Eliciting and Utilizing Willingness-to-Pay: Evidence from Field Trials in Northern Ghana

James Berry, Greg Fischer and Raymond P. Guiteras*

December 2015

Abstract

We demonstrate the use of the Becker-DeGroot-Marschak (BDM) mechanism to elicit precise, individual-level willingness-to-pay and thereby enhance the information generated by randomized experiments. Applying the mechanism to a field experiment studying clean drinking water technology in Ghana, we find that although willingness-to-pay is low relative to the cost, demand is relatively inelastic at low prices; prices do not generate significant sunk-cost effects; and treatment effects are heterogeneous with respect to valuation and consistent with effort as a mediator. We explore differences between BDM and take-it-or-leave-it valuations and make recommendations for effectively implementing BDM in the field.

JEL Classifications: C26, C93, D12, L11, L31, O12, Q51

Keywords: price mechanism, heterogeneous treatment effects, health behavior, Becker-DeGroot-Marschak, field experiments

*Berry: Cornell University, jimberry@cornell.edu; Fischer: London School of Economics, g.fischer@lse.ac.uk; Guiteras: University of Maryland, guiteras@econ.umd.edu. We thank Alex Nisichenko and Abdul-Salaam Sayuti for superb work in coordinating the field activities in Ghana, Kelly Bidwell and Jessica Kiessel at Innovations for Poverty Action, and Mary-Kay Jackson and Susan Murcott at Pure Home Water. We are grateful for helpful comments from Douglas Almond, Nava Ashraf, Abhijit Banerjee, Sylvain Chassang, Vivian Hoffman, Leigh Linden, Dean Karlan, Sheila Olmstead, Gerard Padró i Miquel, Giordano Palloni, Jesse Shapiro, Erik Snowberg, Chris Udry, Sergio Urzua, Roger von Haefen and seminar participants at Colgate University, Columbia University, Duke University, HECER, the Inter-American Development Bank, the London School of Economics, Queens University Belfast, Resources for the Future, Stanford University, Triangle Resource and Environmental Economics Seminar, Universitat Pompeu Fabra, the University of California Santa Barbara, the University of Virginia, and Wesleyan University. S. Aqil Esmail, Kate Loose, Diego Rojas and Ben Zou provided excellent research assistance. This research was supported by the Improving Institutions for Pro-poor Growth research consortium, the International Growth Centre and the Maryland Population Research Center.

1 Introduction

Willingness-to-pay (WTP) for a good, service or amenity is a key parameter for economic or policy analysis. It forms the fundamental building block of demand functions and hence our ability to estimate demand elasticities and consumer surplus. Practically, accurate measures of WTP can inform pricing policy, guiding the magnitude and targeting of discounts or subsidies. They also provide an important intermediate input for the study of technology adoption, social learning, and other issues related to demand formation. WTP and its relationship to a product's benefits are critical for understanding when the price mechanism is useful for allocating goods where their benefits are greatest and when it simply limits access for those with the greatest need.[1] Moreover, WTP provides a link between the structural and treatment effect approaches to policy evaluation and, since price is often a policy variable of interest, is a natural dimension along which to assess heterogeneous treatment effects.[2]

Ideally, one would like a precise measure of each individual's WTP; however, obtaining such a measure is often difficult. If an individual believes her answer to the question, "How much are you willing to pay for this product?" will affect the actual price, she may answer strategically. Alternatively, her answer to a hypothetical question may differ substantially from what she would actually do when confronted with a real-money choice. Economists have considered a range of techniques to elicit a truthful answer, including various stated-preference methods such as contingent valuation and conjoint analysis (Carson and Hanemann, 2005), and revealed-preference methods such as simple take-it-or-leave-it offers, Vickrey auctions, and nth-price auctions (Shogren, 2006). In field experiments, demand is often measured using take-it-or-leave-it (TIOLI) offers, sometimes at randomized prices. TIOLI is transparent and simple to implement, but provides limited information, since the researcher only observes an individual's binary decision to buy or not buy at a single price point. The Becker-DeGroot-Marschak mechanism (BDM), commonly used in experimental

[1]See, for example, Ashraf, Berry and Shapiro (2010), Cohen and Dupas (2010), Karlan and Zinman (2008), and Tarozzi et al. (2014).

[2]See Heckman and Vytlacil (2005); Heckman, Urzua and Vytlacil (2006); and Chassang, Padró i Miquel and Snowberg (2012).

economics, provides an attractive and underutilized alternative. In contrast to TIOLI, BDM elicits an individual's exact willingness-to-pay for a good. It operates much like a second-price auction against an unknown or random price. An individual states her bid for an item, then a random price is drawn from a distribution. If her bid is greater than or equal to the price, she receives the item and pays the price drawn. If her bid is below the price, she pays nothing and receives nothing. For expected utility maximizers, bidding one's true maximum willingness-to-pay is the dominant strategy.[3]

In this paper, we present the results of an experiment in which we used both BDM and TIOLI to study the willingness-to-pay for and impacts of household water filters in a sample of 1,265 subjects from 15 villages in rural northern Ghana. The setting is appropriate for our study, as a high incidence of diarrheal disease and lack of infrastructure for improved drinking water sources has led to promotion of point-of-use water treatment through social marketing. The filter requires effort to use, but if used properly can reduce diarrhea among adults and children who drink from it. After normal marketing efforts, we elicited willingness-to-pay and distributed the filters. We conducted follow-up surveys one month and one year after the sale to measure filter usage and health outcomes related to water quality.

We use our experiment to demonstrate three properties of BDM that make it appealing for field work. First, by providing a precise measure of willingness-to-pay, it allows for direct, non-parametric estimation of demand. This feature allows us to inform pricing policy with direct predictions for take-up at all potential prices, rather than a fixed set of take-it-or-leave-it prices.[4] In our experiment, we find nearly all households are willing to pay something for the filter; however, the median WTP is only 10 to15 percent of the cost of manufacturing and delivery.

Second, BDM induces random variation in both treatment status and price paid, conditional on willingness-to-pay. This allows the researcher to separate the effect of selection (those with a

[3]As demonstrated by Karni and Safra (1987) and Horowitz (2006a), BDM is not necessarily incentive-compatible for non-expected utility maximizers, a feature it shares with auctions.

[4]Note that the standard BDM mechanism estimates demand for only single unit-demand items, such as the water filter studied in this project. By modifying the mechanism to elicit the willingness-to-pay for additional increments of the good, BDM can be used to estimate demand for products where multiple units may be demanded by a single individual. See Hoffmann (2009) for an example.

higher WTP for a product may be more likely to use it or have a greater benefit) from a "sunk-cost effect" (a direct causal effect of price paid). These effects are typically difficult to disentangle, even with randomized prices, since an observed relationship between price and use among those who purchase at that price could include both effects.[5] Combining BDM with survey data on filter use, we find some evidence of a modest screening effect: use at the one-year follow-up increases slightly with WTP over most of the distribution of WTP. We find no evidence that price paid has a causal (sunk-cost) effect on use: conditional on one's willingness-to-pay, the price paid does not affect usage or outcomes.

Third, BDM generates detailed information on heterogeneous treatment effects with respect to willingness-to-pay (Heckman et al., 2006; Heckman and Vytlacil, 2007). Like TIOLI methods with random prices, BDM provides an exogenous source of variation in product allocation, allowing researchers to identify treatment effects of the product. But unlike TIOLI, BDM allows researchers to easily estimate heterogeneous treatment effects, i.e., whether and how treatment effects vary with willingness-to-pay. Intuitively, BDM reveals each individual's willingness-to-pay and then allocates the good to her randomly, conditional on that willingness-to-pay. We demonstrate how BDM can allow researchers to estimate the distribution of marginal treatment effects in a variety of field settings.[6] In our particular context, we find that the filter's benefits, as measured by reductions in diarrhea among children, increase with WTP over most of its distribution, consistent with the pattern of usage.

Balanced against these advantages, BDM faces a number of challenges. It is a novel mechanism with limited experience outside the lab. Moreover, our study presents a severe test of BDM. Numeracy among our subject pool was low and non-standard beliefs about probability were commonplace. We find that even in this setting, subjects were able to readily grasp the mechanism,

[5] One approach, pioneered by Karlan and Zinman (2009), is a two-stage randomization: first, a random initial offer price; second, a surprise discount after the subject has made her purchase decision. BDM generates this two-stage randomization, but does not require the second stage to be a surprise. It can therefore be implemented effectively in environments where individuals know or are likely to discover the allocation mechanism independently.

[6] The ability of BDM to improve information extraction from randomized control trials is emphasized by Chassang, Padró i Miquel and Snowberg (2012), who describe BDM as a type of a "selective trial" and cite an earlier version of this paper as an example of their theoretical mechanism's practical feasibility.

with suitable modification as described in Section 2.2.2, and provided sensible answers.

There is a substantial literature dealing with the implementation and behavior of BDM in university economics labs;[7] however, little is known about the practical applicability of BDM in a field setting.[8] To explore this issue, we randomly assigned respondents to either a BDM or TIOLI sales treatment. Results from both methods of demand elicitation follow a similar pattern; however, TIOLI acceptance rates were above the BDM-estimated demand curve. We explore a number of potential explanations and find that risk aversion accounts for much of the difference between the two experimental mechanisms. The BDM-TIOLI gap decreases across terciles of risk aversion and is close to zero in the least risk-averse third of the population. Overall, we note that neither mechanism precisely replicates typical market interactions and recommend careful, context-specific calibration before using either mechanism to predict market demand; however, we are encouraged by the demonstrated feasibility of BDM to enrich the information generated by randomized control trials.

The rest of the paper proceeds as follows. Section 2 explains the experimental setting and implementation. Section 3 describes the use of BDM for demand estimation and the correlates of willingness-to-pay. Section 4 explores the causal effects of price paid as well as the relationship between willingness-to-pay and use. Section 5 illustrates the use of BDM to estimate heterogeneous treatment effects. Section 6 evaluates the differences in responses between BDM and TIOLI and discusses practical implications. Section 7 concludes.

[7] See, for example, Smith (1982), Keller, Segal and Wang (1993), Bohm et al. (1997), Irwin et al. (1998), Wertenbroch and Skiera (2002), Noussair, Robin and Ruffieux (2004), Mazar et al. (2014) and Cason and Plott (2014). Horowitz (2006b) provides a helpful survey.

[8] Hoffmann, Barrett and Just (2009) use BDM to measure the gap between willingness-to-pay and willingness to accept for bed nets in Uganda. Recent field applications include Cole and Fernando (2012), Cole et al. (2014), Guiteras and Jack (2014), Guiteras et al. (2015a), and Guiteras et al. (2015b).

2 Experimental Setting and Design

2.1 Background

Lack of access to clean water is one of the most significant threats to health and welfare in the developing world, particularly rural Africa. Nearly 40 percent of Africans—and 52 percent of rural Africans—lack access to improved sources of drinking water. This has serious health consequences: diarrheal disease causes nearly 1.8 million deaths worldwide each year and is responsible for 17 percent of deaths of African children under five years of age. Poor water quality also harms the health of the living, contributing to diseases such as schistosomiasis, trachoma and worms. In Ghana, 26 percent of rural households lack access to clean water, and diarrheal diseases are the third leading cause of death for children under five years of age (World Health Organization, 2004, 2011).

Absent improved water sources, a variety of free and socially-marketed point-of-use purification methods have been distributed and sold to households (Clasen et al., 2007). We study the *Kosim* water filter (see Figure A1 in the Appendix), marketed and sold in northern Ghana by Pure Home Water, a non-governmental organization. The filter is highly effective at improving water quality and is appropriate for this context, since it removes particulates and pathogens from water without the use of chemicals or electricity (Miller, 2012). At the time of the study, the average cost of producing a filter and delivering it to a rural household in a village-level distribution was approximately GHS 21 (about $15). Demand for the filter is effectively zero at a break-even price, so the level of subsidy to provide, and consequently the relationship between price and access, use, and outcomes are key concerns for an NGO with a limited subsidy budget.

We offered the filter to 1,265 respondents across 15 villages in northern Ghana between October 2009 and June 2010. To select our sample, we identified villages in Northern Region of Ghana that had limited access to clean drinking water and had not previously been exposed to the *Kosim* filter. Within our study villages, we conducted our baseline survey and sales exercise with women who were primary caregivers of children.[9,10]

2.2 Data Collection and Experimental Setup

2.2.1 Preliminary Activities & Household Survey

MARKETING MEETING.[11] In each study village, we first conducted an initial village meeting, during which we provided a demonstration of the filter and the BDM and TIOLI sales mechanisms, following closely the standard practice of our evaluation partner. Two field staff performed a mock version of both BDM and TIOLI for a token item, such as chocolates or a bar of soap. The field staff also practiced the sales mechanisms with volunteers from the attendees, again for a token item. We informed villagers that a filter would be installed at the village health liaison's home and encouraged them to visit the liaison to see the filter working, taste the water and ask questions. We instructed the attendees that we would visit their households in approximately two weeks to offer them an opportunity to purchase the filter via one of the mechanisms we performed. Attendees were encouraged to discuss with their families what they were willing to pay for the filter. The two-week interim period was consistent with standard sales and marketing practices for the product and was chosen to allow families time to try the filter, determine their WTP, and obtain necessary funds.

On the same day as the marketing meeting, we conducted a comprehensive census of all residents of the village. With this information, we were able to identify the study subjects as defined above and perform random assignment of the sales mechanisms.

WATER QUALITY TESTING. Roughly one week after the village presentation and census, we visited each household to remind them of the upcoming sales visit and to answer any questions they

[9]These were primarily mothers, but occasionally were other relatives caring for children whose parents had migrated, were deceased or were permanently absent for other reasons. We also included pregnant women and women who might become pregnant (married and of childbearing age). This is the subject group of primary interest to the NGO.

[10]Most subjects live in extended patrilineal family compounds, which are small clusters of individual huts, usually enclosed by a wall. Many resources are shared within the compound, although in most cases each woman is responsible for providing water for her own husband and children. As described below, the sales mechanisms were randomized at the compound level and all inference is robust to clustering at the compound level.

[11]See Appendix Figure A2 for an illustrative timeline of the experiment in a typical village.

had about the filter. During this reminder visit, we took a 100 ml baseline sample of their stored drinking water for testing in the lab. Half of the samples were randomly selected to be tested for levels of *E. coli* and turbidity.[12]

HOUSEHOLD SURVEY. Roughly one week after the reminder visit, we conducted a survey and sales visit with each respondent. Respondents were compensated with a GHS 1 cash gift, awarded at the beginning of the survey.[13] The survey collected household demographic information, asset ownership, and education status. The baseline survey also elicited information on water collection and treatment practices, basic health knowledge, and recent episodes of diarrhea among household members.

2.2.2 Filter Sale / BDM Implementation

At the end of the survey, we conducted the sales experiment.[14] Respondents were randomly assigned in equal proportions to either a BDM or TIOLI sales treatment.[15] Treatments were randomly assigned at the compound level, stratified by number of respondents in the compound.[16]

The scripts for the sales were designed to be identical across treatments apart from language related to the specific sales mechanism. Each sale began with a practice round for a bar of soap with retail value of approximately GHS 1. The respondent was then given the opportunity to purchase the soap using the mechanism corresponding to her treatment group. After the practice round was

[12]Logistical constraints prevented testing all water samples. In addition, during these visits we conducted one of two health education treatments in randomly selected households. The first treatment was a general message describing the link between untreated water and health and explaining how the filter helps prevent diseases such as diarrhea. The second treatment provided similar health information but emphasized the dangers to children of untreated water and the potential benefits of the filter to children. The impacts of these treatments on WTP were minimal and are discussed in the Supplementary Materials.

[13]This was awarded in small denomination coins to ensure that respondents could submit reasonably fine-scale bids in the practice WTP game described below. It is possible that a cash gift influenced WTP for the filter by inducing goodwill toward the surveyor. However, because of the length of the survey there was always at least 30 minutes between the gift and the sales offer, which could ameliorate any "house money" effect.

[14]By conducting the sale at the end of a survey on water and health, we may have primed the respondent's demand for the filter. However, it was not feasible to conduct the sale first, because respondents, and especially respondents who were not able to purchase the filter, would quickly lose interest in the survey.

[15]Within each of these two broad categories, we conducted three sub-treatment to understand how the elicited willingness-to-pay might be affected by anchoring effects or strategic behavior. These sub-treatments are further described in Section 6. Results for the sub-treatments within each category are statistically indistinguishable and we group them together for the primary analysis.

[16]Compounds with three or more respondents were grouped into a single stratum.

complete, the respondent was given the opportunity to purchase the *Kosim* filter using the same mechanism.[17]

BDM TREATMENT. First, the surveyor read a brief description of the BDM procedure. We emphasized that the respondent would have only one chance to obtain the filter, could not change her bid after the draw, and must be able to pay that day. The surveyor then played a practice round for the bar of soap. The respondent was asked for her maximum WTP for the bar of soap. The surveyor reminded her that if she drew slightly more than her bid, she would not be able to purchase the soap. She was then allowed to adjust her bid. This process repeated until she was no longer willing to adjust her bid. At this point, the surveyor reminded her that if she drew a price equal to her bid she must be willing and able to make this payment. Before the random price was drawn, the surveyor reviewed various hypothetical outcomes to test her understanding. Once the final bid was established, the price was drawn and the subject either purchased or did not purchase the soap in accordance with the BDM mechanism. This protocol was based on extensive piloting of different procedures to maximize understanding in this population. [18]

The procedure for the filter was similar; however, at the completion of the sales process, the respondent, if successful, paid for the filter and received a receipt that could be redeemed for at a central location in the village, typically the health liaison's home.[19]

We did not require respondents to present the amount of cash they were willing to bid before the draw was made. Rather, we permitted the household to gather the money by the end of that day. Before the draw was made, we asked multiple times whether the respondent would have access to the funds. We did this to maintain realism: households routinely make small loans to

[17]Complete scripts for the BDM and TIOLI treatments are provided in the Supplementary Materials.

[18]Although developed independently, our protocol is similar to the "titration BDM" mechanism tested by Mazar, Koszegi and Ariely (2014). Prices were written on wooden beads and placed in an opaque cup. The subject draws the price herself. The prices were distributed uniformly from 0 to 100 in increments of 10 pesewas (GHS 0.10). We did not inform subjects about the distribution of prices.

[19]The distribution of prices was 0, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 5, 6, 7, 8, 9, 10, 11, 12 in equal proportions. The most salient difference between the procedure for the filter and for the practice round was that the filter was not physically present in front of the respondent during the bidding. We chose not to have surveyors bring the filters to compounds for two reasons: first, they are bulky and could break; and second, there is some instruction on assembly and care that should be given at the time the household receives the filter. This instruction was most efficiently provided at a central location.

each other for purchases. Of the 272 respondents who drew a price less than or equal to their bid, 269 (98.9 percent) completed the purchase. For the three respondents who did not, their failure to purchase appears to have been due to an inability to gather funds, e.g., because a family member was unavailable, rather than because of a lack of understanding.

We also tracked whether a losing respondent attempted (after the price was drawn) to purchase at a price above her final offer and asked all losing respondents whether they wish they had bid more. A substantial share (19.2 percent) exhibit ex post regret, stating that they wished they had bid more.[20] However, only 5.4 percent (18 of 331) actually offered to pay more than their final bid. Taken together, these results suggest that even in an environment with very low literacy and numeracy, the BDM mechanism, with suitable modifications, can be readily understood and incorporated into field experiments. Based on pilot results, we believe that multiple demonstration rounds with a different product or products and the understanding check after respondents stated their bids are essential to successful implementation.[21]

TIOLI TREATMENT. The standard take-it-or-leave it treatment was a simple sales offer at a randomized price. Similar to the BDM treatment, we first conducted a practice round for a bar of soap. We then presented the offer for the filter randomized in equal proportions across three prices: GHS 2, 4, and 6.[22] These prices were the approximate 25th, 50th and 75th percentiles of BDM bids in pilots in similar villages. Before revealing the offer price to the respondents, we emphasized that there would be no bargaining. If they accepted the offer price, respondents were allowed until the end of the day to obtain the necessary cash. If the respondent initially agreed to the purchase but was ultimately unable to obtain the funds, we code her as not purchasing.

[20]We find that ex post regret is highest for those who narrowly missed winning in BDM. Roughly 40 percent of those who missed by GHS 1 or less wished that they had bid more, with this percentage declining to approximately 12 percent among those who missed by GHS 5-10.

[21]The mean time required to complete the BDM sales section, including the practice round for soap and understanding confirmation, was 14.12 minutes, or 4.10 minutes (0.70 s.d.) longer than the 10.01 minutes required for TIOLI.

[22]As described in footnote 15 above, the TIOLI treatment was divided into three sub-treatments. Approximately one-third of respondents received a sub-treatment where soap and filter prices were determined by a draw from a cup. For the remaining respondents, the soap and filter prices were randomly assigned at the compound level. The Supplementary Materials discuss the sub-treatments in detail.

2.2.3 Follow-up Surveys

Two follow-up surveys were conducted approximately one month and one year after the initial sale.[23] Each survey measured whether households had retained the filters they purchased, whether the filters were being used and maintained properly, and recent diarrhea episodes of children aged 5 and under. The one-month follow-up survey was conducted with all households that had been surveyed at baseline. Due to funding constraints, we randomly selected eight of the original 15 study villages to be included in the one-year follow-up survey. We re-surveyed 87.1 percent of households in the one-month follow-up and 90.5 percent of targeted households in the one-year follow-up. Attrition is largely balanced along observable dimensions, although there is some imbalance in observables across attriters and stayers in the one-year follow-up. Most relevant for our treatment effects estimation, attrition in either round is not significantly related to the BDM draw or to the TIOLI price. Section S6 of the Supplementary Materials provides a more detailed analysis of attrition.

2.3 Sample Characteristics and Balance

Table 1 displays summary statistics of the sample and analysis of balance of the randomizations. Column 1 displays means of baseline characteristics for the full sample. Only 9 percent of respondents had ever attended school, and the average number of children aged 0 to 5 was 1.3 per respondent. Diarrhea incidence was relatively high: on average, households had 0.24 episodes of diarrhea among children aged 0 to 5 in the past two weeks. Only 19 percent of households had access to an improved water source year round.[24]

[23] The one-year horizon was chosen to correspond to about half of the filter's expected lifespan of two years. According to our data, 40-50 percent of distributed filters were undamaged and in use as of the one-year follow-up survey.

[24] Due to logistical limitations, water quality data (*E. coli* counts and turbidity measures) was measured for only half of the sample. Since households were randomly selected for water quality testing, this explanatory variable data is, by design, missing completely at random (MCAR). For simplicity, we incorporate into the regression models an indicator variable for whether or not a variable is observed. As demonstrated by Jones (1996), this method can still yield biased estimates even when data are MCAR. Complete-case analysis (i.e., using only those observations without missing water quality measures) is unbiased, but throws away data. We use the complete-case data to estimate bounds on the potential bias in the indicator method. Due to low pairwise correlation between water quality measures and other explanatory variables, these potential biases are small and do not alter the economic or statistical significance of

Columns 2 and 3 display the sample means by treatment (BDM vs. TIOLI), and Column 4 displays the difference in means between the two treatments groups. We observe a few marginally significant differences between the two groups: 0.13 fewer children aged 0 to 5 per household in the BDM treatment ($p < 0.1$), 0.17 more children aged 6 to 17 ($p < 0.1$), 0.07 fewer children aged 0 to 5 with diarrhea in the past two weeks ($p < 0.1$), and 0.55 fewer respondents in the compound ($p < 0.1$).

In Column 5, we check the balance of the BDM draw randomization by regressing the BDM draw on the same set of characteristics as well as the BDM bid. Of the 13 variables in the regression, one is significant at the 10-percent level: a higher number of respondents in the compound is associated with a higher draw ($p < 0.01$). Column 6 regresses the TIOLI price on baseline characteristics. Here, higher assigned prices were associated with more children aged 6 to 17 with diarrhea in the past two weeks ($p < 0.1$) and higher turbidity in stored water ($p < 0.01$).

3 Willingness to Pay, Demand and Correlates

This section describes the advantages of using BDM to elicit each individual's willingness-to-pay, presents the results of our demand estimation, evaluates the correlates of willingness-to-pay, and compares results to those obtained by TIOLI. Many of the benefits of BDM stem from its ability to provide an exact value for each individual's WTP, with precision limited only by the desired granularity of the researcher.[25] In contrast, a TIOLI offer provides only a bound. For example, if a respondent accepts a TIOLI offer of GHS 4, we can only conclude that her WTP was at least GHS 4. Similarly, if she rejects an offer of GHS 6, we can only conclude that her WTP was less than GHS 6. BDM's precision allows for direct, nonparametric estimation of demand at all potential prices and facilitates estimating relationships between WTP and important observables, e.g., wealth and health status. TIOLI could also provide such information but requires a substantially

estimates. All results are robust to multiple imputation (Rubin, 1996).

[25] In principle, it is possible to measure exact maximum WTP down to the smallest available denomination. For example, respondents in our study were free to bid in increments of pesewas (1/100th GHS); however, in practice most final bids were in increments of 50 pesewas (GHS 0.5).

larger sample size: nonparametric estimation of P points on a demand curve at a desired level of precision requires a TIOLI sample approximately P times larger than what is required using BDM.[26] Furthermore, in many settings it may not be feasible to expand the TIOLI sample size sufficiently. For example, we may be interested in heterogeneous responses to treatments or mechanisms at the level of clusters of finite size, e.g., educational interventions or social learning that take place at the village level.

Auctions can also provide precise data on WTP; however, unlike auctions BDM is robust to intra-community conflict or collusion and allows individuals with low WTP to receive the good with positive probability. By varying the price distribution, the researcher can target any desired probability of receiving treatment, subject only to preserving incentives for truthful reporting of preferences.[27]

3.1 Estimating Demand

Figure 1a shows the inverse demand curve generated across all 15 villages using data from the 608 BDM respondents. At each point along the x-axis, we plot the share of BDM subjects whose bid was greater than or equal to p, along with a point-wise 90-percent confidence band robust to clustering within compound. The average bid was GHS 3.0 with a median of GHS 2.5.

There are several features of this inverse demand curve worth noting. First, WTP is almost universally positive: nearly 95 percent of respondents bid at least GHS 1.[28] At the same time, WTP is low relative to the cost of the filter: the median bid of GHS 2.5 corresponds to approximately 10 to 15 percent of the cost of manufacturing and delivery. This result is consistent with the relatively

[26] Alternatively, for a fixed sample size, TIOLI confidence intervals at each point will be wider by a factor of approximately $\sqrt{P}$.

[27] Truthful reporting constitutes a unique optimum if and only if an individual's WTP is in the support of the price distribution. If she believes that all prices are above or below her WTP, she will have no incentive to bid precisely her WTP. Similarly, if the gap between prices is wide, a subject need not reveal her exact WTP between two prices.

[28] "House money" effects could provide one plausible explanation for low elasticity at zero; individuals may be be less price sensitive when spending funds given to them as a participation fee by the surveyors. The sale of soap at a randomized before the filter bid allows us to test for and rule out such effects. We find no relationship between participation fees remaining after soap purchase (i.e., 1-BDM draw for soap among those who purchased soap) and the BDM filter bid, conditional on WTP for soap.

low willingness-to-pay for water treatment and other health goods found in previous work (Kremer and Holla, 2009; Ahuja et al., 2010). Figure 1b displays the price elasticity of demand at prices from 0 to 10 GHS as calculated from the BDM-elicited willingness-to-pay data; demand at low prices is substantially less elastic than has been observed in other settings.[29]

For comparison, Figure 1a also displays the share of TIOLI subjects who purchased at each of the three, randomly assigned TIOLI price points. As with BDM, the 90-percent confidence intervals are robust to clustering within compound. The pattern of demand is consistent across the two mechanisms, but despite the theoretical equivalence of TIOLI and BDM, at each of the three TIOLI price points, the TIOLI purchase rate is greater than the share of BDM respondents bidding that amount or higher(BDM-TIOLI difference: -0.182 at 2 GHS ($p = 0.000$); -0.163 at 4 GHS ($p = 0.002$); -0.100 at 6 GHS ($p = 0.012$)).[30] We discuss the economic significance and possible reasons for these differences in Section 6.

3.2 Correlates of Willingness to Pay

Understanding the relationship between WTP and household characteristics can inform how pricing policies target particular types of households. In this section we demonstrate the use of BDM to estimate this relationship directly and show that BDM-elicited valuations display sensible correlations with several key characteristics. Previous studies have found limited evidence that higher WTP for health goods is related to health characteristics or wealth (Ashraf, Berry and Shapiro, 2010; Cohen and Dupas, 2010; Cohen, Dupas and Schaner, 2015). This is a common problem in the consumer behavior literature: heterogeneity in choice is often only weakly correlated with standard consumer attributes (Browning and Carro, 2007; Nevo, 2011).

[29] Other studies of health goods in developing countries have found mixed evidence of price sensitivity at low or zero prices: Kremer and Miguel (2007), Kremer et al. (2009) and Cohen and Dupas (2010) find that demand falls sharply at any positive price, while Ashraf et al. (2010) and Cohen et al. (2015) find less sensitivity at zero.

[30] See Section S2 of the Supplementary Materials for additional details, including regressions conditioning on our standard set of household controls.

We model the relationship between WTP and baseline characteristics and behaviors as

$$\mathrm{WTP}_{ic} = \alpha_0 + X_{ic}^{'}\beta + \varepsilon_{ic}, \tag{1}$$

where X_{ic} is a vector of characteristics of interest for subject i in compound c, and ε_{ic} is an error term, possibly correlated at the compound level. We consider a variety of characteristics and behaviors. First, we include variables relating to household demographic composition (number of adult males and adult females in the compound, subject's number of children aged 0 to 5 and 6 to 17), an indicator for whether the subject has attended school, and a wealth index (the first principal component of a set of variables on ownership of durables, land and livestock). Second, we include two variables that indicate the number of children aged 0 to 5 and 6 to 17 who have had diarrhea in the past two weeks. Finally, we include four variables relating to access to improved water, water treatment practices, and water quality.

Estimating Equation 1 using our BDM sample is straightforward: we run an ordinary-least-squares regression of the BDM bid on the vector of characteristics. Column 1 of Table 2 presents the results of this regression. The BDM bid is positively related to the number of children 0 to 5 with diarrhea, a result significant at the 10-percent level. One additional child with diarrhea in the household (conditional on the total number of children), increases the BDM bid by GHS 0.55. The BDM bid is also positively related to durables ownership (significant at the 10-percent level), and negatively related to turbidity (significant at the 5-percent level).[31] The p-value of the F-test of the regression equals 0.024, indicating observable characteristics are jointly significantly related to BDM-measured WTP.

These results show that, at least in this context, households' health and wealth influence their willingness-to-pay for health products. By generating precise, individual WTP measures BDM facilitates discovery of these correlations; however, we note that, consistent with the aforementioned

[31]The correlations with children's health status and asset ownership are as one would expect. The turbidity result suggests that household with clearer water have higher WTP. We interpret the latter result as an indication that households who let their water settle, access cleaner appearing sources, or use other methods to remove turbidity have higher demand for the filter. During the marketing phase, households are told that the filter will require more frequent cleaning if using turbid water.

consumer behavior literature, much of the heterogeneity across subjects remains unexplained.

For comparison, we conduct the analogous exercise using TIOLI subjects, estimating a binary discrete choice model via probit. For respondent i assigned price p, we observe

$$\text{buy}_{i,p} = 1\{\text{WTP}_i \geq p_i\} = 1\{\text{WTP}_i - p_i \geq 0\} = 1\left\{\alpha_0 + X_{ic}^{'}\beta + \varepsilon_{ic} - p_i \geq 0\right\}. \tag{2}$$

where $\text{buy}_{i,p}$ is an indicator equal to 1 if respondent i agreed to buy when assigned price p. We normalize the coefficient on price to -1, so that the estimates of the coefficients β are directly interpretable in terms of WTP and are comparable to those obtained by estimating Equation 1 with BDM subjects. Results for the TIOLI subjects are presented in Column 2 of Table 2, with estimated differences between BDM and TIOLI in Column 3.[32] In the TIOLI sample, most of the estimates are statistically indistinguishable from zero, but, notably, the variable indicating children aged 6 to 17 in the household with a recent diarrhea episode has a significantly *negative* association with WTP. The corresponding variable for younger children is also negative but not significant. As shown in Column 3, there are a few significant differences between the estimates for BDM and TIOLI. In the cases where coefficients differ significantly, the BDM coefficient conforms more closely to our prior beliefs. For example, we expect health concerns to be more salient to educated parents and to parents whose young children had more recently had diarrhea. We are unable to explain why the pattern of correlations differs across these two mechanisms and believe further work along this dimension may be helpful in understanding heterogeneity in consumer demand.

4 Screening and Sunk-Cost Effects

The BDM mechanism embeds a double randomization that allows researchers to separately identify two factors that may be important for pricing policy: the *sunk-cost effect*, i.e., the causal effect of price paid conditional on WTP, and the *screening effect*, i.e., the correlation between WTP and

[32] To test for differences in the estimated β, we estimate Equations 1 and 2 as a system using seemingly unrelated regression.

use. Because the price draw is random, we can test for causal effects of price paid by comparing measures of use for subjects with the same WTP but who paid different prices. Similarly, we can test for screening effects by comparing use for subjects with different WTP while controlling, if needed, for price paid. For example, BDM generates the following experiment: consider three subjects, each willing to pay GHS 6 for a filter; one doesn't receive the filter; another pays GHS 6; and the other pays GHS 2. Thus, at every level of WTP above the minimum price, there is variation in both allocation and the price paid conditional on allocation. In this section, we illustrate the usefulness of this feature using the two waves of follow-up data collected approximately one month and one year after the filter sale. We find no evidence of a causal effect of price paid and modest evidence of a positive association between use and WTP.

We collected three objective indicators of use from all subjects who purchased the filter: (i) whether the filter was found in the compound and was undamaged; (ii) whether water was in the plastic storage reservoir; and (iii) whether water was in the clay filter pot. These measures are normalized and aggregated into a single usage/effort index following Kling, Liebman and Katz (2007).

To test for a causal effect of price paid, we estimate

$$\text{use}_{ic} = \alpha_0 + \alpha_1 D_{ic} + \alpha_2 f(\text{WTP})_{ic} + \varepsilon_{ic}, \tag{3}$$

where use_{ic} represents the usage measure, D_{ic} is the respondent's draw, and $f(\text{WTP})_{ic}$ is a cubic polynomial of bid. It is important to control adequately for WTP since, although the price draw was unconditionally random, conditional on receiving the filter it is positively correlated with WTP.

Table 3 presents results from OLS estimation of Equation 3. Panel A shows that there is no detectable effect of the price paid on use in the one-month follow-up. Panel B shows a similar null result in the one-year follow-up data. Taken together, this suggests there is no significant sunk-cost effect.

We next turn to the relationship between WTP and usage in the BDM treatment.[33] We perform

[33] The relationship between WTP and usage in the TIOLI treatment is described in the Supplementary Materials.

this analysis in two ways. First, we regress usage measures on BDM bid among those who purchased the filter. Given the absence of evidence for causal effects of price paid reported above, our primary specification does not control for price paid. The results of these regressions are shown in Table 4. In the short term, use is generally high and there is no evidence for an association between WTP and use. At the one year follow-up, GHS 1 of higher WTP is associated with a 2.7 percentage point higher likelihood of having water in the plastic storage vessel ($p < 0.10$). WTP is positively associated with a higher level of the aggregated use index, but this association is not statistically significant.

Second, we model the relationship between usage and WTP non-parametrically for comparability with our analysis of heterogeneous treatment effects in Section 5 below. We restrict the sample to households with children aged 0 to 5 and estimate the relationship between WTP and the usage indices using kernel regression. Figures 2a and 2b graph this relationship in the one-month and one-year follow-up surveys, respectively. Using the one-year data (Figure 2b), we observe a pattern of increasing usage with respect to WTP over most of the distribution. This pattern is similar to that of the one-year treatment effects and consistent with effort as a mediator of treatment effects as discussed in Section 5.2..

5 Heterogeneous Treatment Effects

5.1 Theory

By eliciting respondents' willingness-to-pay and allocating treatment randomly conditional on this value, BDM provides an easily implementable way to recover the marginal treatment effects (MTEs) introduced by Björklund and Moffitt (1987) and extended by Heckman and Vytlacil (2005, 2007). The intuition is analogous to that of the structural approach developed developed in the prior literature. At each level of WTP, individuals are randomly assigned to either receive the product or not. However, the structural approach requires estimation of a selection equation and then calculating MTEs with respect to a predicted probability of treatment. With BDM we can simply use the price draw as a local instrumental variable conditional on the elicited WTP.

This section describes how we can use this instrument to estimate the distribution of MTEs across the support of the WTP distribution. As Heckman and Vytlacil (2005) detail, the policy object of interest is not always represented by the local average treatment effect from a linear instrumental variable. In our setting, it may be that those who are likely to benefit the most from a product are aware of this and have the resources to pay for it, in which case charging for the product targets those with higher treatment effects (Cohen et al., 2015). On the other hand, it may be that individuals who are most likely to benefit are either unaware of the extent to which they will benefit or are simply too poor or too credit constrained to purchase the product, in which case higher prices will likely restrict access without improved targeting (Cohen and Dupas, 2010). Once we have estimated the distribution of MTEs, the distribution can be integrated with different weights to answer different policy questions. Since price is a natural policy variable in this setting, recovering the distribution of MTEs with respect to willingness-to-pay allows one to easily perform counterfactual policy experiments such as evaluating targeted subsidies or free distributions.

With BDM, this exercise is relatively straightforward. Consider the following outcome equation:

$$y_{jic} = \beta_0 + \beta_1 T_{ic} + u_{jic}, \tag{4}$$

where y_{jic} is the outcome of interest (an indicator for one or more cases of diarrhea in the previous two weeks) for child (aged 0 to 5) j of subject i in compound c, T_{ic} is an indicator for whether subject i purchased a filter, and u_{jic} represents unobservable determinants of health. We suppress covariates for brevity. The classic identification problem comes from the fact that u_{jic} is likely to be correlated with T_{ic}. Unobservable attributes or behaviors affecting child health may be different on average between households that buy or do not buy a water filter. Statistically, $E\left[T_{ic}u_{jic}\right] \neq 0$, which leads to bias when β_1 is estimated using ordinary least squares.

Overcoming this identification problem requires an instrument that is correlated with T_{ic} but uncorrelated with u_{jic}. Both BDM and TIOLI provide such an instrument. In the case of BDM, this is the price draw; in the case of TIOLI, it is the offer price. We denote both as P_{ic}. In both cases, P_{ic} is random, so it is uncorrelated with u_{jic}, and because demand slopes down over our range of P_{ic}, it is correlated with T_{ic}. In a simple linear two-stage least squares (2SLS) framework, the first-stage equation is

$$T_{ic} = \gamma_0 + \gamma_1 P_{ic} + v_{ic}, \tag{5}$$

where P_{ic} is the BDM draw, and the predicted values $\hat{T}_{ic}$ are substituted into Equation 4 to obtain $\hat{\beta}_1^{IV}$.

$\hat{\beta}_1^{IV}$ estimates a local average treatment effect, where the complier group is those whose treatment status would change as a result of experimental variation in the offer price. Define $\beta_1(w)$, the average treatment effect for those with WTP $= w$, and $F_{\text{WTP}}(w)$, the CDF of WTP in the study population. The parameter estimated by instrumenting with a randomized price, β_1^{IV}, is a weighted average of these $\beta_1(w)$, where the weights depend both on $F_{\text{WTP}}(w)$, the distribution of WTP in the study population, and on the range of prices used in the randomization.[34]

Because BDM both reveals WTP and produces random variation in filter allocation at every level of WTP, it is possible to recover more information about $\beta_1(w)$. With a sufficiently large sample, it would be possible to estimate $\beta_1(w)$ at every elicited level of WTP. This is the distribution of marginal treatment effects discussed in the extensive literature on heterogeneous treatment effects. The key advantage of BDM is that it allows us to observe the relevant selection characteristic rather than treating it as a latent variable to be estimated.[35] In practice, we conduct a set of kernel IV regressions, estimating Equations 4 and 5 in the neighborhood of each level of WTP to obtain $\hat{\beta}^{KIV}(w)$ at a set of evaluation points W.

[34]The set of prices in TIOLI and the range of price draws in BDM determine the complier populations. These prices, combined with their distribution and the distribution of WTP in the subject population, determine the weights for calculating the local average treatment effect for a particular experimental mechanism. See Imbens and Angrist (1994) for details on the calculation of these weights.

[35]See, for example, Heckman and Vytlacil (2001), Heckman and Vytlacil (2005) and Heckman et al. (2006). Given an instrument Z, the MTE at z, $\beta(z)$, is the effect on those who are indifferent between treatment and non-treatment when $Z = z$. Our kernel IV estimates exactly this parameter, since by definition those with WTP $= z$ are indifferent between treatment and non-treatment when the price is z. We are grateful to Sergio Urzua and Ed Vytlacil for discussions on this point.

5.2 Application

We first present conventional IV estimates using the the BDM draw as an instrument for take-up among BDM subjects and the randomly-assigned TIOLI price as an instrument for take-up among the TIOLI subjects. These serve as reference points to demonstrate the advantages of using BDM to recover MTEs. We estimate

$$y_{jic} = \beta_0 + \beta_1 T_{ic} + x'_{ic}\beta_2 + u_{jic} \tag{6}$$

by linear two-stage least squares, where y_{jic} is the outcome of interest is an indicator for whether child j of subject i in compound c has had one or more cases of diarrhea in the previous two weeks, T_{ic} is a dummy for whether subject i purchased the filter, and x_{ic} is a vector of covariates. To instrument for the endogenous treatment variable, we estimate the first-stage equation

$$T_{ic} = \gamma_0 + \gamma_1 P_{ic} + x'_{ic}\gamma_2 + v_{ic}, \tag{7}$$

where P_{ic} is the TIOLI offer price for TIOLI subjects and the BDM draw for BDM subjects. Since P_{ic} is random, it is uncorrelated with u_{jic} and therefore it is a valid instrument for treatment.

Panel A of Table 5 presents results from this estimation for our short-term (one-month) follow-up data. In columns (1) and (2), we use only the TIOLI observations; with raw 2SLS in column (1) and adding covariates in column (2); in columns (3) and (4), we use only the BDM observations; and in columns (5) and (6) we pool the TIOLI and BDM data. Between the two elicitation mechanisms, the estimates have the same sign and are of similar magnitude. In Panel B we examine our long-term data, collected in a random sub-sample of half our villages roughly one year after the filter sale. After one year, there is no evidence that the filter was effective at reducing child diarrhea in either the TIOLI or BDM sample. In fact, the IV point estimates are *positive.* The filter appears to have increased the likelihood of childhood diarrhea. While precisely identifying the mechanisms behind this result is beyond the scope of our data, we speculate that it could be driven by compensatory behavior on the part of respondents, as has been found with clean water interventions in other contexts (Bennett, 2012).

The usefulness of BDM becomes apparent when we use the kernel IV approach described above to estimate the relationship between treatment effects and WTP. Beneath the estimated local average treatment effects there is substantial heterogeneity. The outcome variable, as above, is an indicator for whether the child has had one or more cases of diarrhea in the previous two weeks. We estimate treatment effects $\hat{\beta}^{KIV}(w)$ for each GHS 0.1 step from GHS 1 to GHS 6, which correspond approximately to the 0.1 and 0.9 quantiles of WTP in the BDM sample. We use an Epanechnikov kernel and Silverman's rule of thumb to choose the bandwidth.

We present the results in Figure 3. In the top panel (Fig. 3a), we consider the effect at the short-term (one-month) follow-up survey. The null short-term effect in the population overall (Table 5, Panel A, Col. 3 and 4) is consistent across the distribution of WTP, with no group clearly benefiting.[36] In the bottom panel (Fig. 3b), we repeat this analysis using the one-year follow up data. Although standard 2SLS detected no average treatment effect in the sample (Table 5, Panel B, Col. 3 and 4), Figure 3b reveals important heterogeneity. Estimated treatment effects are negative at low levels of WTP, i.e., diarrhea increases. Treatment effects then increase with WTP over most of its distribution, leveling off at about GHS 3.5. Above GHS 5, point estimates again decrease, although this is not precisely estimated.

This pattern resembles that of the relationship between WTP and use (Fig. 2), suggesting that effort (maintaining and using the filter) is an important determinant of impacts.[37] As shown in Figure 2b, low-WTP households used the filter less intensively. Regarding the observed negative treatment effect for those with a low willingness-to-pay, we speculate that drinking water may have become contaminated as a result of improper use. Consistent with Bennett (2012), it is also possible that low-WTP households may have been more likely to engage in compensatory behavior

[36]Since the distribution of WTP is concentrated in the lower values (median 2.5), the effective sample size falls as WTP increases. However, the effect on precision is mitigated by variation in the strength of the instrument with respect to WTP. See the Supplementary Materials for sample sizes and Shea's partial R-squared statistics for each level of WTP at which we took a kernel estimate (Shea, 1997).

[37]See Chassang et al. (2012) for an extensive discussion of how effort may affect estimated treatment effects in randomized experiments and its implications for external validity.

that more than offset the benefits from the filters.

As noted above, the estimated distribution of MTEs allows us to perform counterfactual pricing policy experiments. While we are hesitant to draw general conclusions, the pattern of increasing treatment effects across most of the WTP distribution suggests that higher prices would allocate the filter to those who would benefit most. In fact, because the estimated treatment effects are negative for those with low WTP, a price of about GHS 2.9 would effectively screen out those with negative expected treatment effects. A price of approximately GHS 3.5 would maximize the average treatment effects among those purchasing the filter. Note that such prices still reflect a substantial subsidy, on the order of 85 percent of the cost of manufacturing and delivering the filter.

6 Discussion of Mechanism Effects

The primary aim of this study was to assess the use of BDM to increase the information generated by randomized experiments and improve researchers' ability to identify economic mechanisms. A secondary goal was to compare BDM to TIOLI offers at randomized prices, which generate less information but have been more commonly used in field settings. To that end, we designed the study both to identify differences in elicited willingness-to-pay across the two mechanisms and to explore potential explanations for any differences that we found. The theoretical literature on the BDM mechanism finds that risk-aversion is likely to generate differences between the BDM bid and the maximum acceptable TIOLI offer for non-expected utility maximizers, and this difference is likely to be increasing in risk aversion (Safra et al., 1990; Keller et al., 1993). In our setting, there are multiple sources for deviations from expected utility maximization including loss aversion, ambiguity aversion, subjective beliefs about the pricing mechanism, and non-standard beliefs about probability.[38] We therefore included survey questions and sub-treatments designed to identify individual or mechanism characteristics that could be correlated with any differences.

[38]Among our sample, based on survey responses to questions on hypothetical gambles, 41.6% of subjects exhibit some degree of ambiguity aversion, 30.4% exhibit loss aversion, and 64.6% exhibited at least one of the two.

To examine the relationship between the BDM-TIOLI gap and risk aversion, we divide the sample into terciles of risk aversion and estimate the gap separately for each tercile. To obtain a measure of risk aversion, we first identified the sum at which the subject was indifferent between receiving that sum with certainty vs. a 50-50 gamble for a gain of 8 GHS.[39] We then repeated this exercise in the loss domain (what sum would the subject be willing to pay to avoid a 50-50 gamble for a loss of 8 GHS), and in a gain-loss domain (what sum would the subject be willing to pay to avoid a 50-50 gamble on winning 4 GHS vs. losing 4 GHS, or, if the subject were risk-loving, how much the subject would need to be compensated to forgo such a gamble). In our analysis, we use the first principal component of these three measures but the results are robust to other methods of combining them. We classify subjects by tercile (high risk aversion, medium risk aversion, low risk aversion) and study the BDM-TIOLI gap separately by tercile of risk aversion, both unconditionally and controlling for household observables.

To compare the mechanisms, we collapse the more precise individual willingness-to-pay information from BDM to the binary purchase indicators generated by TIOLI. Our outcome variable is $\text{buy}_{i,p}$, which represents subject i's purchase decision when facing a price $p \in \{2,4,6\}$. For TIOLI subjects, this is just whether they agreed to purchase at the offer price. For BDM subjects, $\text{buy}_{i,p} = 1\{\text{WTP}_i \geq p\}$, where WTP_i is subject i's valuation. As noted above, we classify respondents into terciles by risk aversion, creating RA_i^1, RA_i^2, RA_i^3 to indicate that subject i is in the first (most risk-averse), second, or third (least risk-averse) tercile, respectively. We then estimate

$$\text{buy}_{icp} = \sum_{t=1}^{3} \alpha_p^t \text{RA}_i^t + \sum_{t=1}^{3} \beta_p^t \left(\text{RA}_i^t \times \text{BDM}_i\right) + x_{ic}'\gamma + \varepsilon_{icp}, \tag{8}$$

where BDM_i is an indicator for whether subject i was assigned to the BDM mechanism. For each price p, α_p^t represents the purchase probability for TIOLI subjects in the t-th tercile, while β_p^t represents the "BDM effect" in the t-th tercile. The raw differences (i.e., without controls x_{ic}) are presented in Figure 4. The top panel plots the estimated coefficients $\hat{\beta}_2^1$, $\hat{\beta}_4^1$, $\hat{\beta}_6^1$, with

[39] More precisely, for each of the amounts 0.50 GHS, 1 GHS, ..., 5 GHS, we asked whether they would prefer that sum with certainty, the 50-50 gamble for 8 GHS, or if they were indifferent, and identified the cross-over point.

90% confidence intervals, for tercile 1 of risk aversion (the most risk-averse subjects), while the middle and bottom panels plot the same set of coefficients for terciles 2 and 3 (the least risk-averse subjects), respectively. As Figure 4 makes clear, the BDM-TIOLI gap is largest among the most risk-averse subjects (mean BDM effect -0.200, $p = 0.000$), and has largely closed among the least risk-averse subjects (mean BDM effect -0.051, $p = 0.425$). These results are unconditional, but robust to controlling for a large set of household controls (see Figure A3).

We also hypothesized that the stated prices in the TIOLI treatment could cause respondents to anchor their valuations to those prices. We included three sub-treatments to test for this possibility. In the "anchoring" treatments for both BDM and TIOLI, we informed subjects that the price of the filter in the Tamale town market (the nearest market town) was GHS 20. Based on our pilot results, we believed this information would dominate any conveyed in the TIOLI price, placing both mechanisms on equal footing and allowing us to estimate any anchoring or signaling effects from the offer price. However, this anchoring treatment did not produce any consistent effect on BDM bids or TIOLI purchase behavior.[40]

We also included a "random TIOLI" sub-treatment, in which the TIOLI offer price was drawn by the respondent from a cup of numbered wooden beads, the same mechanism used to determine the BDM price. The aim was to make salient the arbitrariness of the TIOLI prices and reduce the likelihood that they were serving as signals of quality.[41] Based on our pilot results and the evidence that in some settings BDM bids are sensitive to the underlying price distribution (Bohm et al., 1997; Urbancic, 2011; Mazar et al., 2014), we hypothesized that the randomness in the price draw may contribute to the BDM-TIOLI gap, through a failure to reduce compound lotteries, subjects' general discomfort with randomness and ambiguity, or other departures from expected utility maximization. However, demand under the random TIOLI treatment was statistically indis-

[40]There was a significant effect on TIOLI demand at GHS 4 (-0.233, $p < 0.05$), but there was no effect at the other TIOLI prices or in BDM bids. See Section S3 of the Supplementary Materials for further details on the statistical tests, and Figure S3, Table S3 and Table S4 for results.

[41]Unrelated to explaining potential differences between BDM and TIOLI, we also conducted a "market study BDM" treatment in which we told respondents that we were using the information from the study to help decide on the future price of the filter in similar villages. If strategic bidding was important, then this sub-treatment could lead to enhanced strategic bidding and decrease BDM bids.

tinguishable from standard TIOLI, indicating that our efforts to equate the perceived randomness in the two mechanisms had no effect on subjects' purchase behavior.[42] We note, however, that our modifications were designed to increase the perceived randomness of the TIOLI mechanism and speculate that reducing the perceived randomness of the BDM mechanism may narrow the gap. For example, framing BDM as a set of contingent take-it-or-leave-it decisions ("Suppose the price is X. Will you agree to buy, or will you refuse?") may lead subjects to focus more on how they value the product relative to a fixed sum of money and less on the randomization.[43]

Finally, the gap in elicited WTP between BDM and TIOLI also does not appear to be driven by lack of familiarity of the filter and uncertainty of its benefits. Although the sale of soap was primarily intended to be a practice round for the elicitation mechanism, the data provides suggestive evidence of the BDM-TIOLI gap for a more familiar product. Using these data, we find a similar difference in elicited WTP between the mechanisms: BDM predicts between 8 and 45 percentage points lower purchase at the TIOLI price points for the soap (results not shown).

Is the gap between the TIOLI acceptance rates and the demand curve calculated from BDM bids meaningful? This is an important question for future research. In our setting, the differences between mechanisms are locally meaningful but quite small relative to the production costs of the filter. Both mechanisms are research tools for experimental settings, and we recommend caution when using either mechanism to predict market behavior. BDM looks quite different from typical market interactions, and take-it-or-leave-it offers can themselves be unusual in environments where fixed, posted prices are rare and bargaining common.

If the aim is to accurately predict market demand, one should map experimental results to actual market demand. The literature on mechanism effects for price elicitation has largely focused on comparing across mechanisms (e.g., Rutstram, 1998; Noussair et al., 2004). Where BDM has been compared to market demand it appears to generate accurate predictions (Miller et al., 2011), but evidence here is limited. We know of no research at this time that tests both mechanisms as pre-

[42] See Section S3 of the Supplementary Materials for further details on the statistical tests, and Figure S3 and Table S4 for results.

[43] See de Meza and Reyniers (2013) for an application measuring willingness-to-accept compensation (WTA) for beeswax candles among London School of Economics undergraduates.

dictors of actual market demand. Our experience suggests that relationship will likely depend on context and individual characteristics. However, for researchers interested in understanding underlying mechanisms, BDM greatly enriches the information generated by randomized experiments and, implemented along the lines we describe, readily applicable to field settings.

7 Conclusion

This paper demonstrates that BDM can be a useful tool to enhance the information obtained from randomized experiments. We use BDM to elicit a complete demand curve for household water filters in rural Ghana. With minor modifications to the BDM mechanism commonly used in the lab, most notably, guided practice rounds for unrelated products and confirmation checks after individuals state their valuation, the procedure can be readily understood. Even in an environment with extremely low literacy and numeracy, BDM produced sensible results. We then estimate the effects of price paid on use of the filters and estimate how use varies by WTP. Finally, we combine the precise information on WTP, random allocation of the filter conditional on WTP, and data from two follow-up surveys to estimate heterogeneous treatment effects of the filters by WTP.

From the standpoint of pricing policy, our BDM results suggest that small positive prices for the filter do not substantially decrease demand. As prices increase and demand starts to fall, those with the least benefit from the filter are effectively screened out of receiving it. At the same time, median WTP is low relative to the cost of the filter, approximately 10 to 15 percent of cost. While the price mechanism may improve allocative efficiency, a subsidy-free market for this product in this population would not be viable.

Returning to methodology, we observe differences between demand elicited under BDM and TIOLI and note that neither experimental mechanism accurately captures all aspects of typical market interactions. We see significant value in future research that links experimentally elicited demand with actual market behavior. This would be particularly valuable for the BDM mechanism, which can greatly increase the information generated by field experiments in a wide variety of settings. As shown here, this can further enrich the ongoing policy debate regarding the pricing of preventative and curative health products. Moreover, recent work is demonstrating the value of detailed WTP measures in other contexts, such as labor markets (e.g., Guiteras and Jack, 2014), weather insurance Cole et al. (2014), and environmental conservation (e.g., Jack, 2013).

References

Ahuja, Amrita, Michael Kremer, and Alix Peterson Zwane, "Providing Safe Water: Evidence from Randomized Evaluations," *Annual Review of Resource Economics*, 2010, *2*, 237–256.

Ashraf, Nava, James Berry, and Jesse M. Shapiro, "Can Higher Prices Stimulate Product Use? Evidence from a Field Experiment in Zambia," *American Economic Review*, December 2010, *100* (5), 2383–2413.

Bennett, Daniel, "Does Clean Water Make You Dirty? Water Supply and Sanitation in the Philippines," *Journal of Human Resources*, 2012, *47* (1), 146–173.

Björklund, Anders and Robert Moffitt, "The Estimation of Wage Gains and Welfare Gains in Self-Selection Models," *The Review of Economics and Statistics*, 1987, *69* (1), 42–49.

Bohm, Peter, Johan Linden, and Joakin Sonnegård, "Eliciting Reservation Prices: Becker-DeGroot-Marschak Mechanisms vs. Markets," *The Economic Journal*, July 1997, *107* (443), 1079–1089.

Browning, Martin and Jesus Carro, "Heterogeneity and Microeconometrics Modeling," in T. Persson R. Blundell, W. Newey, ed., *Advances in Theory and Econometrics, Vol. 3*, Cambridge, England: Cambridge University Press, 2007.

Carson, Richard T. and W. Michael Hanemann, "Contingent Valuation," in Karl-Goran Maler and Jeffrey R. Vincent, eds., *Valuing Environmental Changes*, Vol. 2 of *Handbook of Environmental Economics*, Elsevier, 2005, chapter 17, pp. 821–936.

Cason, Timothy N. and Charles R. Plott, "Misconceptions and Game Form Recognition: Challenges to Theories of Revealed Preference and Framing," *Journal of Political Economy*, 2014, *122* (6), 1235–1270.

Chassang, Sylvain, Gerard Padró i Miquel, and Erik Snowberg, "Selective Trials: A Principal-Agent Approach to Randomized Controlled Experiments," *American Economic Review*, 2012, *102*, 1279–1309.

Clasen, Thomas, Lawrence Haller, Damian Walker, Jamie Bartram, and Sandy Cairncross, "Cost-Effectiveness of Water Quality Interventions for Preventing Diarrhoeal Disease in Developing Countries," *Journal of Water and Health*, 2007, *5* (4), 599–608.

Cohen, Jessica and Pascaline Dupas, "Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment," *Quarterly Journal of Economics*, February 2010, *125* (1), 1–45.

_ , _ , and Simone Schaner, "Price Subsidies, Diagnostic Tests, and Targeting of Malaria Treatment: Evidence from a Randomized Controlled Trial," *American Economic Review*, 2015, *105* (2), 609–45.

Cole, Shawn A. and A. Nilesh Fernando, "The Value of Advice: Evidence from Mobile Phone-Based Agricultural Extension," Technical Report, Harvard Business School Finance Working Paper No. 13-047 2012.

— , **Daniel Stein, and Jeremy Tobacman**, "Dynamics of Demand for Index Insurance: Evidence from a Long-Run Field Experiment," *American Economic Review Papers and Proceedings*, May 2014, *104* (5), 284–290.

de Meza, David and Diane J. Reyniers, "Debiasing the Becker DeGroot Marschak valuation mechanism," *Economics Bulletin*, June 2013, *33* (2), 1446–1456.

Guiteras, Raymond and B.K. Jack, "Incentive, Selection and Productivity in Labor Markets: Evidence from rural Malawi," *NBER Working Paper 19825*, 2014.

Guiteras, Raymond P., David I. Levine, Stephen P. Luby, Thomas H. Polley, Kaniz Khatun e Jannat, and Leanne Unicomb, "Disgust, Shame and Soapy Water: Tests of Novel Interventions to Promote Safe Water and Hygiene," *Journal of the Association of Environmental and Resource Economists*, 2015. Forthcoming.

— , — , **Thomas H. Polley, and Brian Quistorff**, "Credit Constraints, Discounting and Investment in Health: Evidence from Micropayments for Clean Water in Dhaka," October 2015.

Heckman, James J. and Edward J. Vytlacil, "Policy-Relevant Treatment Effects," *The American Economic Review*, 2001, *91* (2), 107–111.

— **and** — , "Structural Equations, Treatment Effects, and Econometric Policy Evaluation," *Econometrica*, 2005, *73* (3), 669–738.

— **and** — , "Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast their Effects in New," in J. Heckman and E. Leamer, eds., *Handbook of Econometrics*, Vol. 6 of *Handbook of Econometrics*, Elsevier, September 2007, chapter 71.

Heckman, James, Sergio Urzua, and Edward J. Vytlacil, "Understanding Instrumental Variables in Models with Essential Heterogeneity," *The Review of Economics and Statistics*, 2006, *88* (3), 389–432.

Hoffmann, Vivian, "Intrahousehold Allocation of Free and Purchased Mosquito Nets," *American Economic Review*, 2009, *99* (2), 236–41.

— , **Christopher B. Barrett, and David R. Just**, "Do Free Goods Stick to Poor Households? Experimental Evidence on Insecticide Treated Bednets," *World Development*, 2009, *37* (3), 607–617.

Horowitz, John K., "The Becker-DeGroot-Marschak Mechanism is Not Necessarily Incentive Compatible, Even for Non-random Goods," *Economics Letters*, 2006, *93* (1), 6–11.

— , "What Do We Know About the Performance of the Becker-DeGroot-Marschak Mechanism?," 2006.

Imbens, Guido W. and Joshua D. Angrist, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, March 1994, *62* (2), 467–475.

Irwin, Julie R., Gary McClelland, Michael McKee, and William D. Schulze, "Payoff Dominance vs. Cognitive Transparency in Decision Making," *Economic Inquiry*, 1998, *36* (2), 272–285.

Jack, B.K., "Private Information and the Allocation of Land Use Subsidies in Malawi," *American Economic Journal: Applied Economics*, 2013, *5* (3), 113–135.

Jones, Michael P., "Indicator and Stratification Methods for Missing Explanatory Variables in Multiple Linear Regression," *Journal of the American Statistical Association*, 1996, *91* (433), 222–230.

Karlan, Dean and Jonathan Zinman, "Credit Elasticities in Less-Developed Economies: Implications for Microfinance," *The American Economic Review*, 2008, *98* (3), 1040–1068.

_ and _ , "Observing Unobservables: Identifying Information Asymmetries With a Consumer Credit Field Experiment," *Econometrica*, 2009, *77* (6), 1993–2008.

Karni, Edi and Zvi Safra, "Preference Reversal and the Observability of Preferences by Experimental Methods," *Econometrica*, 1987, *55* (3), 675–685.

Keller, L. Robin, Uzi Segal, and Tan Wang, "The Becker-DeGroot-Marschak Mechanism and Generalized Utility Theories: Theoretical Predictions and Empirical Observations," *Theory and Decision*, March 1993, *34* (2), 83–97.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz, "Experimental Analysis of Neighborhood Effects," *Econometrica*, 2007, *75 (1)*, 83–119.

Kremer, Michael and Alaka Holla, "Pricing and Access: Lessons from Randomized Evaluations in Education and Health," in "What Works in Development: Thinking Big and Thinking Small," Brookings Institution Press, 2009.

_ and Edward Miguel, "The Illusion of Sustainability," *Quarterly Journal of Economics*, August 2007, *122* (3), 1007–1065.

_ , _ , Sendhil Mullainathan, Clair Null, and Alix Zwane, "Making Water Safe: Price, Persuasion, Peers, Promoters, or Product Design?," 2009. Unpublished Manuscript.

Mazar, Nina, Botond Koszegi, and Dan Ariely, "True Context-dependent Preferences? The Causes of Market-Dependent Valuations," *Journal of Behavioral Decision Making*, 2014, *27* (3), 200–208.

Miller, Klaus M., Reto Hofstetter, Harley Krohmer, and Z. John Zhang, "How Should Consumers' Willingness to Pay Be Measured? An Empirical Comparison of State-of-the-Art Approaches," *Journal of Marketing Research*, 2011, *48* (1), 172–184.

Miller, Matthew Rhodes, "Hemispheric Ceramic Pot Filter Evaluation and Quality Assurance Program in Northern Ghana," Thesis, Massachusetts Institute of Technology 2012.

Nevo, Aviv, "Empirical Models of Consumer Behavior," *Annual Review of Economics*, 2011, *3* (1), 51–75.

Noussair, Charles, Stephane Robin, and Bernard Ruffieux, "Revealing Consumers' Willingness-to-Pay: A Comparison of the BDM Mechanism and the Vickrey Auction," *Journal of Economic Psychology*, 2004, *25* (6), 725–741.

Rubin, Donald B, "Multiple Imputation After 18+ Years," *Journal of the American Statistical Association*, 1996, *91* (434), 473–489.

Rutstram, E., "Home-grown Values and Incentive Compatible Auction Design," *International Journal of Game Theory*, 1998, *27* (3), 427–441.

Safra, Zvi, Uzi Segal, and Avia Spivak, "The Becker-DeGroot-Marschak Mechanism and Non-expected Utility: A Testable Approach," *Journal of Risk and Uncertainty*, June 1990, *3* (2), 177–190.

Shea, John, "Instrument Relevance in Multivariate Linear Models: A Simple Measure," *The Review of Economics and Statistics*, 1997, *79* (2), 348–352.

Shogren, Jason F., "Experimental Methods and Valuation," in Karl-Goran Maler and Jeffrey R. Vincent, eds., *Valuing Environmental Changes*, Vol. 2 of *Handbook of Environmental Economics*, Elsevier, 2006, chapter 19, pp. 969–1027.

Smith, Vernon L., "Microeconomic Systems as an Experimental Science," *The American Economic Review*, December 1982, *72* (5), 923–955.

Tarozzi, Alessandro, Aprajit Mahajan, Brian Blackburn, Dan Kopf, Lakshmi Krishnan, and Joanne Yoong, "Micro-loans, Insecticide-Treated Bednets, and Malaria: Evidence from a Randomized Controlled Trial in Orissa, India," *American Economic Review*, 2014, *104* (7), 1909–41.

Urbancic, Michael B., "Testing Distributional Dependence in the Becker-DeGroot-Marschak Mechanism," November 2011. Manuscript, U.C. Berkeley Department of Economics.

Wertenbroch, Klaus and Bernd Skiera, "Measuring Consumers' Willingness to Pay at the Point of Purchase," *Journal of Marketing Research*, May 2002, *39* (2), 228–241.

World Health Organization, "Water, Sanitation and Hygiene: Links to Health," http://www.who.int/water_sanitation_health/publications/facts2004/en/ 2004.

_ , *World Health Statistics 2011*, Geneva Switzerland: World Health Organization, 2011.

Table 1: Sample Composition

	Mean			Diff.	Regressions	
	Full Sample (1)	BDM (2)	TIOLI (3)	BDM-TIOLI (4)	BDM Draw (5)	TIOLI Price (6)
Number of respondents in compound (census)	3.593 [2.323]	3.305 [1.816]	3.859 [2.683]	−0.554* (0.323)	0.236*** (0.079)	−0.051 (0.045)
Husband lives in compound	0.794 [0.404]	0.792 [0.406]	0.796 [0.403]	−0.004 (0.022)	0.453 (0.367)	−0.243 (0.168)
Number of children age 0-5 in household	1.135 [0.978]	1.069 [0.941]	1.196 [1.008]	−0.127* (0.073)	0.195 (0.159)	0.028 (0.078)
Number of children age 6-17 in household	1.303 [1.282]	1.389 [1.304]	1.224 [1.258]	0.165** (0.084)	0.028 (0.129)	−0.013 (0.047)
Number of children age 0-5 with diarrhea in past two weeks	0.243 [0.525]	0.208 [0.487]	0.277 [0.557]	−0.069* (0.035)	−0.372 (0.376)	0.075 (0.128)
Number of children age 6-17 with diarrhea in past two weeks	0.049 [0.272]	0.050 [0.302]	0.048 [0.241]	0.002 (0.016)	−0.499 (0.417)	0.463* (0.267)
Respondent has ever attended school	0.090 [0.286]	0.079 [0.270]	0.100 [0.301]	−0.021 (0.016)	−0.025 (0.515)	−0.077 (0.195)
First principal component of durables ownership	0.132 [1.555]	0.059 [1.512]	0.198 [1.592]	−0.139 (0.126)	−0.046 (0.091)	0.005 (0.056)
All-year access to improved water source	0.187 [0.390]	0.196 [0.397]	0.179 [0.384]	0.017 (0.038)	−0.126 (0.376)	0.119 (0.252)
Currently treats water	0.115 [0.319]	0.109 [0.312]	0.120 [0.325]	−0.011 (0.024)	0.567 (0.468)	0.048 (0.257)
E. coli count, standardized	−0.052 [0.949]	−0.026 [1.012]	−0.076 [0.887]	0.050 (0.089)	−0.102 (0.162)	0.038 (0.120)
Turbidity, standardized	−0.065 [0.997]	−0.099 [0.922]	−0.032 [1.063]	−0.068 (0.096)	−0.008 (0.178)	0.224*** (0.081)
BDM Filter Bid (GHS)					−0.093 (0.062)	
Number of households	1265	607	658		607	658
Number of compounds	558	275	283		275	283

Notes: Columns 1, 2 and 3 display sample means in the full sample, BDM treatment and TIOLI treatment, respectively. Column 4 displays the differences in means between the BDM and TIOLI treatments. Column 5 displays the results of a regression of BDM draw on the listed characteristics. Column 6 displays the results of a regression of TIOLI price on the listed characteristics. Missing values of independent variables in Columns 5 and 6 are set to 0, and dummy variables are included to indicate missing values. Standard deviations in brackets. Estimated standard errors, clustered by compound, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Correlates of Willingness to Pay

	OLS	Probit	
	BDM-Estimated (WTP)	TIOLI (Purchase)	Diff.
	(1)	(2)	(3)
Number of respondents in compound (census)	0.053	-0.089^{***}	0.143^{**}
	(0.060)	(0.034)	(0.069)
Husband lives in compound	-0.005	-0.463^{*}	0.457
	(0.245)	(0.244)	(0.346)
Number of children age 0-5 in household	0.067	-0.066	0.133
	(0.113)	(0.092)	(0.145)
Number of children age 6-17 in household	0.018	0.197^{**}	-0.179^{*}
	(0.068)	(0.080)	(0.105)
Number of children age 0-5 with diarrhea in past two weeks	0.550^{*}	-0.260	0.809^{**}
	(0.286)	(0.170)	(0.333)
Number of children age 6-17 with diarrhea in past two weeks	-0.187	-0.663^{*}	0.475
	(0.220)	(0.355)	(0.417)
Respondent has ever attended school	0.604	-0.535^{**}	1.138^{**}
	(0.413)	(0.236)	(0.475)
First principal component of durables ownership	0.128^{*}	0.099	0.029
	(0.074)	(0.071)	(0.103)
All-year access to improved water source	-0.307	-0.259	-0.049
	(0.250)	(0.265)	(0.364)
Currently treats water	0.560	0.246	0.313
	(0.373)	(0.270)	(0.460)
E. coli count, standardized	-0.123	0.134	-0.258
	(0.109)	(0.161)	(0.194)
Turbidity, standardized	-0.190^{**}	0.076	-0.267^{*}
	(0.085)	(0.123)	(0.150)
P-value: All coeffs = 0	0.0104	0.0047	0.0001
R-squared	0.053		
Number of subjects	607	658	1265
Number of compounds	275	283	556

Notes: Column (1) displays coefficients from a linear regression of directly reported willingness to pay (the BDM bid) on baseline characteristics. Column (2) reports probit estimates of purchase decisions. As discussed in Section 4, by restricting the coefficient on price to equal -1 in the probit estimation, the estimated coefficients can be interpreted in terms of willingness to pay. Missing values of the independent variables are set to 0, and dummy variables are included to indicate missing values. Standard errors clustered at the compound (extended family) level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Casual Effect of Prices

	Filter Present and Undamaged (1)	Storage Vessel Contains Water (2)	Clay Pot Contains Water (3)	Usage Index (4)
A. Short-term effects				
Draw	0.017	0.037*	-0.003	0.043
	(0.018)	(0.022)	(0.024)	(0.038)
Mean Dependent Variable	0.877	0.753	0.728	-0.003
R-squared	0.017	0.014	0.008	0.012
Observations	235	235	235	235
B. One-year effects				
Draw	-0.005	0.025	0.018	0.026
	(0.035)	(0.032)	(0.032)	(0.051)
Mean Dependent Variable	0.641	0.486	0.380	0.066
R-squared	0.006	0.027	0.009	0.007
Observations	142	142	142	142

Notes: The sample includes those subjects in the BDM treatment who purchased the filter, i.e., drew a price less than or equal to their bid. Each column presents the results of a separate regression of the dependent variable, listed in the column heading, on BDM draw and a cubic function of BDM bid. See Section 5 for discussion of data. Usage index is the average of the normalized values of the three individual usage measures. Usage measures are observed by enumerator at indicated follow-up survey. Standard errors clustered at the compound (extended family) level in parentheses.

Table 4: Screening Effect of Prices

	Filter Present and Undamaged (1)	Storage Vessel Contains Water (2)	Clay Pot Contains Water (3)	Usage Index (4)
A. Short-term effects				
Bid	-0.010	-0.008	-0.009	-0.022
	(0.010)	(0.012)	(0.013)	(0.021)
Mean Dependent Variable	0.877	0.753	0.728	-0.003
R-squared	0.006	0.002	0.003	0.006
Observations	235	235	235	235
B. One-year effects				
Bid	0.013	0.027*	-0.013	0.018
	(0.014)	(0.014)	(0.012)	(0.021)
Mean Dependent Variable	0.641	0.486	0.380	0.066
R-squared	0.005	0.023	0.005	0.005
Observations	142	142	142	142

Notes: The sample includes those subjects in the BDM treatment who purchased the filter, i.e., drew a price less than or equal to their bid. Each column presents the results of a separate regression of the depend variable, listed in the column heading, on the willingness to pay, i.e the subject's bid in BDM. See Section 5 for discussion of data. Usage index is the average of the normalized values of the three individual usage measures. Usage measures are observed by enumerator at indicated follow-up survey. Standard errors clustered at the compound (extended family) level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

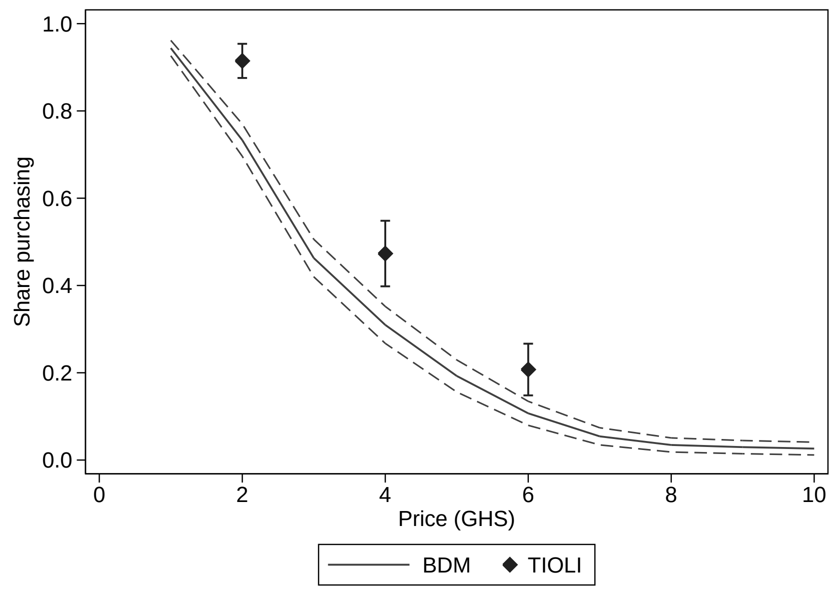
Table 5: Instrumental Variables Estimates of Treatment Effects:
Dependent variable: Child aged 0-5 has had diarrhea over previous two weeks

	TIOLI subjects		BDM subjects		Combined all subjects	
	(1)	(2)	(3)	(4)	(5)	(6)
A. Short-term effects						
Filter Purchase	-0.100*	-0.098*	-0.049	-0.058	-0.072**	-0.079**
	(0.054)	(0.051)	(0.050)	(0.043)	(0.035)	(0.034)
Mean Dependent Variable	0.149	0.149	0.142	0.142	0.145	0.145
R-squared	.	0.079	.	0.103	.	0.057
Observations	665	665	579	579	1,244	1,244
B. One-year effects						
Filter Purchase	0.148	0.220**	0.090	0.108	0.105	0.134**
	(0.099)	(0.100)	(0.089)	(0.090)	(0.067)	(0.068)
Mean Dependent Variable	0.215	0.215	0.262	0.262	0.241	0.241
R-squared	.	0.093	0.006	0.118	0.003	0.066
Observations	266	266	273	273	539	539
Controls	NO	YES	NO	YES	NO	YES
Village FEs	NO	YES	NO	YES	NO	YES

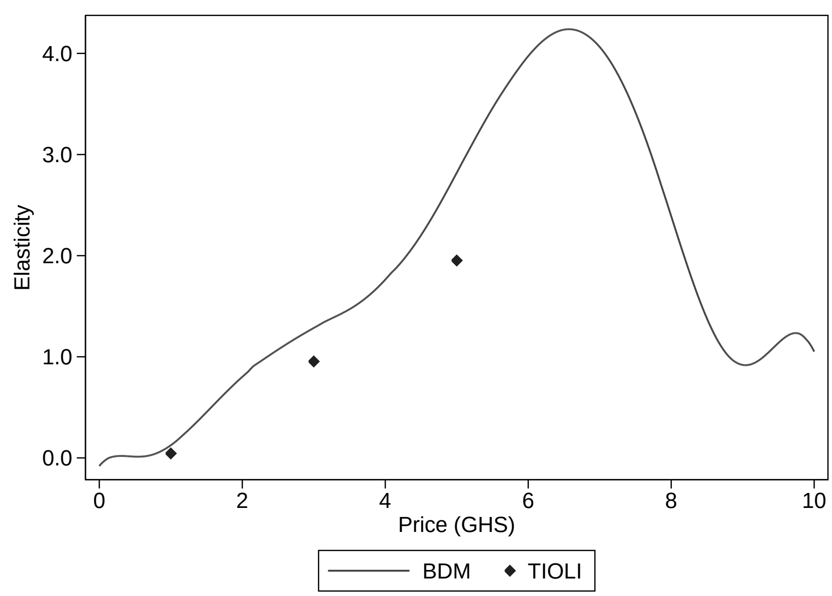
Notes: Each column displays the results of a 2sls regression of child diarrhea status at the child level on filter purchase, where filter purchase is instrumented for by random BDM draw for BDM subjects and by randomly assigned TIOLI price for TIOLI subjects. Controls include all variables (other than BDM draw) listed in Table 1. Missing values of independent variables are set to 0, and dummy variables are included to indicate missing values. Standard errors clustered at the compound (extended family) level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 1: Demand and Elasticity

(a) Inverse Demand Curve



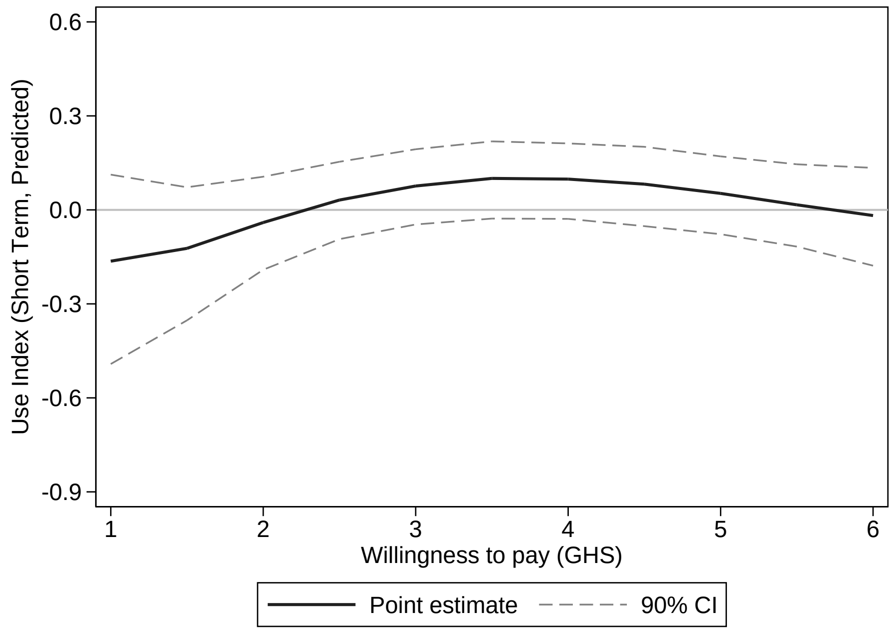
(b) Price Elasticity of Demand



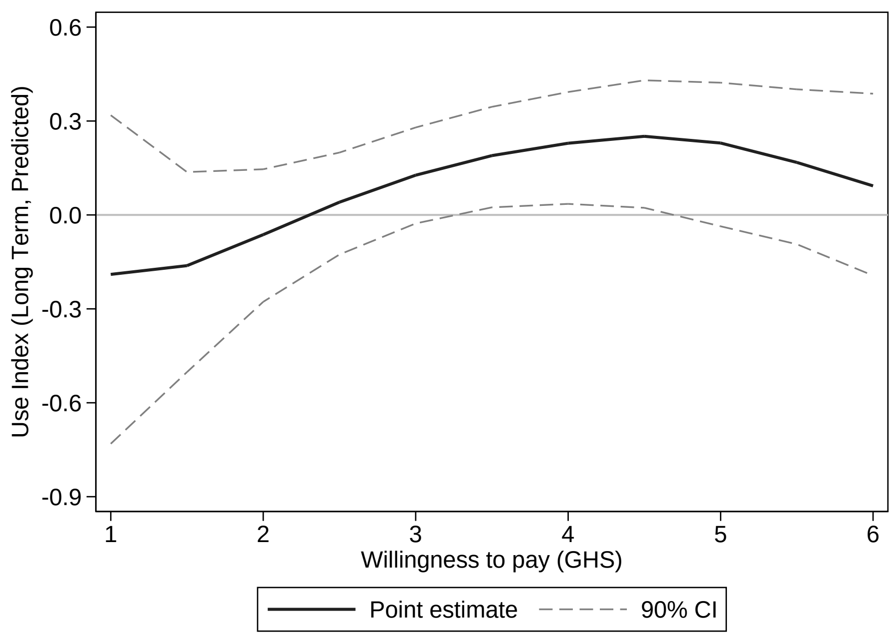
Notes: The top panel plots the BDM demand curve, with a 90% confidence band, and take-it-or-leave-it (TIOLI) demand at three price points (2, 4 and 6 GHS), with 90% confidence intervals. The BDM demand curve indicates the share of respondents with a BDM filter bid greater than or equal to the indicated price. The TIOLI markers indicate the share of respondents who purchased the filter at the corresponding (random) price. Point-wise inference from logit regressions (at prices GHS 1, 2, . . . , 10 for BDM, 2, 4, 6 for TIOLI). Standard errors clustered at the compound (extended family) level. 607 BDM observations. 658 TIOLI observations, of which 246 at a price of 2, 224 at a price of 4, and 188 at a price of 6. The bottom panel plots demand elasticities among BDM and TIOLI respondents. The BDM elasticity is calculated by a local polynomial regression, using an oversmoothed Epanechnikov kernel. The TIOLI elasticity is an arc elasticity calculated between GHS 0-2, 2-4 and 4-6 and plotted at the midpoint of each segment (GHS 1, 3 and 5, respectively). For both BDM and TIOLI, demand at a price of zero is is assumed to be 1.

Figure 2: Relationship between Use and Willingness to Pay
BDM Purchasers with Children 0 to 5

(a) One-Month Follow-Up



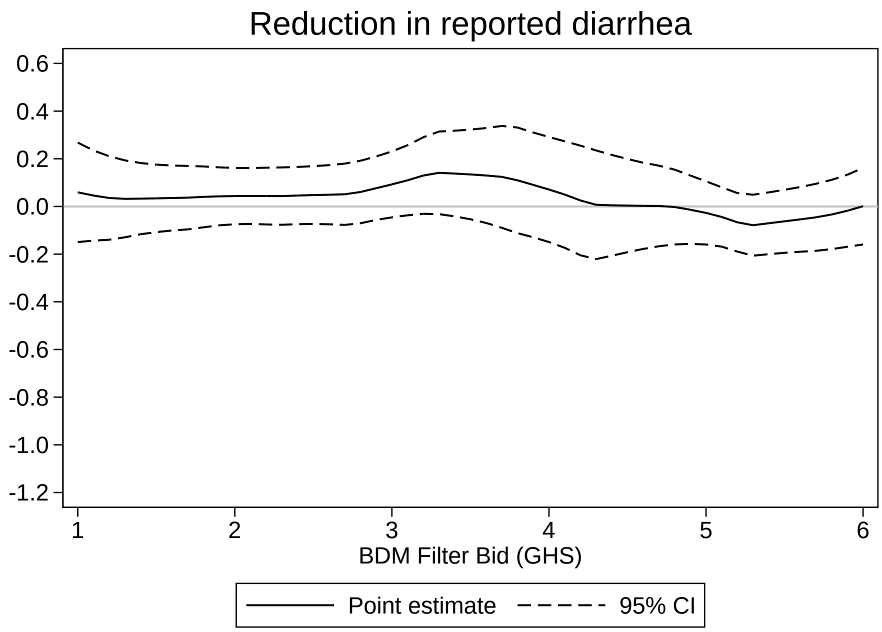
(b) One-Year Follow-Up



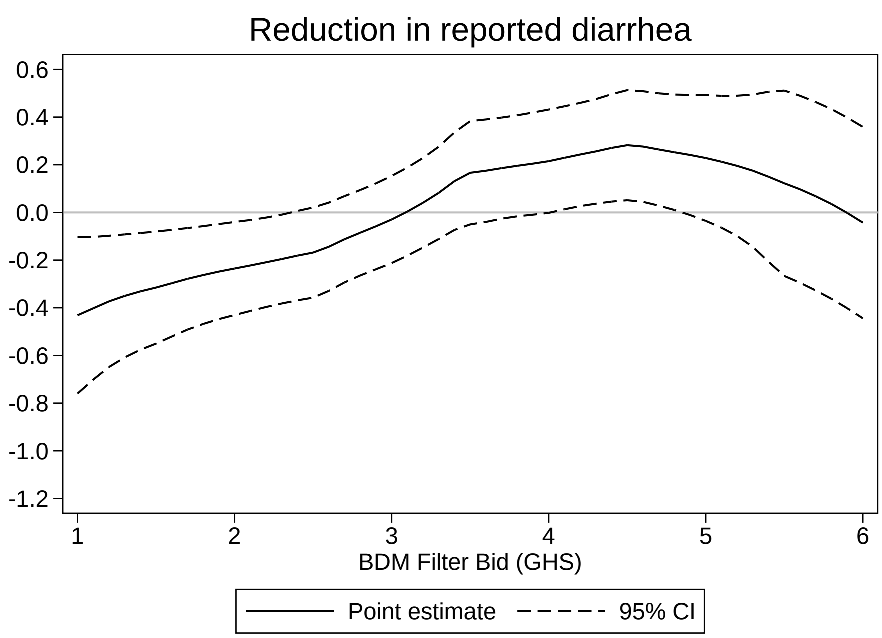
Notes: These figures show predicted values from a kernel regression (local polynomial of degree 1) of an index of use measures on the household's willingness-to-pay (WTP), as stated in the BDM sale. The index consists of indicators for whether the filter was observed in the compound, whether the safe storage container contained an appreciable amount of water (at or above the level of the spigot), and whether the ceramic pot contained water. These measures are standardized and averaged as in Kling, Liebman and Katz (2007). The sample consists of households that won a filter in the BDM sale and have one or more children age 0 to 5. Confidence intervals robust to clustering at the compound (extended family) level are computed by bootstrapping, resampling compounds with replacement (1,000 repetitions).

Figure 3: Kernel IV Estimates of Treatment Effects

(a) Short-term: One-Month Follow-Up



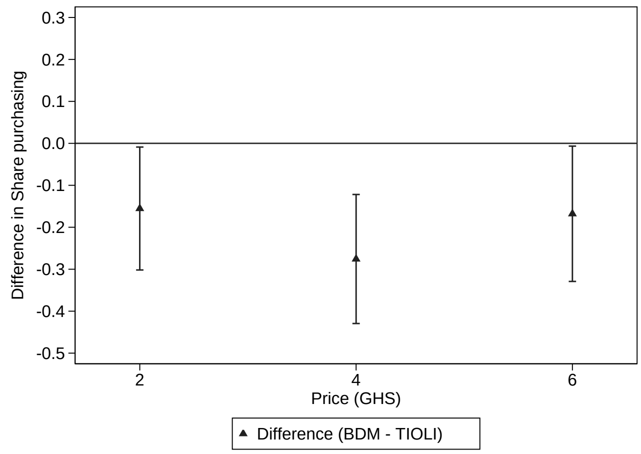
(b) Long-term: One-Year Follow-Up



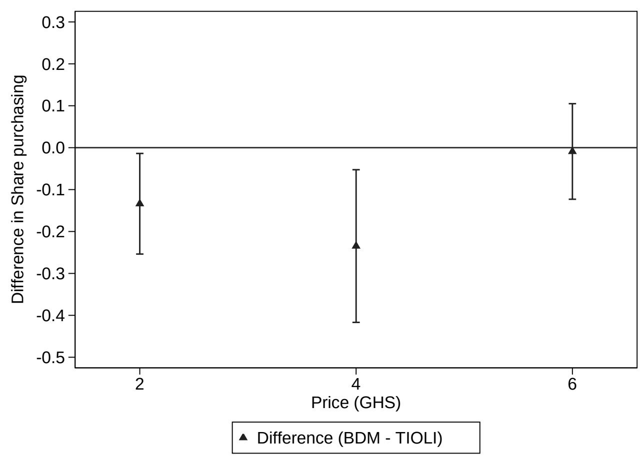
Notes: These graphs present estimated treatment effects (reduction in diarrhea among children age 0 to 5) as a function of willingness-to-pay (WTP). Estimates are by linear two-stage least squares at WTP = 1.0, 1.1, . . . , 6.0, weighting observations by their distance from the evaluation point using an Epanechnikov kernel. The endogenous treatment variable is an indicator for whether the household purchased a filter, and the exogenous instrument is the household's BDM draw. Standard errors are clustered at the compound (extended family) level. See Section 5.2 for details, and the Supplementary Materials for ancillary statistics (sample sizes and instrument strength).

Figure 4: BDM–TIOLI gap by tercile of risk aversion

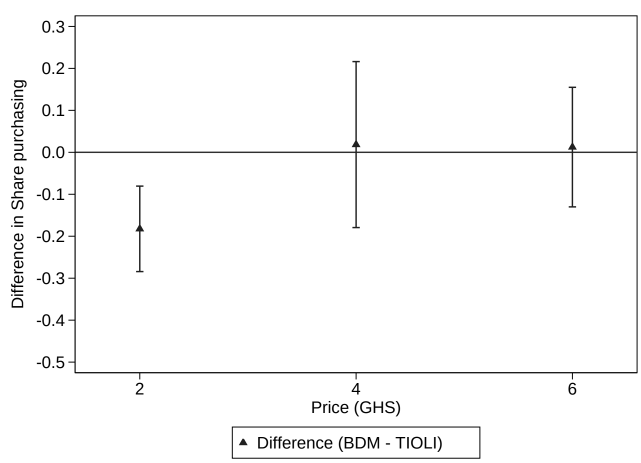
(a) Tercile 1 (most risk-averse)



(b) Tercile 2



(c) Tercile 3 (least risk-averse)



Notes: These figures plot the difference between the share of BDM subjects and the share of TIOLI subjects agreeing to purchase at each TIOLI price (GHS 2, 4, 6), separately by tercile of risk aversion. The results here are unconditional, see Figure A3 for robustness checks with additional controls.

A Appendix Figures

Figure A1: The *Kosim* filter

Figure A2: Experimental Timeline for a Typical Village

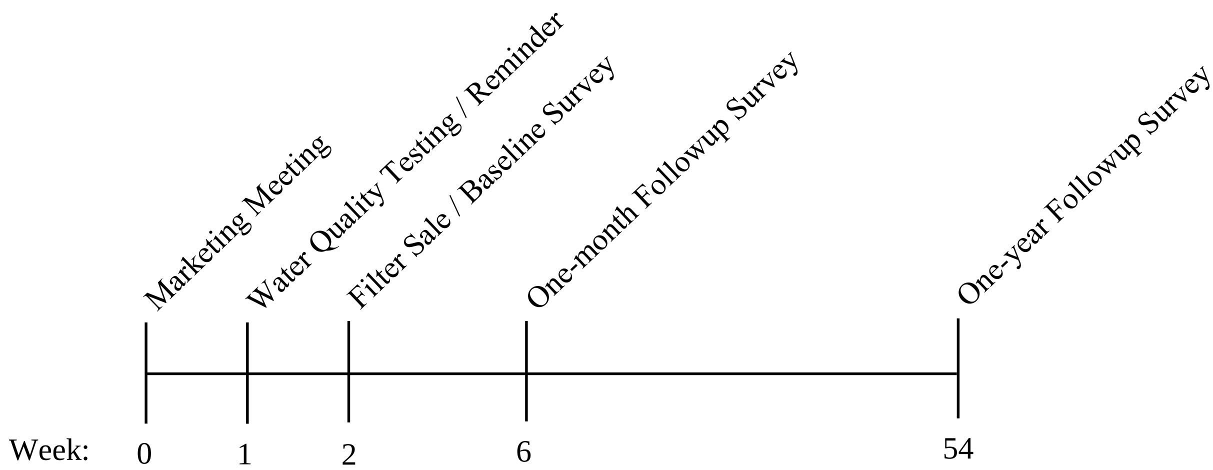
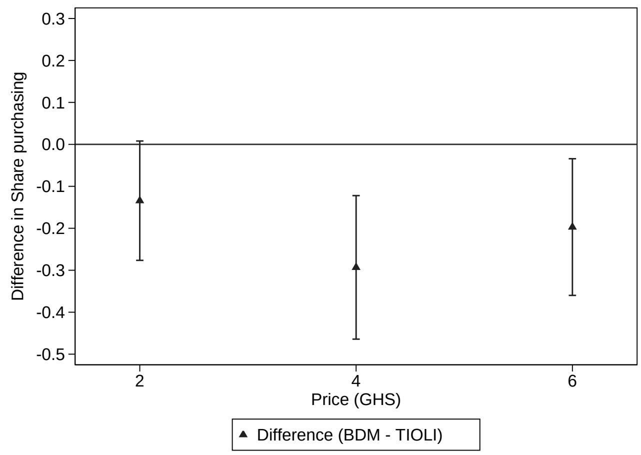
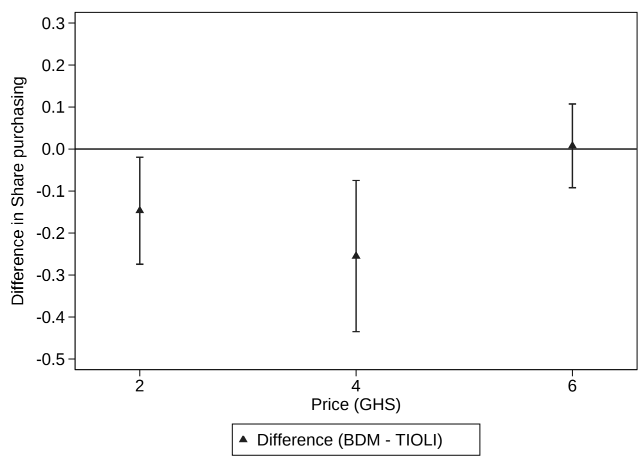


Figure A3: BDM–TIOLI gap by tercile of risk aversion
Robustness check: with controls, including ambiguity aversion

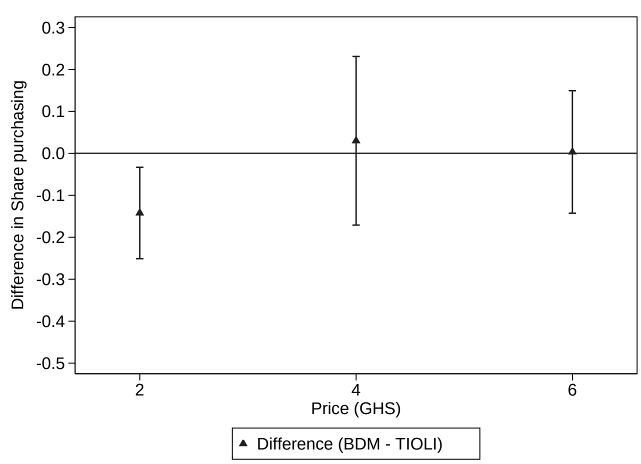
(a) Tercile 1 (most risk-averse)



(b) Tercile 2



(c) Tercile 3 (least risk-averse)



Notes: These figures plot the difference between the share of BDM subjects and the share of TIOLI subjects agreeing to purchase at each TIOLI price (GHS 2, 4, 6), separately by tercile of risk aversion. The regression includes the standard set of household controls and our measure of ambiguity aversion, described in Section 6.

B BDM Script

Section numbers refer to survey instrument. For full text of all sales treatments, see the Supplemental Materials.

J. REGULAR_BDM

READ EXACTLY FROM SCRIPT. DO NOT SAY ANYTHING THAT IS NOT IN SCRIPT.

READ:

- We would like to sell you a filter but the price is not yet fixed. It will be determined by chance in a game we are about to play.

- You will not have to spend any more for the filter than you really want to.

- You may even be able to buy it for less.

Here is how the promotion works:

- I will ask you to tell me the maximum price (***dan kuli***) you are willing to pay (***ka a ni sagi dali***) for the *Kosim* filter (koterigu di mali lokorigu).

- In this cup, I have many different balls with different numbers on them.

- The numbers represent prices for the filter.

- Then I will ask you to pick a ball from the cup, and we will look at the price together.

- If the number you pick is less than or equal to your bid, you will buy (***ani too dali***) the filter and you will pay the price you pick from the cup.

- If the number you pick is greater than your bid, then you cannot buy the filter.

- You will only have one chance to play for the filter.

- You cannot change your bid after you draw from the cup.

- You must state a price that you are actually able to pay now.

- We will practice in one moment, but for now, do you have any questions?

Answer any questions respondent has.

J.1 REGULAR_BDM PRACTICE

REMEMBER: Get respondent to state HIGHEST price they are WILLING AND ABLE to pay right now.

NOTE: Refer to p.23 for correct Dagbani translation of Cedi amounts.

- Before we play for the filter, let's practice the game. We'll play the same game, but instead of playing for the filter, we will play for this bar of soap. ***Show respondent soap.***

1) What is the maximum amount (dan kuli) that you are willing to pay for this soap? **[*Respondent states price X*]**

2) Now, if you pick a number that is less than or equal to X, you will buy the soap at the price you pick. If you pick a number greater than X, you will not be able to purchase the soap, even if you are willing to pay the greater number. You cannot change your bid after you pick a price. Do you understand?

3) Please, tell me - if you pick [X+5 peswas] now, what happens? ***If respondent does not give correct answer, explain the rules again and then ask question again.***

4) And if you pick [X-5 peswas] now, what happens? ***If respondent does not give correct answer, explain the rules again and then ask question again.***

5) If you draw [X+5], will you want to purchase the soap for [X+5]?

IF YES: → **5)**
IF NO: → **6)**

6) Do you want to change your bid to [X+5]?

IF YES: OK, your new bid is [X+5]. → **2) [*use* X+5 *as new* X]**
IF NO: → **6)**

7) So, is X truly the most you would want to pay?

IF YES: → **7)**
IF NO: → **1)**

8) If you pick X, you must be able to pay X. Are you able to pay X now?

IF YES: → **J.1.1**
IF NO: What is the maximum price you are willing **and able** to pay now? → **2) [*use new X*]**

→ ***Record respondent's Final Bid*** (J.1.1, page 29)

9) Could you please fetch the amount you have stated you are willing to pay and show it to me?
Wait for respondent to fetch money and check to see she has enough funds for Final Bid.

10) Now you will pick a price from the cup. If you pick X or less, you will buy the soap at the price you pick. If you pick more than X, you will not be able to buy the soap. Are you ready to pick a ball?

Mix balls in cup, hold cup above eye level of respondent and have her pick a ball without looking.

11) Now you can draw a ball from the cup. ***Let respondent draw ball. Together, look at the ball and read the price picked.* [*Drawn price is Y*]**

→ ***Record Drawn Price*** (J.1.2, page 29)

12) Let us look at the ball together.

→ ***Record if Drawn Price is lower/equal to or higher than Final Bid Survey*** (J.1.3, page 29)

a. **[*If Y <= X*]:** The price is Y which is [less than/equal to] the amount you said you would be willing and able to pay for this soap. You can now buy the item at this price.

→ ***Exchange payment for soap*.**

b. **[*If Y > X*]:** The price is Y, which is greater than the amount you said you would be willing to spend. You cannot purchase the soap.

13) Do you have any questions about the game?

Address any questions or concerns respondent has. Make sure she understands rules of game.

J.2 REGULAR_BDM FILTER SALE

REMEMBER: Get respondent to state HIGHEST price they are WILLING AND ABLE to pay right now.

NOTE: Refer to p.23 for correct Dagbani translation of Cedi amounts.

Read:

- Now you will play to buy the filter
- Recall the community meeting on [day of community meeting]
- Have you thought about how much you are willing to pay for the filter?
- Do you have the funds available now?

Let's begin:

1) What is the maximum amount (dan kuli) that you are willing to pay for this filter? **[*Respondent states price X*]**

2) Now, if you pick a number that is less than or equal to X, you will buy the soap at the price you pick. If you pick a number greater than X, you will not be able to purchase the soap, even if you are willing to pay the greater number. You cannot change your bid after you pick a price. Do you understand?

3) Please, tell me - if you pick [X+1 cedis] now, what happens? ***If respondent does not give correct answer, explain the rules again and then ask question again.***

4) And if you pick [X-1 cedis] now, what happens? ***If respondent does not give correct answer, explain the rules again and then ask question again.***

5) If you draw [X+1], will you want to purchase the filter for [X+1]?

IF YES: **→ 5)**
IF NO: **→ 6)**

6) Do you want to change your bid to [X+1]?

IF YES: OK, your new bid is [X+1]. **→ 2) [*use* X+1 *as new* X]**
IF NO: **→ 6)**

7) So, is X truly the most you would want to pay?

IF YES: **→ 7)**
IF NO: **→ 1)**

8) If you pick X, you must be able to pay X. Are you able to pay X now?

IF YES: **→ J.2.1**
IF NO: What is the maximum price you are willing **and able** to pay now?
→ 2) [*use new X*]

→ ***Record respondent's Final Bid*** (J.2.1, page 29)

9) Could you please fetch the amount you have stated you are willing to pay and show it to me?
Wait for respondent to fetch money and check to see she has enough funds for Final Bid.

10) Now you will pick a price from the cup. If you pick X or less, you will buy the filter at the price you pick. If you pick more than X, you will not be able to buy the filter. Are you ready to pick a ball?

Mix balls in cup, hold cup above eye level of respondent and have her pick a ball without looking.

11) Now you can draw a ball from the cup. ***Let respondent draw ball. Together, look at the ball and read the price picked.*** **[*Drawn price is Y*]**

→ ***Record Drawn Price*** (J.2.2, page 29)

12) Let us look at the ball together.

→ ***Record if Drawn Price is lower/equal to or higher than Final Bid*** (J.2.3, page 29)

a. **[*If Y <= X*]:** The price is Y which is [less than/equal to] the amount you said you would be willing and able to pay for this filter. You can now buy the filter at this price.

→***Receive payment for filter. Record filter tracking code on survey*** (I.2.5, page 29). ***Record filter tracking code on receipt and give it to respondent. Inform her of where and when she can pick up the filter.***

b. **[*If Y > X*]:** The price is Y, which is greater than the amount you said you would be willing to spend. You cannot purchase the filter.

→***Go to Household Survey question*** J.24, page 29