be randomized into a discounted price group (treatment) or regular price group (control) for a period of 4 weeks. All households in the treatment and control group will be allowed to purchase water at their assigned rates during the intervention period. At the end of the intervention period, we will return to the same 516 households for an end-line survey and make an additional offer for water deliveries at market price. We will track consumption (purchase) among these households for the following three weeks.

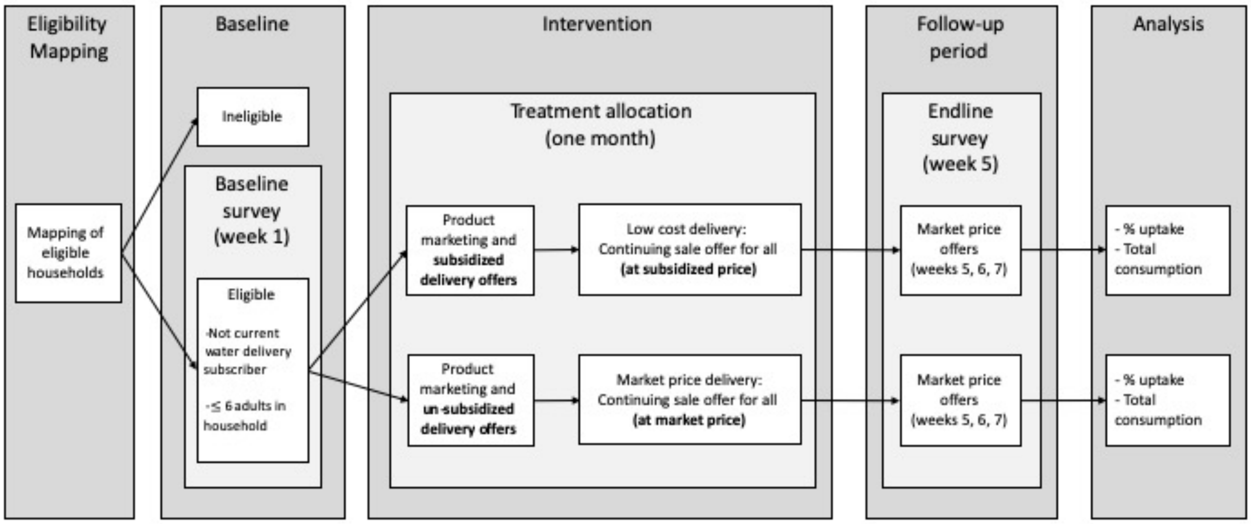

Figure 1. Research design

Conceptual diagram

In order to explore the dynamic of the impact of price subsidies on future consumption, we use the following simplified conceptual diagram in Figure 2. First, households in the the treatment group are offered the product at a subsidized rate at time 1. Respondents must choose whether to take up the subsidized product or not. Indeed, depending on the amount of the price subsidy and the household's priors about the quality of the expected utility returns of the good, the household may *not* choose to purchase. In our case, we assume that households modeled in Figure 2 below face a reduced price for the good at time 1 and have salient and high levels of information about the good provided to them during the intervention marketing visit (described below), thus we assume that a large number of households (though not all) will take up the product to allow for an intent-to-treat analysis.

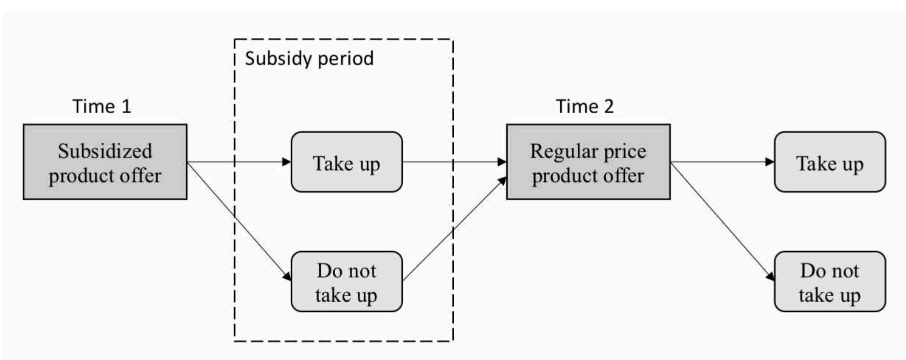

Figure 2. Decision model

In time 2, after the subsidy period has ended, all treatment households are revisited, offered a positive price for the product and must decide whether to purchase at the increased price. It is this purchase decision that is the primary interest of our analysis. Households in this period engage in a decision-making process that is based on several factors, including but not limited to: i) the price of the good, ii) information they have gathered about that good and its utility during the reduced-price provision of the good in the intervention period (updating any priors they may have had about the good), and iii) the price of the good vis-à-vis the price they previously faced (in other words, the gain-loss utility). However, thinking of this analysis at two discrete time points limits the extent to which we might explore multiple pathways through which subsidies can impact future demand for products.

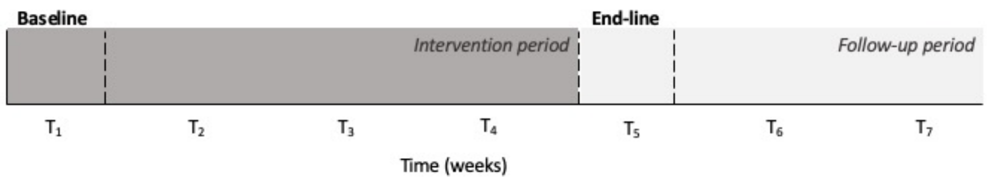

Figure 3. Intervention and measurement period

The study period will last seven weeks (see Figure 3 above), beginning with a baseline survey in week 1. It is during this initial visit that households will be randomly divided into treatment versus control conditions through random price draws (market rate versus subsidized rate drawn from a hat), thus all participants will know that they had an equal, random chance of receiving a low or full price for the product for the first four weeks. All households will be revisited for end-line interviews at week 5, immediately after the subsidy period has ended, and offered full price purchase of water deliveries. Deliveries will also be tracked in each of the seven weeks, and the product will be available to all households in treatment and control groups during all seven weeks of the study. Our primary research question (below) will cover the full study period by comparing uptake between treatment and control groups both at week 7 to see if uptake changes over time.

Primary Research Question:
What is the impact of subsidies for potable water delivery subscriptions on future uptake of subscriptions once subsidies have ended?

Hypotheses:
H_0*: There will be no difference in uptake of water delivery subscriptions between the subsidized and the non-subsidized group after subsidies end, suggesting that neither anchoring nor positive learning effects dominate.*

H_1*: Uptake of water delivery subscriptions will be higher in the subsidized group compared to the no-subsidy group after subsidies end, suggesting that positive learning effects dominate.*

H_2*: Uptake of water delivery subscriptions will be lower in the subsidized group compared to the no-subsidy group after subsidies end, suggesting that anchoring or negative learning effects dominate.*

Sub-aims

It is difficult to formally disentangle learning and anchoring effects when comparing uptake only between the baseline and follow-up period. Indeed, previous studies dealing with more durable products have been forced to rely on one-time product offers at baseline versus end-line in order to examine the effect of subsidies on future demand for products requiring only one-time purchase (Dupas 2014; Fischer et al. 2018). In this study, we benefit both from participant households being located within the new catchment area of our partner NGO (who has built a new water plant completed in June 2019) and from the ability to regularly revisit families and track their consumption, which repeats at regular intervals among consumers. Thus, we will exploit repeated observations in an attempt to quantify these two effects separately. Rather than formally test hypotheses around these sub-aims, by examining trends in uptake across each week of the study we can begin to draw several important conclusions that may help us to better understand the dynamic of learning versus anchoring as mechanisms through which subsidies act on uptake.

First, we will observe and quantify change in uptake due to learning effects among the treatment group during the subsidy period by comparing any change in uptake from week 1 to week 4 (in each week before the subsidies expire). Because the price shock has not happened among the treatment group by the last week of the subsidy period, any change in uptake over this period should be a reflection of learning effects only – either increased uptake from positive learning and spillovers or decreased uptake from negative learning. This leads to our first research sub-question:

Research sub-question 1:
Among the treatment group, how does the proportion of up-takers change during the 4-week subsidy period related to learning effects?

Hypotheses:
H_0*: There is no change in uptake during the subsidy period, suggesting that learning effects have no relationship with uptake among treatment households*

H_1*: There is a positive change in uptake during the subsidy period, suggesting that learning effects have a positive relationship with uptake among treatment households*

H_2*: There is a negative change in uptake during the subsidy period, suggesting that learning effects have a negative relationship with uptake among treatment households*

Among the subsidized treatment group, if we observe no change in uptake during the subsidy period, then we will fail to reject H_0, suggesting that any learning effects were not strong enough to influence uptake during this period. If we observe a positive, statistically significant change in uptake over the period, we will fail to reject H_1, suggesting that learning effects may have a positive relationship with uptake during the subsidy period.[1] Alternatively, if we observe a statistically significant decrease in uptake over this period, this could suggest that negative learning effects had a negative relationship with uptake over the period. We can interpret the coefficient on any week as the proportion of the change in uptake related to learning. Rather than take the mean slope across all weeks during the intervention period, we intend to track week-to-week change over this time frame (i.e. uptake in week 1, week 2, week 3 and week 4). Tracking change in uptake during each week will allow us to examine whether uptake changes more in certain weeks than in others, since we might expect to see uptake decrease more between weeks 2 and 3 (when hardware payments are due) than at any other point during the subsidy period. Gaining a clear picture of

[1] We have little reason to expect to see an increase in uptake among the treatment group during this period (suggesting some kind of spillover of positive learning between neighbors), though it remains a possibility.

week-to-week changes during this subsidy period will be essential for interpreting overall study results as they relate to the timing of learning effects.

Next, we turn our attention to the period immediately before and after the price change (from week 4 to week 5) where we seek to observe anchoring effects among the treatment group. Though any observed change in uptake from week 4 to week 5 among treatment households does not negate the presence of negative or positive learning effects over that time frame, we can expect that most learning effects will have taken place during prior weeks. Thus, we can quantify anchoring effects among these households by comparing uptake at week 4 to week 5 (for the immediate impact of a price shock). This leads to our second research sub-question:

Research sub-question 2:
Among treatment households, what is the proportion of change in uptake between week 4 and week 5 (after the price shock) related to anchoring effects?

Hypotheses:
H_0*: There is no change in uptake from week 4 to week 5, suggesting that there is no relationship between change in price and uptake among treatment households*

H_1*: There is a positive change in uptake from week 4 to week 5, suggesting that there is a positive relationship between change in price and uptake among treatment households*

H_2*: There is a negative change in uptake from week 4 to week 5, suggesting that there is a positive relationship between change in price and uptake among treatment households*

If we do not observe a statistically significant difference in uptake between weeks 4 and 5, we will fail to reject H_0, suggesting that there is no relationship between the price change and uptake. If we observe a positive, statistically significant change in uptake from week 4 to 5 we would fail to reject H_1, and a negative, statistically significant change in uptake would lead us to fail to reject H_2. In any case, we can interpret the difference in coefficients from week 4 to week 5 as the proportion of the change in uptake likely related to anchoring effects.

It is worth reiterating that any observed change in uptake from week 4 to week 5 among treatment households does not negate the presence of negative or positive learning effects over that time frame. However, we could assume that most learning should have been washed out (or already led to changes in uptake) during the subsidy period between weeks 1 and 4, and any observed change in uptake between week 4 and 5 would be primarily due to anchoring effects. However, our ability to learn anything about possible anchoring versus learning from the change between week 4 and 5 will depend largely on the slope in uptake from week 1 to week 4.

It is also worth note that for either of these sub aims, there may be difference in change in uptake between the treatment and control groups (a dynamic that is more akin to our primary research question) at any point in the first 5 weeks, however a differential change does not suggest that positive or negative learning did not take place in either group. Further, we have no strong priors that a change in uptake should be differential between the two groups during the subsidy period itself (through week 4), thus, the more interesting question for our sub-aims is the extent to which uptake changes among the subsidy group during the intervention period as compared to week 5. It is also possible that positive or negative learning could take place in the treatment group during the subsidy period, though not be powerful enough to motivate a change in purchase behavior during that time – i.e. that these effects will bear out only after the price change in week

5. In any case, we cannot make causal claims from these tests. Rather, they will help us to tell a more complete story about the dynamics playing out.

3. Pilot study findings

In preparation for the full RCT described herein, we conducted a pilot study of 162 neighboring households in March and April of 2019 to test elements of the intervention, the baseline and follow-up surveys and the overall study design (preregistration AEARCTR-0003829). While the findings from this pilot are discussed in more detail elsewhere, a several outcomes of this process are worth note.

First, the idea of subscriptions (paying up-front for deliveries over the course of several days or weeks) was not found to be universally palatable among subjects. As a result, the final study design includes either 1) a 30-day subscription paid up-front, or 2) a pay-as-you-go 'punch card' for treatment households good for 30 subsidized water deliveries over the subsequent 30 days after the initial offer (whichever comes first). Control households (facing the market price) will also receive punch cards of either variety as a reminder, and to account for any salience effects that may result from being visited or selected as part of the study, or from simply having a nice looking punch-card. Second, we determined that outcome measurements for uptake should be considered both as a discrete one-time measure (uptake or not of water deliveries at follow-up interview) as well as over three weeks after the follow-up interview (proportion of households receiving deliveries per week). To track usage, we will rely both upon administrative records from the NGO, as well as randomly timed audits to participating households to ensure that they are receiving deliveries as scheduled. Although we plan to continue tracking household usage beyond this three-week follow-up period to see whether the effects of the intervention are sustained, we caution that relying on administrative records alone may be subject to some reporting error from delivery drivers who may not complete their assigned routes in the time they claim.

Third, during the pilot study we determined that participants enrolled in the study must first purchase the hardware required for water deliveries (20 litre bottle and dispenser equal to roughly 250 INR at market price). We also determined that purchase of these materials is a strong potential barrier to enrollment at their current prices, thus using a random price auction we were able to construct a demand curve for this hardware, conditional on knowing what delivery price the family would face. This helped us to determine an optimal pricing scheme for both the hardware and the water deliveries among treatment and control families to achieve our desired baseline levels of uptake. In this case, we were able to determine how demand for hardware shifts when households are randomly allocated into either a full price or subsidized price group, but during a colder period of the year when demand for water is somewhat lower than during the hot summer months when the full study will take place. Thus, we anticipate that demand will be higher on average across both treatment and control groups for the full study, and we will price our hardware offers accordingly.

Fourth, we determined that liquidity is a problem for some families who would like to participate in the study but cannot afford the full cost of the hardware. Thus, following prior research by Levin and others (2012), we have developed a scheme in which families can place an initial deposit on the hardware and be afforded a short time period in which to pay the remaining price. We plan to make the hardware product fully refundable during this payment period to allow for those who are unable to gather sufficient funds to receive a full refund of their initial deposit, conditional on the hardware being returned in clean and functional condition. The following sections outline the pre-analysis plan for this randomized controlled trial beginning in June of 2019 and ending in August of the same year.

4. Study details

The geographic area of the full study has been identified through help from the SHRI water delivery team. The convenience sample for this study was determined by budgetary and logistic constraints regarding the geographic scope and feasibility of the intervention/enumeration area. This was discussed in detail between the PI, the institutional partner SHRI, and DCOR survey team. Thus, we have chosen a sample area comprising two small communities near the village of Sukhpur in the Supaul district of Bihar. The sample will consist of a mix of Hindu and Muslim households, several dominant income generating activities (farming, animal husbandry, salaried positions, temporary labor), and generally low-income households. We intentionally sample these communities as potential beneficiaries of SHRI's services.

Household mapping in this geographic area will take place in the days immediately before the survey is to be launched. Representatives from the survey team will be dispatched to households in the survey area and enroll them based on the following preconditions: they must i) not already be regular consumers of SHRI water delivery or have received delivery during the pilot study, ii) be willing to take part in a baseline and follow-up survey as well as a marketing experiment, and iii) have a maximum household size of 6 adults (where households in Bihar are defined as family units utilizing the same cooking area, thus multiple households could reside in the same dwelling but be considered separate).

These selection criteria will ensure that households in study are comparable in prior exposure to delivered water, in size and in the quantity of water consumed. Larger households may require multiple daily bottle deliveries and therefore face a different price and level of exposure than smaller households. A maximum size of 6 adults ensures that 2-3 liters per person per day replaced by bottled water would not exceed 20-liters of water per day. Respondents will be asked whether they are willing to take part in a marketing experiment regarding water delivery, and whether the survey team can return to provide a short health presentation and a promotion to families, return at the end of one month to conduct an end-line survey, and visit periodically over the course of 7 weeks to monitor deliveries. Only families responding in the affirmative will be interviewed at baseline. Household units will be considered in the traditional unit of observation as a family unit residing in a dwelling with no less than one cookstove or kitchen devoted to providing daily meals for that group. Thus, several households could (and often do) reside on the same plot given that each uses a separate kitchen facility to prepare meals. Following IRB requirements, we will explain to families that the study entails minimal risk to human subjects, and that participating households may benefit from a chance to receive a discount on water delivery during the study.

5. Timeline Baseline survey and intervention

Enumerator training

Training will take place at a local rental hall in Supaul, the district capital, starting in June 2019 and last from 4 to 6 days. Topics will cover the baseline survey and intervention. Enumerators will receive instruction in each of the survey modules as well as the all elements of the intervention including the role of local social marketing teams and water delivery drivers. During the baseline and follow-up surveys, enumerators will work closely with social marketers, drivers and SHRI management staff to implement the study. At least one afternoon will also be devoted to pilot testing the survey questions in nearby households that will not be involved in the final study.

Baseline survey

Enumerators will recruit households and collect baseline data over seven to ten days from a sample of 516 total households. Baseline surveys will cover demographics, socio-economic characteristics, employment and income, water use and sanitation, health and wellbeing indicators, and household GPS coordinates. To examine potential learning about the water delivery product over time, respondents will also participate in a survey module employing a conjoint analysis (discussed in greater detail in Section 8 with our secondary research questions). Respondents will be offered discrete hypothetical choices about water given the pre-defined set of qualities we anticipate being most important to their purchase decisions. Each household will be asked to evaluate a random sub-set of three characteristic possibilities each. The baseline survey materials including the conjoint analysis script are included with our preregistration on AEA's preregistration website.

Social marketing intervention

At the end of the baseline survey, social marketing teams will be asked to join the enumerators to provide a marketing demonstration about the water delivery products to all 516 households surveyed at baseline. These social marketers are young men between the ages of 18 and 25 who have been selected from the local community to impart information about the clean water product and share their own experiences and education about the importance of clean water use. If feasible, this messaging may include a short video explaining the possible routes of diarrheal disease transmission in local Hindi language. After social marketers deliver their messages, they will offer respondents a free taste test of the potable water product. Then, they will give a demonstration of the high iron content of the local water sources as compared to SHRI's bottled water. The marketer will fill two cups with water. The first from the same pump the family uses for water consumption. The second, from SHRI's bottled water. Marketers then collect guava leaves (older leaves work better for the demonstration), crush them in their hands, and place the crushed leaves in each of the two cups of water. Tannins in the guava leaves react with iron content in the pump water to turn the water black (if arsenic is present in the water, it will react to turn water a bright blue color). This reaction takes roughly two minutes, and families can repeat the experiment themselves if they choose. Social marketers then allow ample time for families to ask questions about the water product, the delivery service, or any of the health benefits of purchasing water deliveries.

Random price offer & hardware purchase

Finally, the enumerators will invite households to participate in the random price offer for water delivery. Enumerators will each have a bag with ten small manila envelopes. Within each envelope will be a laminated card worth either 5 or 10 rupees per bottle for delivery price (see Figure 3), and each respondent will have a 50:50 chance of drawing either a 5 or 10 rupee per delivery punch card. In this way, the sample of 516 households will be randomly allocated into one of two arms.

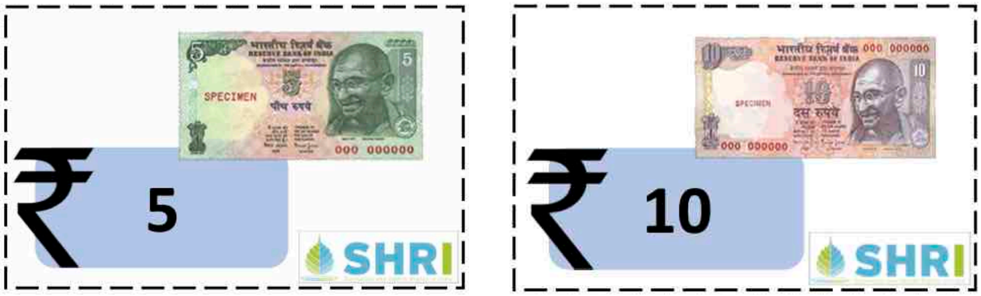

Figure 3. Random price drawing cards

- **Arm 1 –** *Social marketing visit with the option of 50% subsidy for water delivery*
    - Households will be offered a subsidized price of 5INR / delivery with punch-card coupons for one month. At the end of one month, the price will increase to 10INR / bottle for delivery for this group.

- **Arm 2 –** *Social marketing visit with the option of non-subsidized water delivery*
    - Households will be offered the market rate of 10INR / delivery with punch-card coupons for one month. At the end of one month, the price will remain the same for this group.

After the respondent draws a card, they will then be asked to choose whether they would like to participate by purchasing a bottle and dispenser. Bottles and dispensers will be sold for an up-front price of 250 INR (market price for this hardware is 280 INR, thus we are offering a 30 INR (~10%) discount if purchased in the next two weeks). If a household would like to participate but cannot afford the initial investment for the bottle and dispenser, they will be allowed to place a 50% deposit for the hardware and will be given the following two weeks to pay off the remaining 50% of the purchase price. As an incentive, the 10% discount on the hardware will expire once the two weeks have elapsed, and the price will return to 280INR. Following Levine and Cotterman (2012), during the two-week trial period, participating households who place an initial deposit will have the chance to return the bottle and stand for a full refund so long as neither piece of hardware has been damaged or misused, and as long as each is returned in a clean condition. To limit abuse of this trial period, all purchasing households will be asked to sign or provide a finger print on a document ensuring that they understand the terms of the loan period, and that they will be responsible for the full price of the product with no refund if the hardware is damaged beyond the possibility of reuse or lost or stolen.

Households choosing not to participate by purchasing (or placing a deposit) on a bottle/stand will receive a coupon good for future purchase during the subsidy period at the delivery rate designated during the random price draw. Figure 4 below shows this coupon, which serves both as a reminder of the price that the family faces during the subsidy period based on the random price draw, and as an accountability mechanism so that families are not tempted to later falsify the price they drew when approached for future sales (though we will keep a record of the price drawn by each family as well).

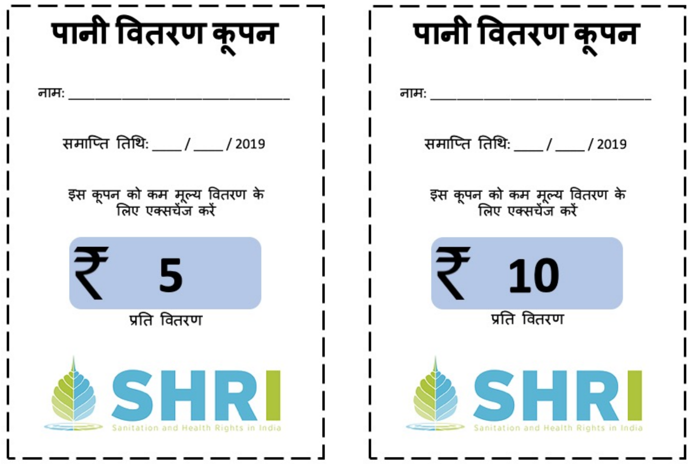

Figure 4. Non participant coupons for possible future purchase

Households choosing to participate by purchasing a bottle and dispenser will be issued a laminated punch card corresponding to the price that they drew during the auction (see Figure 5 below). These punch cards will be identical in size and appearance except for the printed price faced by each household. The punch cards will be laminated with a string or band to attach them to the narrow neck of the water bottle in an effort to keep punch cards from being misplaced. Upon each daily delivery, the driver will punch one 'bottle' image on the card until the end of the month-long period or after 30 deliveries, whichever comes first. Punch cards will include the head of household name and household ID, and households will be instructed that punch cards are non-transferable to other households and have no cash value.

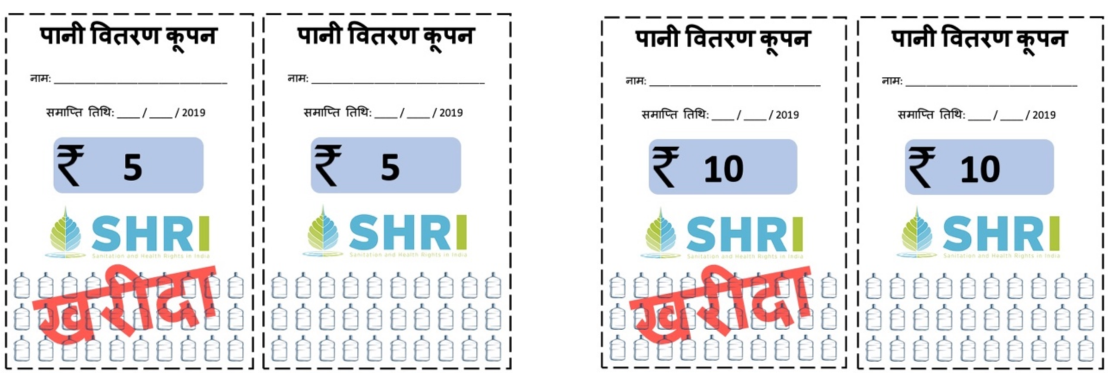

Figure 5. Punch cards

Importantly, since we wish to encourage high rates of uptake during the subsidy period, we will offer families one of two options: 1) pay the full price of a month of deliveries up-front and receive one bottle for free, or 2) pay-as-you-go for all 30 bottles. A laminated card carried by the enumerators, with prices corresponding to the random price drawn, will be used to explain this concept (see Figure 6, below). Punch cards for the former option will be overwritten in red with the phrase "paid" so that delivery drivers will not ask for payment upon delivery, and deliveries will take place every day for 31 days, while the latter option will require payment on delivery (see Figure 5 above). At the end of the household visit, families who choose to participate by purchasing water delivery but do not pay up front will be asked how often they would like to receive deliveries (every day, every other day, etc.) and a delivery schedule will be established (with a strong suggestion of daily deliveries). If families opt to purchase less frequently than once per day when asked at baseline, we may revisit these households during the intervening weeks to ask if they would like to change the frequency of their delivery schedules.

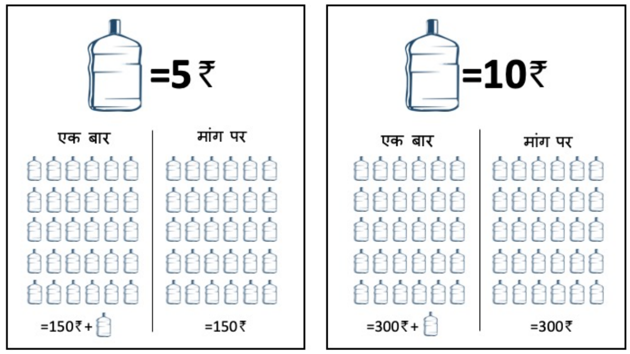

Figure 6. Up front versus pay-as-you-go explanation card

All of these visual reminders of the price faced by families depending on their random draw (Figures 3 – 6) should help serve to reinforce the price faced both during the baseline interview, and at subsequent household visits.

6. Household auditing, repeat offers and hardware purchase

Using administrative delivery records on the households who choose to purchase filtered water delivery, we will be able to track continued purchase of water delivery to see if delivery persists or wanes over time. We will track consumption during the one-month treatment period, as well as for at least three weeks after the subsidies expire. However, these records do not account for demand among households who did not choose to purchase at baseline. Therefore, we will hire our team of social marketers to continue conduct weekly audits of all participating households in the sample, and to revisit non-participating households to encourage enrollment. Each of our 5 social marketers will be responsible for visiting roughly 104 families per week over the course of the study. In visits to non-participating households, auditors will simply ask if the family has changed their mind and would like to purchase water from the NGO. Among the participating households, auditors will ensure that the NGO's delivery records are correct.[2] During these household visits to participating families, auditors will compare SHRI's administrative records to the number of deliveries reported purchased by the household over that time. The punch card can also be used for comparison. We will also use this opportunity to ensure that households are receiving deliveries (or delivery offers) at the desired time-intervals (frequency), and at the appropriate time of day (early morning in most cases). In the event of discrepancies, auditors will communicate with the delivery supervisor who will, in turn, direct delivery drivers to complete any missing deliveries.

[2] It will be especially important that all households are visited during the audits for weeks 4 and 7 to ensure that uptake during these weeks is measured such that all 500 households have received the product offer and either accept or reject. During weeks 1 and 5 the survey enumeration teams will ensure that all households are revisited and offered product sales during the baseline and end-line interviews, respectively.

Table 1. Example master delivery ledger

Household	July - 1		July - 2		July - 3		July - 4		July - 5		July - 6		July - 7		July - 8		July - 9	
	D	A	D	A	D	A	D	A	D	A	D	A	D	A	D	A	D	A
hhid_001	1	1	0	0	1	0	1	1										
hhid_002	0	0	0	0	0	0	0	0										
hhid_003	1	1	0	0	0	0	1	1										
hhid_004	1	1	1	1	1	1	1	1										
hhid_005	1	1	0	0	0	0	1	1										
hhid_006	0	0	0	0	0	0	0	0										
hhid_007	0	0	0	0	0	0	0	0										
hhid_008	0	0	0	0	0	0	0	0										
hhid_009	1	1	0	0	1	1	0	0										
hhid_010	1	1	1	1	1	0	1	1										
hhid_011	0	0	0	0	0	0	0	0										
hhid_012	1	1	0	0	0	0	1	0										
hhid_013	0	0	0	0	0	0	0	0										
hhid_014	1	1	0	0	1	1	0	0										

Table 1 shows a hypothetical example of the master delivery ledger for tracking home use. All households are listed in the first column regardless of delivery participation. Households who have purchased delivery are bolded for convenience (001, 003, 004, 005, 009, 010, 012, and 014), while all other households have declined to purchase deliveries. Shaded cells indicate the frequency of delivery desired by each household (as determined at baseline). For example, household 001 requests delivery every other day, while household 004 requests daily delivery, and household 003 requests delivery every third day. Some households might change their requested delivery schedule (such as household 012 who changes from every 3 days to every other day starting on July 4th). On each day of the intervention and follow-up periods we will have a record of deliveries reported from drivers – columns "D" of the ledger, marked complete with an "1" and incomplete with a "0". Meanwhile, data collected by auditors will be reported in columns "A" of the ledger (also marking "1" for complete and "0" for incomplete). In the example provided in Table 1, deliveries have taken place through July 4th, and audits have been conducted up until that day. These audits will allow us to triangulate between driver and household reported deliveries to give consistent measures of the number of deliveries received by each house per week of the intervention. Social marketers will also be asked to periodically revisit non-participating households to see if they have changed their minds about participation in the daily deliveries.

Finally, among households who chose an initial deposit for bottle and dispenser, after two weeks from the baseline survey social marketers will revisit to ensure full payment of the hardware costs. These costs must be paid off by the end of the second week to ensure that there are no price effects lingering in weeks 3 and 4 that may additionally influence the choice of uptake at week 5 after the price increases. If these households cannot pay the full price of the bottle and dispenser after two weeks from the date of baseline visit, they will be asked to return the bottle and dispenser and, if clean and undamaged, they will be refunded their initial deposits for the hardware. These families will continue to be revisited over the course of the study and allowed to re-enroll but must pay the full price for bottle and dispenser up front.

7. Endline survey and follow-up product offer

Data collection will take place immediately following the cessation of the 4-week subsidy period for water delivery and cover the same time-varying items as in the baseline survey, including questions about water collection, treatment and use, health, and opinions about safety. This survey will also contain another round of the conjoint analysis identical to the one undertaken during the baseline interview to see if household preferences have changed regarding the characteristics of the product versus the status-quo hand pump water among those who purchased versus those who did not. At the end of the visit, enumerators will ask all respondents they are willing to purchase additional water deliveries at the prevailing market rate (10 INR/bottle). All households will continue to receive random audits over the subsequent three weeks to see if scheduled water deliveries are taking place, and whether any families who had not participated have changed their minds and wish to purchase hardware and water deliveries from the NGO.

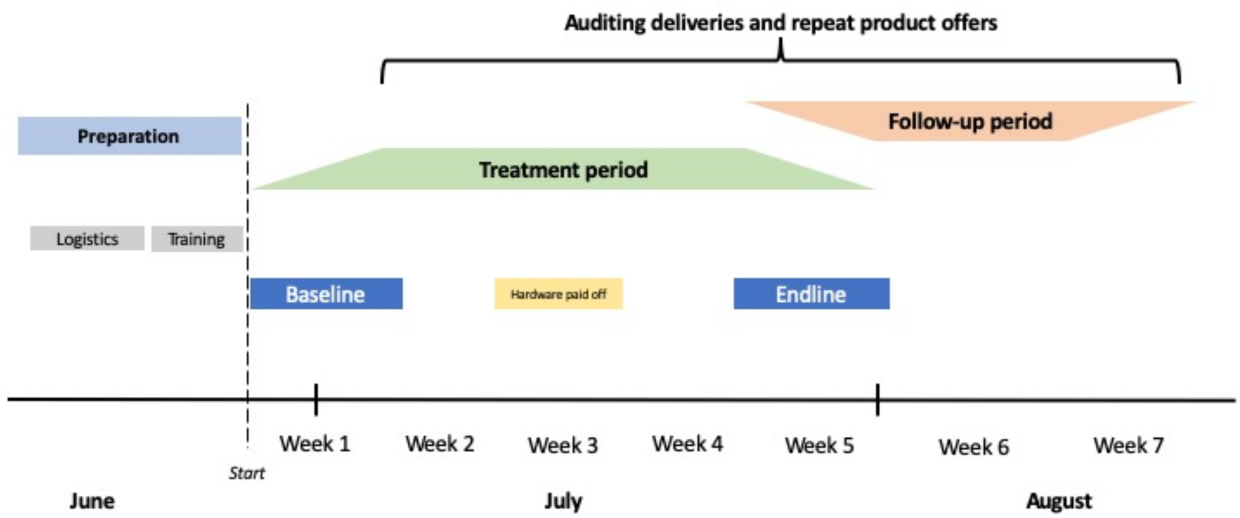

Figure 7. Study timeline

Figure 7 provides an outline of the proposed timeline for data collection, program implementation and follow-up social marketer visits. We will include a preparation period of roughly 10 days to coordinate delivery drivers, raw materials, and auditing procedures. Training will begin in late June and last roughly 4 days. Training will be provided for 13 enumerators, 1-2 field supervisors, 5 social marketers, two delivery drivers, and one delivery supervisor. The baseline will begin at the end of the training period and last for 7 to 10 days (sampling roughly 50 households per day (or 4-5 households per enumerator per day). At the end of the baseline survey period, enumerators will leave the field site and program staff including those from the NGO and the social marketers will coordinate deliveries and weekly visits/audits of study households to ensure deliveries take place as demanded and offers continue among non-participants. After two weeks from the first day of baseline data collection participating families will be required to have paid off the full price of the bottle and dispenser to continue with water deliveries.

Roughly one month after the initial baseline training begins, enumerators will return to the field site to begin training for the follow-up survey. The follow-up survey training will be shorter (for a shorter instrument) and last only 3-4 days. At the end of the training period, precisely 30 days from their baseline visits, all 516 households will be revisited by enumerators for the follow-up survey. Surveys will operate under the same time-frame as the baseline survey where households will only be approached for follow-up interviews after 30 days have elapsed since their baseline visits.

8. Analysis

Balance

Balance between the two treatment arms will be checked using a student's t-test comparing the covariates listed in Table 2 between treatment and control groups at the end of the baseline. We will also check this balance at the end of each day of baseline data collection to ensure the randomization procedure is happening properly in real time to allow for any adjustments in field data collection. First, we will evaluate balance across demographic characteristics including the household size, whether the family speaks Hindi as a first language in the household, the age of the head of household, and the number of children under the age of 18 residing in the household. Next, we will evaluate balance based on socio-economic characteristics including whether the household head has received any primary education, a generated wealth index score (derived from the average number of durable assets owned by the household), whether the dwelling is electrified, and whether the dwelling (including roof, walls and floors) are made of improved materials. Finally, we will check for balance based on water use variables including whether the household has ever purchased treated water delivery in the past, whether the household ever treats their water before drinking, and whether any children within the household have experienced diarrhea within the last two weeks.

Daily checks of the data during baseline will also include balance of treatment arms for each enumerator to ensure that enumerators are giving preference of one coupon or the other during the random price draw. Enumerators will be responsible for geographic blocks of households, therefore those within a given enumerator's catchment area may have similar demographic or socioeconomic characteristics. Thus, checks will be run on the percentage of respondents interviewed by each enumerator that are grouped into treatment vs. control conditions each day for each enumerator, as well as whether there is any relationship between socioeconomic status and treatment allocation. Specifically, we will be sure to examine balance between SES indicators from table 1 for each enumerator separately. Any patterns that may be found will be addressed with the enumeration supervisor at the end of each day of data collection.

Table 2. Baseline covariates required for balance

Variable name	Measurement	Description
Demographic variables		
Household size	Continuous	Number of household members (any age)
Hindi	Binary	Binary variable equaling 1 if the primary language spoken in the home is Hindi and 0 otherwise
Age of household head	Continuous	Age of the reported head of household
Number of children under 14 currently living in household	Continuous	Total number of children under age 14 residing in the household
SES variables		
Household head has any primary education	Binary	A binary variable measuring 1 if head has achieved any primary education and 0 otherwise
Asset index	Continuous	A mean score generated by the average ownership of 26 household assets between 0-1: (Σ assets)/26
Electricity	Binary	A binary variable measuring 1 of household has electricity and 0 otherwise
Improved structure	Continuous	A mean score generated by the average of three measures of improved structure, between 0-1: (Σ roof, walls, floors)/3 Improved roof (=1 if metal/GI, wood, cement, asbestos, concrete, shingles, tiles, brick; 0 otherwise); Improved walls (=1 if cement, stone, concrete, bricks, blocks, singles, metal, asbestos, and =0 otherwise) Improved floor (=1 if floors are parquet, polished wood, ceramic tile, cement, carpet, polished stone, marble, granite), and 0 otherwise).
Other covariates		
Ever received water delivery	Binary	A binary variable measuring 1 if household has ever purchased or received a water delivery in the past, and 0 otherwise
Ever treat drinking water	Binary	A binary variable measuring 1 if the household reports ever treating drinking water before consumption and 0 otherwise
Diarrhea	Binary	A binary variable measuring 1 if any children in the household have experienced diarrhea in the last two weeks and 0 otherwise

Outcome measurements

The main study outcome is uptake of potable water from SHRI. This outcome will be measured in one of two ways (see Table 3): The first is the percentage of households purchasing water in each of the study weeks during the intervention and follow-up periods. For our primary research question, we will report this measure for week 7 of the follow-up period, capturing the net impact of subsidies on uptake of water deliveries over the full program period. This requires that all households are visited in week 7 and offered water delivery sales. We can also capture this measure in weeks 5 and 6 of the follow-up period to see if uptake changes over time after the price shock for our primary research question. This will also be the preferred outcome for our secondary aims on anchoring and learning discussed previously (measured at weeks 1 through 4 and weeks 4 versus 5, respectively).

Our second measure of uptake is captured during our endline survey in the 5th week of deliveries (the week following the subsidy period). During the interview, subjects are asked whether they intend to purchase water delivery or not at 10/INR per bottle. This is measured using variable (currently FS1a (for those who did not purchase during the intervention period) and FS1c (for those who did purchase during the intervention period) of the endline survey – see AEA trial preregistration for endline survey questions). While this is a compelling measure because it is taken directly from survey enumerators, it may not be as reliable as delivery reports as it presupposes that respondents will purchase water without direct observation of actual purchase. This could present measurement error since some families may report wanting to purchase water yet change their minds after the interview, or vice-a-versa.[3]

Finally, we can measure water consumption as the number of bottles delivered to each family per week. Our primary week of interest in this case is again week 7. Although this measure provides more granularity in measuring the impact of subsidies on demand, it may also be subject to greater risk of misreporting and confounding. First, this measure relies on accurate reporting from both delivery drivers and social marketers. It also depends on accurate recall from interviewed family members visited by social marketers who may not have been present for all household deliveries during a given week. Further, the number of bottles consumed by each family per week will depend on the number of individuals in each household as well as whether the household opted for up-front payment versus pay-as-you-go, thus this outcome measure requires controlling for a number of factors like household size. Although this measure offers greater specificity it may also be subject to several sources of bias.

Table 3. Primary outcome measurements

Outcome measurement	Variable type	Description
Uptake		
1. Weekly uptake	Binary measures [0,1] each week	Did the family purchase water at all (yes/no) during the week; delivery reports validated by audits of participants and re-offers among non-participants
1a. Endline uptake	Binary measure [0,1] of intent to purchase at endline interview	Did the family respond that they intend to purchase water (yes/no) during the endline interview (week 5)?
Consumption		
2. Weekly consumption (no. of bottles delivered)	Count measure of the number of deliveries purchased per week	Total number of deliveries purchased each week; delivery reports validated by audits among participants and re-offers among non-participants

It is important to note that these three outcome measurements examine different variations on the same underlying question of demand. Outcome measure 1a (endline uptake) examines

[3] This occurred with several households during out pilot study, leading us to place a higher preference on the first triangulated measure of uptake.

uptake at a discrete moment where individuals in the treatment group are faced with the choice of purchasing (or not) immediately after the shock of the price increase (though they may anticipate this shock) and may not purchase at that exact moment. Figure 8 illustrates how we might conceptualize the second discrete outcome measurement (1a) as the choice of whether or not to purchase water at two points in time. In this case, if uptake at endline is greater in the treatment group than in the control group, we can conclude that there was some positive learning among subsidy recipients that motivated them to purchase at a higher rate (fail to reject H_1). If uptake is not statistically different from the control group, we can conclude that the intervention and subsequent price shock did not have a measurable effect on demand after the intervention (fail to reject H_0). Finally, if uptake is lower than in the control group, we can conclude that either anchoring or negative learning effects resulted from the intervention and subsequent price shock (fail to reject H_2). While we can control for uptake over the subsidy period, this discrete measure is less descriptive in terms of consumption over time.

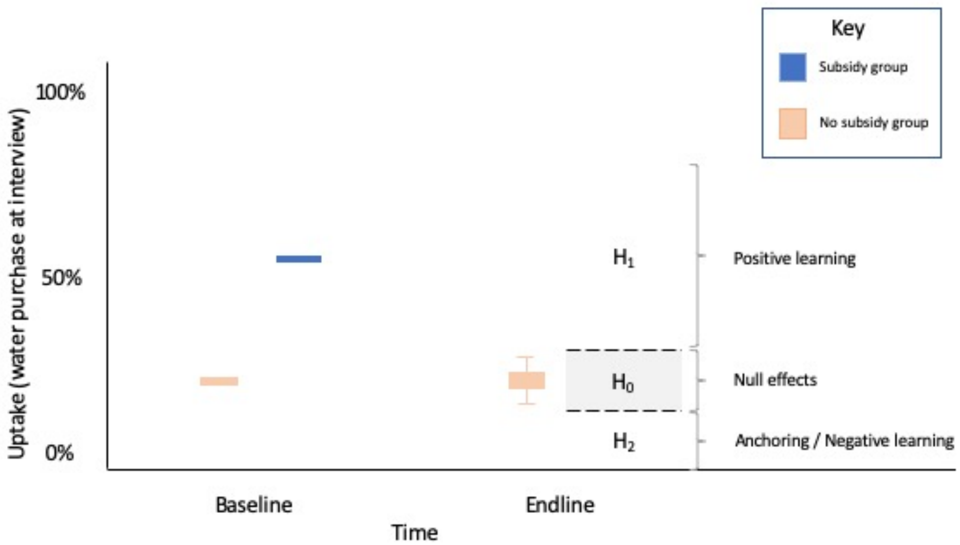

Figure 8. Potential outcomes under discrete measurement

Meanwhile, our preferred outcome measure (#1, weekly uptake) provides a weekly counting of whether the family consumes (yes/no) over the course of each week during the study, as recorded by household audits and delivery reporting. Ideally, we will have reliable data for this measure in each of the seven weeks (with offers made to each household in each week), with particular emphasis placed on week 7 for our primary research question. Weekly consumption (measure #2) examines demand in terms of the overall quantity of water purchased by each household in each week (number of bottles). Figure 9 below illustrates how we might conceptualize these measures over time during the intervention and follow-up periods, with the baseline survey taking place at week 1 (T_1) and endline survey taking place at week 5 (T_5). For measure 1, the y-axis is the percent of households purchasing in each week. For measure 2, the y-axis is the average number of deliveries purchased per week.

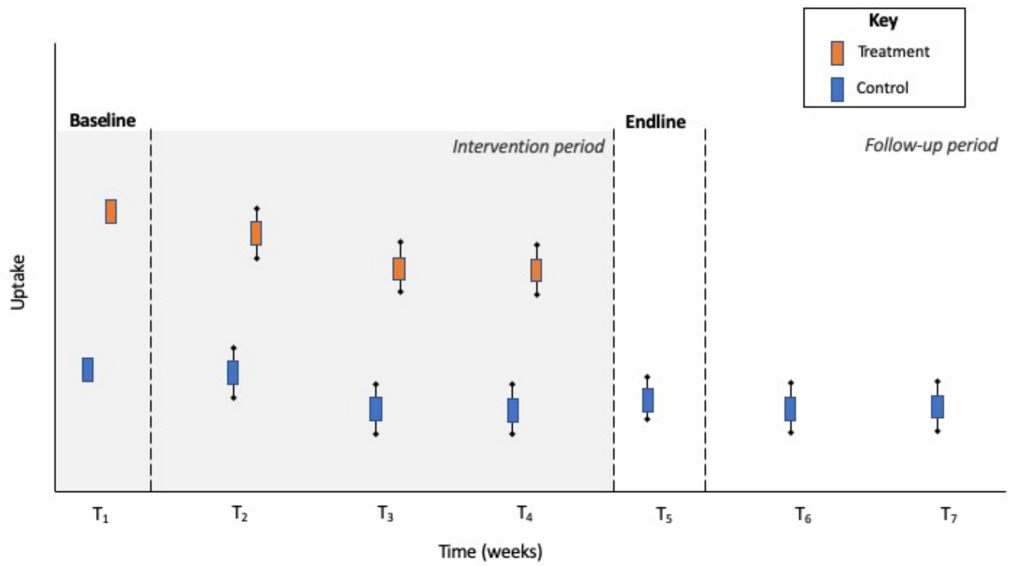

Figure 9. Weekly outcome measurements

Analysis strategy

Our analysis strategy begins with an unadjusted (naïve) model reporting the mean comparison in uptake from our RCT at week 7 between the subsidy and non-subsidy group. Equation (1), below models a simple student's t-test of uptake as Y_{ijt} (percent of households purchasing any water in week 7) for household i in hamlet j at time $t = 7$, with intercept β_0, for treatment arm $Arm1_{ijt}$, and a normally distributed error term η_{ijt}.

$$Y_{ijt} = \beta_0 + \beta_1 Arm1_{ijt} + \eta_{ijt} \tag{1}$$

Next, we will consider equation 2, which again examines the net impact of subsidies on uptake in week 7 using a logistic regression model that controls for water delivery uptake at all during the one-month subsidy period using a lagged dependent variable Y_{ijt-1}, as well as a vector of any unbalanced household-level covariates X_{it} from Table 2 (we will also examine the choice of covariates to include in this model using lasso), ϵ_{jt} is a vector of hamlet-level fixed effects[4] for hamlet j and η_{it} is an individual-level error term.

$$Y_{ijt} = \beta_0 + \beta_1 Arm1_{ijt} + \beta_2 Y_{ijt-1} + \beta_3 X_{it} + \epsilon_{jt} + \eta_{it} \tag{2}$$

We can will consider both of the above equations for Y_{ijt} with time $t = 5$ for uptake at endline using both the triangulated measurement of percent of households purchasing water during week 5 as well as whether households report intending to purchase water at endline (measure 1a). This will allow us to see whether demand shifts over time (from week 5 to week 7), as well as whether the reported desire to uptake aligns with reliable triangulated measures of uptake from delivery reports and audited records.

[4] Analogous to an error term for hamlet-level fixed effects.

Finally, for our primary research question we will also examine delivered water consumption measured as the number of deliveries purchased per household in week 7 between treatment and control groups. As stated previously, a naïve measure of this outcome will be biased based on household size and other potential confounders, thus we introduce the following equation 3 as a negative binomial count data model to account for the new outcome measure. Count data models are more appropriate to our outcome measure than ordinary least squares or logistic regression because the dependent variable is non-continuous and the difference between values is standard. Further, we choose a negative binomial count model over Poisson because we expect our outcome to contain a large number of zero values.

$$Y_{ijt} = f(\beta_0 + \beta_1 Arm1_{ijt} + \beta_2 Y_{ijt-1} + \beta_3 X_{it} + \epsilon_{jt} + \eta_{it}) \tag{3}$$

Here, Y_{ijt} is the total number of bottles delivered in week 7.[5] The right-hand-side values are the same as in equation 2, except that we now express them as a function to account for non-linear distribution of our outcome variable. We also recognize that the quantity of water consumed will depends, in-part, on the number of individuals consuming water, thus we will include household size in the vector X_{it} in the final regression. See table 4 below for an example of how results will be reported in tabular format.

Table 4. Exemplar results table for primary research question (binary and count uptake at week 7[#])

Variables	un-adjusted model			adjusted model 1			adjusted model 2		
	coef.	s.e.	p-value	coef.	s.e.	p-value	coef.	s.e.	p-value
Difference in uptake at week 7									
Any uptake during subsidy period									
Covariates									
Hindi									
Household size[*]									
Age of household head									
Number of children <14 y.o.									
Head primary education									
Asset index									
Electricity									
Improved structure									
Ever received water delivery									
Ever treat water									
Children <5 with diarrhea last 2 weeks									
Vector of hamlet fixed effects[$]									
n									

Note: [#]Unadjusted model contains no covariates, selection of covariates used in the adjusted model 1 (logistic regression) and adjusted model 2 (negative binomial count model) will be based on balance t-tests and lasso procedures; [*]adjusted model 2 will include a covariate for the number of persons per household; [$]Hamlet fixed effects included as a robustness check in adjusted model 1 and adjusted model 2; Adjusted model 1 and two use robust standard errors.

[5] We may also examine Y_{ijt} in week 5 to compare the immediate post-price shock decision made by the household to consumption over time.

Analysis of sub-aims

In order to investigate our two sub-aims around the change in uptake among the treatment group only, we propose the following analysis. In a data representation where each observation represents a household-period, we will regress Y_{ijt}, uptake for family i, in hamlet j, in week t by an intercept, $Week_{ij}$ an indicator for time in weeks, X_{it} a vector of baseline household-level covariates, ϵ_{jt} hamlet level fixed effects, and η_{it} an individual level error term using robust standard errors.

$$Y_{ij} = \beta_0 + \beta_1 Week_{ij} + \beta_2 X_{it} + \epsilon_{jt} + \eta_{it} \quad (4)$$

For sub-aim 1 (quantifying the week-to-week change in uptake during the subsidy period), we will examine uptake where $Y_{ijt=2}$, $Y_{ijt=3}$ and $Y_{ijt=4}$ to determine the size and significance of the coefficient for each. In order to examine sub-aim 2, we will quantify the difference between $Y_{ijt=4}$ and $Y_{ijt=5}$ in a separate regression with the same functional form and with reference week 4. The choice of household-level covariates X_{it} will be determined using cross-validated lasso and be drawn from the list of potential covariates from Table 2. The results of this procedure will be displayed as in the exemplar Table 5, with coefficients for weeks 2-4 uptake specific to sub-aim one, and the coefficient for week 5 relevant to sub-aim two – both comparing change in uptake and associated significance level to the reference week 1. A second test will be run isolating uptake only at weeks 4 and 5 to examine the change in uptake between those weeks, as additionally relevant to sub-aim 2.

Table 5. Exemplar results table for research sub-aims 1 and 2

	Sub-aim 1			Sub-aim 2		
Variables	Coef.	s.e.	p-value	Coef.	s.e.	p-value
Week 2 uptake$						
Week 3 uptake$						
Week 4 uptake$						
Week 5 uptake$#						
*Covariates**						
Hindi						
Household size						
Age of household head						
Number of children <14 y.o.						
Head primary education						
Asset index						
Electricity						
Improved structure						
Ever received water delivery						
Ever treat water						
Children <5 with diarrhea last 2 weeks						
Vector of hamlet fixed effects						
n						
R^2						

Notes: $ coefficients for weekly uptake express difference (and level of significance) from the reference week 1; # In sub-aim 2, only week 5 will be included in the table with reference week 4 omitted; *Choice of covariates depends upon results of lasso procedure

Sample Size Calculation

We use the following formula in Equation (4) from Djimeu and Houndolo (2016) and Bloom (1995) to determine the required number of households per arm for our randomized controlled trial experiment with individual-level binary outcomes. We discuss this calculation in detail as a conservative estimate of power for the study:

$$n = \{\frac{P}{T\delta^2} + \frac{-P+1}{-T+1}(-t_1 - t_2)^2\} \tag{4}$$

Where, for a two-tailed test, $\alpha = 0.05$ and our desired power is $\beta = 0.80$, both by convention, thus $t_1 = 2.04$ and $t_2 = 0.85$.[6] We expect the true population proportion of the outcome in the absence of the intervention (i.e. the uptake at time t=1 for the no-subsidy group) to be $P = 0.15$ and the proportion of the study sample randomly assigned to treatment to be $T = 0.5$.[7,8] Finally, we follow previous research in this area (Dupas 2014; Fischer et al. 2018) which finds effect sizes of between 7 and 11 percentage points, respectively. Further, we consider that an economically relevant effect size would be a 10pp difference in uptake between treatment and control groups at week 7. Smaller effect sizes, though interesting, might suggest that the use of subsidies would be unlikely to create an adequate return on investment. Thus, we aim to be powered for a minimum detectable effect (MDE) size of 10 percentage point difference in uptake between the treatment arm and control ($\delta = 0.1$). Given these parameters and assuming a naïve model with no covariates to improve precision of our estimates, the equation above dictates that a 2-arm trial would require a minimum 500 households. In our pilot study (taking place over three weeks) we experienced a 1.2% attrition rate (2/162 households) in a neighboring community. Since households in the full study exhibit similar characteristics and the time-frame for the full study is only 7 weeks in total (4 for baseline and follow-up interviews, and three for continued delivery data), we expect similar attrition over the full study period. **Thus, plan to sample a total of 516 total households to account for a similar rate of attrition in our sample.**

There is some uncertainty around the potential uptake of the control groups during follow-up. In our case, we expect that at time $t = 1$ uptake will be similar between the partial- and no-subsidy groups (roughly 15%).[9] However, the minimum detectable effect size is sensitive to the rate of uptake in the control group (for example, to be powered to detect a 10pp difference in uptake between groups, a base change in uptake among the control from 15 to 20 percent, would require an increase of 88 households for the sample (588 in total, plus 12 to account for attrition for a total of 600 households), and an increase from 20 to 25 percent would require an additional

[6] Given a minimum sample size of 30 decided beforehand to determine t values.

[7] Following Hasselblad (2016), equivalent arms (1:1:1) should produce sufficient power, as the tradeoff with other ratios (e.g. 3:3:4) are only marginally better.

[8] I also acknowledge that the inclusion of covariates to increase the precision of the estimated effect of the treatment on uptake would help to decrease the needed sample size, examined further below.

[9] We estimate that uptake among the partial-subsidy group will be around this level based on findings from our pilot study. Willingness to pay auctions (see: Baker et al. 1954) were conducted for a combination of the hardware (bottle and dispenser) and water deliveries among a similar subset of respondents to determine the optimal price for the partial subsidy group. We found that when water was priced at 10 INR per delivery, roughly 15% of respondents were willing to purchase the bottle and dispenser at a price of 200 INR. In the case of this study, summer months might see higher demand for water (since temperatures are substantially higher), thus, we may choose to price hardware at closer to market rates of 250 INR to achieve around 15% uptake in the control arm.

70 households (or 658, or 671 accounting for attrition) – see Figure 10 for an illustrative example. However, these changes in control group uptake (to 20% and to 25% respectively), given no change in sample size would only equate to a change in MDE from 0.097 to 0.105 and 0.111, respectively. Thus, we explore sensitivity of these parameters in Figure 9 below to consider the tradeoff between the desired MDE and the corresponding sample size required.

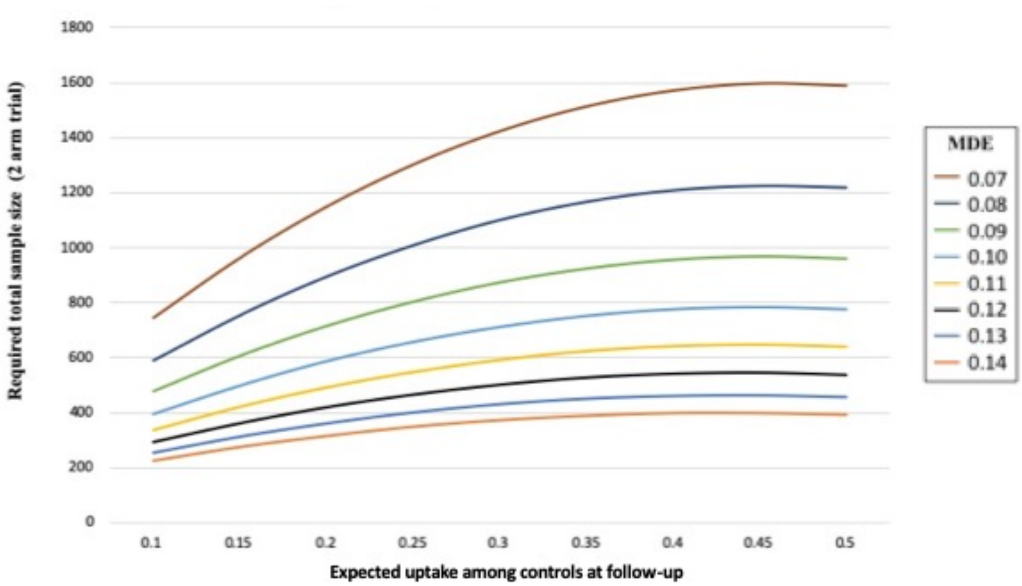

Figure 10. Required sample for effect size scenarios

Considering results from similar studies helps to explain the MDE this study should be powered to detect. Dupas (2014) finds a 7-percentage point difference between treatment and control given a relatively smaller change in prices faced by consumers in the first period (average of all values less than 20% versus all values more than 20% market price) than in this proposal (100% vs 50% of market value), thus this may present an upper bound of possible outcomes.[10] Fischer and colleagues (2018) find a 10-percentage point difference in uptake for a pooled sample of all products sold during the second period between their full-subsidy and sale arms (9 percentage points in the for-profit arm and 12 percentage points in the not-for profit arm). The effect size for uptake can be further disaggregated for each product, showing a range of possible negative impacts as small as -5.4pp for Zinkid and as large as -24pp for Panadol and -18pp for Elyzole. As such, ability to detect a 10pp change in uptake between arms would be within the demonstrated range of effect sizes seen thus far in the literature.

We also explore the potential impact of including covariates for additional power in table 5, below, using Equation (5) from Djimeu and Houndolo (2016):

$$n = \{\frac{P}{T\delta^2} + \frac{-P+1}{-T+1}(-t_1 - t_2)^2(-R^2 + 1)\} \quad (5)$$

[10] Dupas (2014) finds that for insecticide treated nets that market for $3.80 each, baseline adoption of ITNs was 7% among those receiving no subsidy and increased to over 60% when the price decreased to 19.7% of the original purchase price, and to 98% when the price was zero. At follow-up (one year later), she finds that households who had received ITNs for 20% or less than market price at baseline were 7.2 percentage points more likely to purchase an ITN at end-line than those who had received nets for more than 20% of market price previously.

For example, using the preferred number of observations (500 across two arms before accounting for attrition) and assuming an R^2 of 0.10 would decrease the minimum detectable effect size (MDE) from 0.090 to 0.085. Thus, with the inclusion of additional (level-1) covariates to account for outcome variance, we may be able to considerably increase the precision of our estimates. See Table 5 for a breakdown of possible R^2 assumptions.

Table 5. MDE given proportion of outcome variance explained by level-1 covariates

n	R^2	*MDE*
500 total households across 2 arms	0.00	0.090
	0.05	0.087
	0.10	0.085
	0.15	0.083
	0.20	0.080
	0.25	0.078
	0.30	0.075
	0.35	0.072

Several other power concerns are worth note. SHRI currently distributes around 70,000 liters of water during the peak summer season and between 30-40,000 liters during the cooler winter months. This capacity will expand through upgrades currently being made to the current water infrastructure, though the extent of this potential expansion is still somewhat unknown. The extent to which demand will increase during the summer months from what we observed during the pilot period in March is also uncertain. Finally, SHRI has relied on purchasing water from competitors to cover customer deliveries in the event of equipment malfunctions or other outages. Of course, any expansion for this trial that involved purchasing water from another producer would require that deliveries continue to be managed by SHRI delivery drivers, and that these deliveries be randomly allocated across treatment arms so as not to introduce any differential bias due to water taste, temperature, quality or other characteristics. Further, households receiving any deliveries of water from competitors must be tracked carefully to ensure that we can control for this process during analysis of data.

9. Secondary research questions

There are two fundamental shortcomings with the proposed primary research question and corresponding hypotheses tested above. First, alternative hypotheses 1 and 2 contain potentially competing effects (anchoring vs. learning) that are not mutually exclusive. Thus, even if an anchoring effect dominates the overall purchase decision, it may still do so in the presence of learning. Second, although we mostly expect scope for positive learning about the product, the direction of a learning effect is theoretically ambiguous (either positive or negative). Thus, if we fail to reject hypothesis 2 (evidence of anchoring or negative learning effects), given the current study design we will be unable to distinguish between anchoring and negative learning as the pathway through which this result is observed.

Therefore, it is critical to understand more about the potential channels through which learning might operate, and which of these plays the strongest influence on the purchase decision (whether positive or negative) could be very useful in understanding the effect of our intervention.

For example, in the published literature (and in my findings from formative interviews in the summer of 2018), respondents have expressed that the most important considerations about their purchase decisions (aside from cost) is that water be “clean, tasty and simple to use” (Brouns, et al. 2013, 12), and that if water is not delivered promptly at a cold temperature, customers will be far less likely to purchase.

To investigate the relative power of these learning pathways, we propose conducting a conjoint analysis (a form of discrete choice experiment) in which respondents are asked – given a set of characteristics about a good – to express a preference for one possible ‘scenario’ or combination of these characteristics over another. This analysis will allow us to determine the relative importance of each of a pre-defined list of characteristics of the product in terms of respondents’ stated preferences. Bridges et al. (2011), Johnson et al. (2013) and Hauber et al. (2016) review variations on this method and offer a set of guidelines for their application in various health settings. We also draw from examples in the literature (in particular, Ryan and Farrar 2000; Taylor and Villas-Bosas 2015; Zannolini et al. 2018). In this case, during the baseline and follow-up surveys, respondents will be asked to express their preferences for bottled water delivery given a randomly allocated, pre-defined set of characteristics about that water. These qualities are derived from the literature and focus group discussions, are most important to consumers regarding water purchase decisions.

Questions and hypotheses

The goal of this exercise will be to answer the following secondary research question:

1. *What characteristics of water delivery subscriptions play the strongest role in beneficiary preference for the product?*

A natural extension of this question is to determine whether these preferences change over time, particularly among those who experience the product through subsidy versus those who do not. Thus, we propose the following secondary research question #2, examining both the size and direction of any learning effects over time (though we caution that our sample sizes, determined based on other considerations, may not be sufficiently large to account for measurement error of these changes over time):

2. *How does the preference for characteristics of water delivery subscriptions change over time between those who receive the product versus those who do not?*

Secondary question #2 represents a treatment-on-the-treated type of analysis in which we compare *takers* to *non-takers*, as described previously. Though, as part of this question we can also examine any changes in preference among those who were offered subsidized water delivery subscriptions versus those who were not (an intent-to-treat analysis). The implications being that, although ITT results are likely to be less than TOT results, these findings may still inform policy regarding expectations around the effect of offering subsidies to groups versus presuming that effects are only significant among those actually receiving subsidies.

Following the results of focus group discussions and previous literature, the characteristics that seem most likely to either a) comprise strong naïve priors about the quality of the good, or b) be experienced during the subsidy period upon repeated delivery among takers. These characteristics comprise the quality of the product m such that, $m = f(p, t, d, h, c, n)$, where p refers to the price of the product, t refers to the taste of the product, d refers to the convenience of the product being delivered, h refers to the health benefits of the product, c refers to the temperature of the product, and n refers to the social pressure resulting from whether ones neighbors consume the product.

Secondary research question study design

To test these hypotheses, we undertake the following procedure. First, we take each of these qualities, with the addition of a variable for price (following Bridges et al. 2011) and discretize them simply into the following valuations: In the case of price, we use a categorical variable with standard interval values of 0INR, 3INR, 6INR and 9INR.[11] For each, these valuations are: t, taste (good; not good); d, convenience (on demand; call for delivery); h, health quality (will not cause sickness; may cause sickness); c, temperature (cold; warm), and n, neighbors (my neighbors use; my neighbors don't use). Each of these qualities can be split into competing scenarios as displayed in Table 6 below.

Table 6. Product characteristic possibilities

		(A)	(B)	(C)	(D)
i.	Price	0 INR	3 INR	6 INR	9 INR
ii.	Taste	Tastes nice	Tastes bad		
iii.	Convenience	On demand	Must call to order		
iv.	Health	Safe	Unsafe		
v.	Temperature	Cold	Warm		
vi.	Neighbors	My neighbors use	Neighbors don't use		

Next, we generate a subset of scenarios of product characteristics, varying qualities in each scenario as shown in Table 7 below until we arrive at $(4 * 2 * 2 * 2 * 2 * 2) = 128$ scenarios:

Table 7. Product characteristic scenarios

		1	2	3	4	5	6	7	8	…	128
i.	Price	A	A	A	A	A	A	B	B		D
ii.	Taste	A	A	A	A	A	B	A	B		B
iii.	Delivery	A	A	A	A	B	A	A	A	…	C
iv.	Health	A	A	A	B	A	A	A	A		B
v.	Temperature	A	A	B	A	A	A	A	A		B
vi.	Neighbors	A	B	A	A	A	A	A	A		B

Rather than compare each possible combination of scenarios (128*127 = 16,256) as in a discrete choice experiment, we instead follow the example of Ferrar and Ryan (2000) and create a single 'status-quo' scenario against which each household will state its preference. Specifically, households will be asked to imagine that their current water supply exhibits the qualities of a status quo scenario. This will be a 'theoretical' hand pump-base water source that has the qualities described in Table 7 below, which are in line with the experiences of most households in the region. Notably, this 'theoretical' pump may not reflect the experiences of all households, given that both the taste of water and health risks can vary from pump to pump. However, it will not be

[11] In this hypothetical scenario, we will tell respondents not to consider the price of the hardware when expressing their preferences between different options.

cognitively burdensome for households who do not experience these exact pump qualities each day to imagine the scenario. Respondents will also be asked to imagine that all the pumps in nearby households have the exact same qualities (to limit the possibility of moving to a next-door neighbor's pump for water). Otherwise, the price, delivery, temperature and neighbor qualities of the status quo are very much the same across all communities. Thus, respondents will be asked to compare the status quo option to a random selection from among the alternatives.

Table 8. Status-quo decision

	quality	theoretical pump	description
i.	Price	A	Free
ii.	Taste	B	Tastes like iron
iii.	Delivery	A	On-demand
iv.	Health	B	Unsafe
v.	Temperature	A	Cold
vi.	Neighbors	A	Neighbors use

Next, following Bridges et al. (2011), we eliminate all product characteristic scenarios (as from Table 8 above) that are strongly dominated by the status quo decision. For example, a delivered water product for a positive price that tastes bad, requires request for delivery, can make respondents sick, and is not used by neighbors will certainly be dominated by the status quo. Next, because it is impractical to ask each household to choose or rank preferences from among all possible scenarios, we will instead offer a random subset of three possible choices to each household and ask at each choice to select between a random scenario and a status-quo scenario (following Ferrar and Ryan 2000). This reduced set of choices will be randomized across all respondents in the baseline survey in order to answer secondary research question 1. These choices will be balanced across treatment arms, and each respondent will be asked to choose either between the status quo condition, one randomly assigned scenario, or a 'no preference' option.

One of these choices is illustrated in Figure 11 below. Every randomly selected choice contains the same set of images on the left, with the picture of the pump at the top. In this case, each of the possible product characteristics in this random alternative (#90) is different from those in the status quo: The alternative costs 6INR instead of 0INR; the alternative has a good taste instead of a poor taste, the alternative requires a phone call to have water delivered while the status quo can be used on demand, the alternative is safe for use among children while the status quo might cause illness, the temperature of the alternative is warm, while the temperature of the status quo is cold, and finally all the neighbors in the community use a different source than the alternative (pumps) while the neighbors in the status quo community use the same source (pumps).

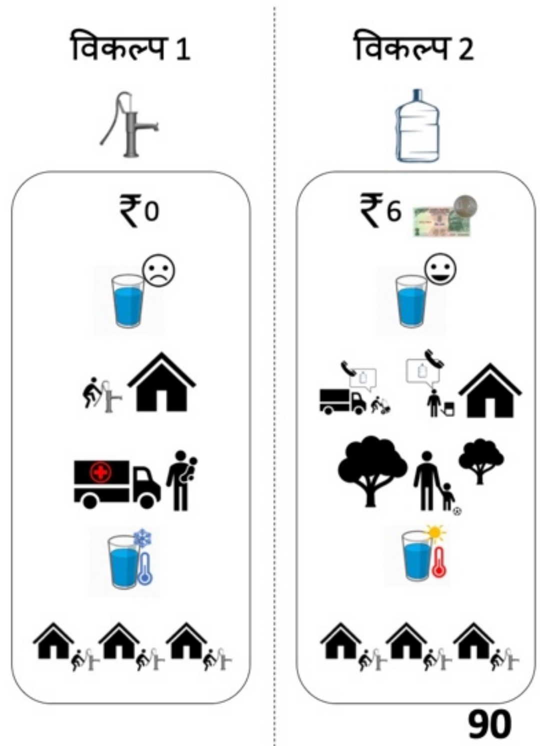

Figure 11. Example of one potential choice set faced by a household

Each household will be given three possible scenarios comparing the status quo with a random alternative scenario. Households will be asked to choose either a) the status quo, b) the alternative scenario, or c) no preference between the two. The result of this process will be a set of binary variables, one for each choice possible within each of the $i - vi$ characteristics above, with a subset of 3 choice-set observations for each household at time $t = 1$ (and 3 additional observations at time $t = 2$, described for secondary research question #2, below).

Estimation

To analyze the results of this experiment in time $t = 1$, we specify the following multivariate logistic regression in Equation (6), below:

$$g(\mu)_{ij} = \alpha + \beta_1 p_{ij} + \beta_2 t_{ij} + \beta_3 d_{ij} + \beta_4 h_{ij} + \beta_5 c_{ij} + \beta_6 n_{ij} + \beta_7 X1_{ij} + \beta_8 X2_{ij} + \epsilon_{ij} \quad (6)$$

Where, for household i and choice j, $g(\mu)_{ij}$ is the log-odds of taking water delivery over the status quo (the 'theoretical' household pump), α is the intercept, p is price, t is taste, d is convenience, h is health, c is temperature, n is neighbors, $X1$ is a vector of household-level covariates, $X2$ is a set of community- or hamlet-level fixed effects, and ϵ is an error term with extreme value distribution. To test each of the hypotheses in secondary research question #1, we will examine the marginal output of each corresponding coefficient, β_1 through β_6, for statistical significance (dummies for each category within covariates).[12] Statistically significant coefficients at the 5% level signify a failure to reject the corresponding hypothesis regarding that characteristic. The sign of the coefficient will signify the direction in which respondents value the particular characteristic.

In order to answer secondary research question #2 (*Do preferences change over time among households who received deliveries versus those who did not?)* we will repeat this procedure during the follow-up survey, again randomly distributing a subset of choice scenarios to respondents and asking them to select from between alternatives. We can then compare the outcomes of the regression both at time $t = 1$ and $t = 2$ for those who received deliveries versus those who did not. Specifically, we will pool data across time and formally test whether the pre- versus post- coefficients are statistically different both (a) jointly as a group, and (b) for each individual attribute.

[12] Note that coefficients will exist for each variable category such that β_1 is actually $\beta_1 p_1 = 3INR$; $\beta_1 p_2 = 6INR; \beta_1 p_3 = 9INR$ with $0INR$ as the reference category, thus there will be a total of 9 covariates, one for each alternative characteristics category.