NC STATE UNIVERSITY

Center for Environmental and Resource Economic Policy
College of Agriculture and Life Sciences
https://cenrep.ncsu.edu

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

James Berry, Greg Fischer, Raymond Guiteras

Center for Environmental and Resource Economic Policy
Working Paper Series: No. 18-016
May 2018

Suggested citation: Berry, James, Greg Fischer and Raymond P. Guiteras, "Eliciting and Utilizing Willingness-to-Pay: Evidence from Field Trials in Northern Ghana," CEnREP Working Paper No. 18-016, May 2018, http://go.ncsu.edu/cenrep-wp-18-016.

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

James Berry, Greg Fischer and Raymond Guiteras*

May 2018

Abstract

Using the Becker-DeGroot-Marschak (BDM) mechanism, we estimate the willingness-to-pay (WTP) for and impact of clean water technology through a field experiment in Ghana. Although WTP is low relative to the cost, demand is relatively inelastic at low prices. In the short-run, treatment effects are positive—the incidence of children's diarrhea falls by one third—and consistent throughout the WTP distribution. After a year, usage has fallen, particularly for those with relatively low valuations. Strikingly, the long-run average treatment effect is negative for those with valuations below the median. Combining estimated treatment effects with individual willingness-to pay measures implies households' valuations of health benefits are much smaller than those typically used by policymakers. Finally, 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: University of Delaware, jimberry@udel.edu; Fischer: London School of Economics, g.fischer@lse.ac.uk; Guiteras: North Carolina State University, rpguiter@ncsu.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, Gharad Bryan, Sylvain Chassang, Vivian Hoffman, B. Kelsey Jack, Leigh Linden, Dean Karlan, Sheila Olmstead, Gerard Padró i Miquel, Giordano Palloni, Jesse Shapiro, Erik Snowberg, Chris Udry, Sergio Urzúa, Roger von Haefen, Ed Vytlacil 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. The Online Appendices referenced in the text are posted at `https://dataverse.harvard.edu/file.xhtml?fileId=3151066`. Scripts and survey forms for all sales treatments are posted at `https://dataverse.harvard.edu/file.xhtml?fileId=3151083`.

1 Introduction

Unsafe drinking water is a significant threat to health and welfare in the developing world. Approximately 30 percent of the world's population lacks access to safe water (WHO and UNICEF 2017), and diarrheal disease kills nearly 1.4 million people per year, including over 500,000 children under age five (WHO 2016). The problem is especially acute in sub-Saharan Africa, where diarrheal disease causes nearly 10 percent of deaths of children under age five, and 41 percent of the rural population drinks water from unimproved sources (WHO 2016; WHO and UNICEF 2017). Rural infrastructure improvements, such as bore wells or spring protection, have suffered from poor governance, frequent outages, and recontamination of water between collection and consumption (Wright et al. 2004; Miguel and Gugerty 2005; Kremer et al. 2011), leading to interest in household water treatment as a potentially attractive alternative. Simple, relatively inexpensive technologies, such as chlorination and filtration, are known to be microbiologically effective, and have reduced diarrhea in controlled field trials (Clasen et al. 2015).

Despite these potential benefits, demand for household water treatment is typically low (Ahuja et al. 2010). This is an example of a general puzzle in development economics—households appear to underinvest in seemingly beneficial technologies across many domains (Foster and Rosenzweig 2010; Dupas 2011; Jack 2011; Cole et al. 2013). When demand is low, measuring willingness-to-pay (WTP) provides a key input for pricing policy, guiding the magnitude and targeting of subsidies. Furthermore, understanding the relationship between WTP and a product's benefits is critical for distinguishing when the price mechanism allocates goods where their benefits are greatest and when it simply reduces access (Kremer and Holla 2009). In addition, combining measures of WTP with estimated treatment effects can yield insights into how households value health.

We study the demand for and impact of a household water filter in a field experiment with 1,265 households in rural northern Ghana. The filter requires effort to use, but if used properly produces safe drinking water for the household. After normal marketing efforts, we made sales offers to households and distributed the filters to those who purchased it. We conducted follow-up surveys one month and one year after the sale to measure filter use and health outcomes related to water quality.

In our study we used the Becker-DeGroot-Marschak mechanism (BDM, Becker et al. 1964) to elicit precise measures of WTP. In BDM, an individual states her bid for an item. Then a random price is drawn. If the random price is greater than her bid, she cannot purchase the product. If the random price is less than or equal to her bid, she purchases the product, but pays the random price draw rather than her stated bid. Because the subject's stated WTP affects only whether or not she purchases the item, not the price she pays, BDM is *incentive-compatible*: the subject's dominant strategy is to bid her true maximum WTP.[1] In contrast to take-it-or-leave-it (TIOLI) offers, which yield only a bound on WTP, BDM produces an exact measure. In addition, BDM induces random variation in both treatment status and price paid, conditional on WTP. This allows researchers to separately identify screening and sunk cost effects.[2] Embedded in a field experiment, BDM can thus extract richer information than is typically available, but with the potential cost of added complexity. To assess the performance of BDM in a field setting, we randomly allocated half the households to a BDM sales treatment and half to a more traditional sales treatment using a TIOLI offer at a random price.

This study makes five main contributions. First, we measure demand for clean water technology in a highly relevant population facing a stark decision: how much of their scarce resources should they allocate to a product that improves poor water quality? Although demand estimates can provide important information on welfare and how to prioritize policy, measuring demand in developing countries is difficult because revealed-

[1] Deviations from expected-utility maximization may lead a subject's optimal bid to deviate from her true maximum WTP (Horowitz 2006a), which we discuss in Section 6 below.

[2] Screening effects and sunk-cost effects typically cannot be separately identified, either in observational data or through TIOLI offers. In this paper, we focus on screening effects because we find no evidence of sunk-cost effects. We provide a detailed discussion and analysis of sunk-cost effects in Appendix E.

preference tools such as hedonic valuation or compensating differentials rely on strong assumptions of complete markets (Greenstone and Jack 2015). This paper adds to a small but growing literature on measuring demand for health goods directly through sales to households.[3] Similar to previous research for other preventative health products, we find that demand is low. Median WTP is only 10 to 15 percent of the manufacturing cost, and demand is close to zero at a break-even price. However, we find that almost all households have positive WTP, and demand is relatively inelastic at low prices.

Second, we estimate the causal effect of receiving the filter on child health by using exogenous variation in filter allocation provided by our sales exercise. In the short run, we find that the filter reduces the probability that a child aged five or under has a reported case of diarrhea in the previous two weeks by about 7 percentage points, which is a substantial reduction relative to the baseline rate of 21 percent. However, these benefits do not persist. In fact, we estimate the average treatment effect of the filter at our one-year follow-up visit to be negative: diarrhea *increased*.

Third, we shed light on this surprising finding by estimating the distribution of treatment effects with respect to WTP. The importance of understanding and estimating marginal treatment effects (MTEs) has been emphasized by Heckman and Vytlacil (1999), Heckman and Vytlacil (2007) and Chassang et al. (2012), both for policy analysis and for uncovering structural economic parameters. Estimating MTEs typically requires strong structural assumptions or multiple or multi-valued instruments. In contrast, by jointly eliciting WTP and generating exogenous variation in treatment conditional on WTP, BDM allows us to estimate the distribution of MTEs with respect to WTP in a simple and transparent way. We find that after one year, the benefit of the filter is increasing in WTP, and the negative effect occurs in households with below-median WTP. The pattern of filter use resembles the pattern of treatment effects: households with low WTP were less likely to be using

[3] See, for example, Ashraf et al. (2010), Cohen and Dupas (2010) and Guiteras et al. (2015). Ito and Zhang (2016) provide an alternative approach using observational data, carefully isolating the price premium for goods with varying environmental benefits.

the filter after one year, suggesting that household behavior, in particular effort with respect to proper maintenance and use of the filter, is an important mediator of benefits. These findings have two important policy implications. First, in this sample, charging a positive price would allocate the filter to households where it is beneficial. Second, it underscores the importance of household behavior. Even technologically sound health products may not achieve their potential without appropriate household inputs (Brown and Clasen 2012; Hanna et al. 2016).

Fourth, we combine our data on demand with the estimated health impact of the filter to calculate demand for health. This contributes to the limited set of revealed-preference estimates for the value of health in low-income countries (Greenstone and Jack 2015). Because we have precise revealed-preference WTP data as well as WTP-specific impacts, we can estimate the distribution of demand for health. Using our short-run estimates, we find that median WTP to avert one episode of children's diarrhea is USD 1.12. With additional assumptions, this implies a median WTP of USD 3,604 to avoid one statistical child death or USD 40 to avoid the loss of one disability-adjusted life year. Consistent with Kremer et al.'s (2011) calculations based on the health effects of spring protection in Kenya, our estimates are well below standard cost-effectiveness thresholds.

Fifth, by randomizing households to either BDM or TIOLI sales treatments we can compare the two WTP-elicitation mechanisms. Although BDM has the potential to enhance the information gained from field experiments, little is known about its performance in the field. BDM has been extensively used in laboratory settings in experimental economics, but anomalous behavior among subjects has been observed, such as sensitivity to the distribution of draws (Bohm et al. 1997; Mazar et al. 2014) or misunderstanding of the dominant strategy (Cason and Plott 2014). It is therefore an open question whether BDM's potential advantages outweigh its potential drawbacks. We present what is, to our knowledge, the first direct comparison of BDM and the more common TIOLI in a developing-country field setting.[4] Encouragingly, results from both methods of demand elicitation follow a similar pattern and imply similar price elasticities. Furthermore, the cross-validated, predictive power of BDM estimates for TIOLI behavior is comparable to that of TIOLI itself. However, TIOLI acceptance rates are above the BDM demand curve. We explore a number of potential explanations and find that risk aversion accounts for much of the difference between the two methods.

The paper proceeds as follows. Section 2 describes the experimental setting and the data. Section 3 describes the pattern of demand for the filter. Section 4 presents the health impacts of the filter and heterogeneous treatment effects by WTP. Section 5 discusses policy counterfactuals and the WTP for children's health. Section 6 compares the BDM and TIOLI mechanisms and discusses implications for future research using BDM. The final section concludes.

2 Experimental Setting and Design

We study the *Kosim* water filter (Figure A1), marketed in northern Ghana by Pure Home Water, an NGO. The *Kosim* product consists of a clay pot treated with colloidal silver and a plastic storage container with a tap. The filter has been shown to be highly microbiologically effective in field trials, removing more than 99 percent of pathogens (Johnson 2007). This effectiveness is sustained with proper use. Field tests one to three years after purchase have found that well-maintained filters remove more than 95 percent of pathogens (Clopeck 2009). At the time of the study, the cost of production and delivery to a rural household in a village-level distribution was about GHS 21 (USD 15). Demand for the filter is close to zero at this price, so the relationships between price and access, use, and impacts 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 be-

[4] Section 6 summarizes the theoretical and experimental literature studying behavior under BDM.

tween October 2009 and June 2010. To select our sample, we identified villages that had limited access to clean drinking water and had not previously been exposed to the *Kosim* filter. Within these villages, we conducted our baseline survey and sales exercise with women who were primary caregivers of children.[5] Figure A2 provides an illustrative timeline for a typical village.

2.1 Data Collection and Experimental Design

2.1.1 Preliminary Activities & Household Survey

MARKETING MEETING. In each study village, we held an initial village meeting. The NGO conducted its usual demonstration and marketing of the filter, and our field staff demonstrated the sales mechanisms. During these demonstrations, field staff performed mock versions of BDM and TIOLI for a token item, such as chocolate or soap. The staff also practiced the sales mechanisms with volunteer attendees, again for a token item. We informed villagers that a filter would be installed at the village health worker's home and encouraged them see it in use, taste the water, and ask questions. We announced that we would visit households in two weeks to offer the filter via one of the two sales mechanisms and encouraged them to discuss with their families what they were willing to pay for the filter. The two-week interim period was 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 village census to identify study subjects and randomize the sales treatments.

WATER QUALITY TESTING. One week after the marketing meeting, we visited each household to remind them of the upcoming sale and to answer questions. In all households, we collected a 100 ml sample of drinking water. Budget constraints prevented

[5]These were primarily mothers, but occasionally were others caring for children whose parents had migrated 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 main group of interest to the NGO.

testing all samples, so we tested levels of *E. coli* and turbidity in a randomly-selected half of the samples.

HOUSEHOLD SURVEY. One week after the reminder visit, we conducted a survey and sales visit with each subject. Subjects were compensated with a GHS 1 cash gift at the beginning of the survey. This was given in small coins so that respondents could submit fine-scale bids in the practice rounds 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. The survey collected demographic information, asset ownership, information on water collection and treatment practices, basic health knowledge, and recent episodes of diarrhea among household members.

2.1.2 Filter Sale

At the end of the survey, we conducted the sales experiment. Respondents were randomly assigned in roughly equal proportions to either a BDM or TIOLI sales treatment.[6] Treatments were randomized at the compound level, stratified by number of respondents in the compound.[7] Each sale began with a practice round in which we offered the respondent the opportunity to purchase a bar of soap with retail value of about GHS 1 using her assigned sales mechanism. After the practice round, we offered the respondent the *Kosim* filter using the same mechanism. If the mechanism resulted in a sale, the subject paid for the filter and received a receipt that could be redeemed for the filter at a central location in the village, typically the health liaison's home. To maintain realism—households routinely make small loans to each other for purchases—we permitted households to gather

[6]Within each of these two broad categories, we included three sub-treatments, described in Appendix J, to examine mechanisms underlying potential differences between BDM and TIOLI responses. However, demand was statistically indistinguishable by sub-treatment, and we group sub-treatments together for the primary analysis.

[7]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 mother is responsible for providing water for her husband and children. All inference is robust to clustering at the compound level.

the money by the end of the day. If the respondent initially agreed to the purchase but was ultimately unable to obtain the funds, we code her as not purchasing. Scripts for both mechanisms are provided in Appendix A.

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 to bid her maximum WTP for the soap. The surveyor then asked the respondent if she would want to purchase the soap if she drew slightly more than her bid. The respondent was then allowed to adjust her bid. This process repeated until she was no longer willing to adjust her bid. Next, the surveyor reminded her that if she drew a price equal to her bid she must be willing and able to make this payment. At several points during the process, the surveyor reviewed various hypothetical outcomes to test the respondent's understanding. These confirmation steps differ from the normal BDM procedures used in labs; however, we found during piloting that they greatly increased subjects' understanding and comfort with the mechanism. Once the final bid was established, the price was drawn and the subject either purchased or did not purchase the soap. The procedure for the filter was similar.[8]

Consistent with the general sales process described above, we did not require respondents to present the amount of cash they were willing to bid before the draw was made. However, before the draw, we asked multiple times whether the respondent would have access to the necessary funds. 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 unexpected inability to gather funds, for example because a family member was unavailable.

[8]Prices were written on wooden beads and placed in an opaque cup. The subject drew the price herself. For soap, the prices were distributed uniformly from 0 to 100 in increments of 10 pesewas (GHS 0.10). For the filter, 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. In neither case did we inform respondents of the distribution.

TIOLI TREATMENT. The standard TIOLI treatment was a simple sales offer at a randomized price. We emphasized that there would be no bargaining. We first conducted a practice round for a bar of soap. We then presented the offer for the filter randomized across three prices: GHS 2, 4, and 6, which were the approximate 25th, 50th and 75th percentiles of BDM bids in piloting.

2.1.3 Follow-up Surveys

We conducted follow-up surveys one month and one year after the sale.[9] Both surveys obtained caretaker reports on diarrhea over the previous two weeks among children aged five and under. Among households that purchased the filter, surveyors recorded objective indicators of the filter's condition and use. In the one-year survey, we also measured risk aversion, ambiguity aversion, digit span, and other preferences and beliefs that we hypothesized could be related to perceptions of the two sales mechanisms. Appendix B provides additional detail.

The one-month survey was conducted in all 15 study villages. Due to funding constraints, we randomly selected eight villages for the one-year survey. We re-surveyed 87.1 percent of targeted households in the one-month follow-up and 90.5 percent in the one-year follow-up. Attrition is largely balanced along observable dimensions, although there is some imbalance in observables across attritters and stayers in the one-year follow-up. Most relevant for our treatment effects estimation, attrition is not related to the BDM draw or to the TIOLI price. See Appendix C for detail.

2.2 Sample Characteristics and Balance

Table 1 displays summary statistics. Column 1 displays means of baseline characteristics for the full sample. Only 9 percent of respondents had ever attended school, and the

[9] Given good maintenance practices, in particular regular cleaning of the filter element, the filter's useful life is expected to be two years. We chose the one-year horizon as half this expected life. In practice, 40-50 percent of filters were found to be undamaged and in use after one year.

average number of children aged 0 to 5 was 1.1 per respondent. Diarrhea incidence was relatively high: on average, households had 0.24 episodes of diarrhea among children aged 0 to 5 in the previous two weeks. Only 19 percent of households had access to an improved water source year round.[10]

Columns 2 and 3 display sample means by treatment (BDM or TIOLI), and Column 4 tests differences between the two. There are a few marginally significant differences: 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 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 0.1 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 Demand for Filters

This section describes the demand for water filters measured through sales to households. Here, our focus is on the pattern of demand estimated through either the BDM or TIOLI mechanisms. Section 6 compares the two mechanisms in detail.

Figure 1a shows the inverse demand curve generated across all 15 villages using data from all 608 BDM and 658 TIOLI subjects. For the BDM observations, we plot for each price p the share of subjects whose bid was greater than or equal to p. For the TIOLI

[10]Due to budget constraints, water quality (*E. coli* and turbidity) 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).

subjects, we show the share who purchased at each of the three randomly-assigned price points, $P = 2, 4, 6$.

There are several features of this inverse demand curve worth noting. First, WTP is almost universally positive: across the full sample, 95 percent of respondents were willing to pay at least GHS 1.[11] At the same time, WTP is low relative to the cost of the filter: the median BDM 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 low WTP 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 WTP data, and, for TIOLI subjects, the arc price elasticity of demand from 0 to 2, from 2 to 4, and from 4 to 6. In both groups, demand at low prices is relatively inelastic. In fact, demand is price inelastic up to roughly the median of the WTP distribution. While the lack of a steep drop in demand above a price of zero is largely consistent with existing estimates of demand for health products, the demand curve we estimate exhibits less price sensitivity than what has been found in much of the prior literature (Dupas and Miguel 2017).

4 Health Impacts and Heterogeneous Treatment Effects

This section presents estimates of the filter's impact on children's diarrhea. In Section 4.1, we present standard IV estimates of treatment effects using the random offer price as an instrument for TIOLI subjects and the random price draw as an instrument for BDM subjects. In Section 4.2, we discuss the importance of estimating heterogeneous treatment effects (HTEs) and how we can use BDM to estimate HTEs, in particular the relationship

[11] "House money" effects could provide one plausible explanation for high demand at small positive prices; 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 (computed as 1 minus the BDM draw for soap among those who purchased soap) and the BDM filter bid, conditional on WTP for soap.

between treatment effects and WTP. In Section 4.3, we apply this method and uncover important heterogeneity: benefits and WTP are positively related in our one-year follow-up data. Section 4.4 shows that a similar positive relationship exists between use and WTP, and Section 4.5 further investigates the mechanisms behind the observed patterns of treatment effects.

4.1 Average Effects on Child Health

We begin with the basic treatment effects equation

$$y_{jic} = \beta_0 + \beta_1 T_{ic} + \varepsilon_{jic}, \tag{1}$$

where y_{jic} indicates 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 dummy variable indicating whether subject i purchased the filter, and ε_{jic} captures unobservable determinants of y. The coefficient of interest is β_1, the effect of purchasing a filter on children's diarrhea.

To instrument for the endogenous treatment variable, we estimate the first-stage equation

$$T_{ic} = \gamma_0 + \gamma_1 P_{ic} + v_{ic}, \tag{2}$$

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 ε_{jic} and therefore is a valid instrument for treatment. Table A1 presents the linear probability model estimates of the first-stage equation. In all samples and specifications, price strongly predicts treatment, with a 1 GHS reduction in price leading to a 9.3 to 18.4 percentage point increase in the probability of treatment, and an F-statistic on the excluded instrument between 72.3 and 665.6.

Panel A of Table 2 presents linear two-stage least squares estimates from our short-term (one-month) data for the pooled, TIOLI, and BDM samples. Using the pooled data, the likelihood of diarrhea in the two weeks before the survey is reduced by about 7 percentage points, comparable to estimates from other trials (Ahuja et al. 2010). The estimates for TIOLI and BDM subjects are similar. The TIOLI point estimates are slightly higher, but not statistically different.

In Panel B of Table 2 we examine our long-term data, collected in a random sub-sample of half our villages. After one year, there is no evidence of health benefits. In fact, the point estimates are positive: the filter appears to have increased the likelihood of children's diarrhea. The effect is only statistically significant with controls, but the point estimates are consistently positive and economically meaningful across specifications and samples. We explore this finding in the following subsections.

4.2 Heterogeneous Treatment Effects: Theory

The standard IV approach of the previous subsection estimates a single average treatment effect. As discussed by Heckman and Urzúa (2010), this may not be the parameter of interest. In our setting, understanding the relationship between benefits and WTP is critical for pricing policy. 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. 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 restrict access without improved targeting (Cohen and Dupas 2010). Standard IV methods cannot address these questions. However, because BDM both elicits respondents' WTP and allocates treatment randomly conditional on WTP, it provides a simple way to estimate the relationship between benefits and WTP.

Consider the following econometric model, adapted from Heckman et al. (2006), which generalizes (1) to allow β_1 to vary by WTP :[12]

$$y = \beta_0 + \beta_1 (w) T + \varepsilon, \tag{3}$$

where $\beta_1 (w)$ is the marginal treatment effect for those with WTP $= w$, and WTP has distribution $F_{WTP}(w)$. Let $\overline{\beta}_1 = E_{F_{WTP}} [\beta_1 (w)]$ be the average treatment effect in the population, and let $\tilde{\beta}_1 (w) = \beta_1 (w) - \overline{\beta}_1$ be the difference between $\beta_1 (w)$ and this average.

Now, consider the usual case where WTP is unobserved. The estimable model is

$$y = \beta_0 + \overline{\beta}_1 T + u, \tag{4}$$

with compound error term $u = \tilde{\beta}_1 (w) T + \varepsilon$. OLS estimation of (4) be biased if

$$E [Tu] = E \left[T \left(\tilde{\beta}_1 (w) T + \varepsilon \right) \right] \neq 0. \tag{5}$$

There are two potential sources of bias. The first is selection on levels, $E [T\varepsilon] \neq 0$, in which treatment status is correlated with unobservable determinants of y in the absence of treatment. The second is selection on gains: if WTP and benefits of treatment are correlated, then $E \left[T\tilde{\beta}_1 (w) \right] \neq 0$.

Selection on levels is traditionally addressed by an instrument: a source of variation in treatment that is uncorrelated with unobservables. One natural candidate is a randomized price. Let $Z \in \{P_L, P_H\}$ be a randomized price, which for simplicity takes on two values, $P_L < P_H$. If demand is downward-sloping, then $\Pr (T|P_L) > \Pr (T|P_H)$, so the instrument is relevant: Z is correlated with T. The instrument is valid if

$$E [Zu] = E [Z\varepsilon] + E \left[Z\tilde{\beta}_1 (w) T \right] = 0. \tag{6}$$

Since Z is random, $E [Z\varepsilon] = 0$, which solves the problem of selection on levels. How-

[12] We provide a more complete treatment in Appendix D.

ever, the problem of selection on gains remains. Since $T = 1\{\text{WTP} > Z\}$, if there is a relationship between WTP and gains then $E\left[Z\tilde{\beta}_1(w)\,T\right] \neq 0$.

The discussion above shows that in the presence of selection on gains, IV using the offer price Z will not produce a consistent estimate of $\bar{\beta}_1$. By the LATE theorem of Imbens and Angrist (1994), IV does provide a consistent estimate of the average effect on the compliers: those whose treatment status is changed by the instrument. In this case, this is the group with $P_L \leq \text{WTP} \leq P_H$, the population that would buy a filter at P_L but not at P_H. Formally, instrumental variables using Z as the instrument estimates

$$\beta_1^{IV}\left(P_L \leq \text{WTP} \leq P_H\right) = \int_{P_L}^{P_H} \beta_1(w)\, dF_{WTP}(w),$$

the average $\beta_1(w)$ between P_L and P_H weighted by $F_{WTP}(\cdot)$. As argued in Heckman and Vytlacil (2005) and Heckman and Vytlacil (2007), this may not be a useful parameter, since it only tells us the effect of changing price from P_H to P_L in a population with WTP distributed $F_{WTP}(\cdot)$.

BDM provides a simple method to estimate $\beta_1(w)$.[13] Intuitively, BDM reveals the respondent's WTP, and then the BDM draw randomizes treatment.[14] With a large enough sample, we could estimate the function $\beta_1(w)$ nonparametrically by comparing outcomes of winners and losers at each WTP. In practice, our sample is not large enough to condition

[13]Heckman et al. (2006) show that the distribution of $\beta_1(w)$ can be estimated even though WTP is typically not observed directly. Their *local instrumental variables* (LIV) method estimates a propensity score model in a first step and then regresses the outcome of interest on the propensity score. Our BDM approach has the advantage of observing WTP directly, rather than inferring it through a first-step selection equation. This increases power, as we show in Appendix D.2. We focus on price as an instrument for comparability with our application. However, the LIV method of Heckman et al. (2006) also applies to non-price instruments. Note that continuous, many-valued, or multiple instruments will be required to estimate the propensity score flexibly. Furthermore, interpretation is more subtle with non-price instruments since the heterogeneity in treatment effects is estimated with respect to unobservables. See further discussion in Appendix D.1.

[14]Using the BDM draw as an instrument requires the exclusion restriction that the draw does not directly affect the outcome of interest. This would be violated if there are wealth effects, since the draw determines the price paid. This is a common problem in IV estimation and applies to the randomized TIOLI price as well (Jones 2015). Similarly, a causal effect of price paid on use may also violate the exclusion restriction. However, as shown in Appendix E, we do not observe a causal effect of price paid on use.

on exact WTP. Instead, we compute kernel-weighted linear 2SLS estimates on a WTP grid. For $w \in \{\underline{w}, \ldots, \bar{w}\}$, we estimate

$$
\begin{aligned}
y_i &= \beta_1 (w) T_i + \varepsilon_i \\
T_i &= \psi Z_i + \nu_i,
\end{aligned}
$$

assigning higher weight to observations with WTP closer to w.

4.3 Heterogeneous Treatment Effects: Application

The kernel IV approach reveals substantial heterogeneity with respect to WTP. 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 kernel-weighted treatment effects $\hat{\beta}_1 (w)$ 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.[15] See Figures A4 and A5 for the first stage and tests of instrument strength.

We present the results in Figure 2, where we reverse the sign of $\hat{\beta} (w)$ so benefits are represented as positive. In the top panel (Fig. 2a), we consider the effect at one month. The point estimates are positive, consistent with Table 2, although not statistically significant at any level of WTP, and there is little heterogeneity in benefits. In the bottom panel (Fig. 2b), we repeat this analysis using the one-year data. Figure 2b shows important heterogeneity: the perverse negative effect of the filter occurs in the population with below-median WTP. The estimated benefit increases with WTP, becoming positive at roughly GHS 3 and peaking at roughly GHS 4.5. Above GHS 4.5, point estimates decrease, although confidence intervals are wide and the decrease is not statistically significant.

While the sample size and the flexibility of the kernel IV both limit the precision of our

[15] Non-parametric estimators are prone to bias at boundaries. Restricting to the 0.1 and 0.9 quantiles of WTP reduces this risk. Furthermore, our estimator is analogous to a local linear regression rather than a local constant regression, and local linear regressions are less subject to boundary bias (Li and Racine 2007, Chap. 2.4).

estimates, we can reject that the one-year treatment effects for those with WTP $= 4$ and those with WTP $= 2$ are equal (estimated difference $\hat{\beta}(4) - \hat{\beta}(2) = 0.450$, standard error of estimated difference 0.141, $p = 0.001$). If we impose that the heterogeneous treatment effect be linear, the interaction or slope term is statistically significant (point estimate 0.170, std. err. 0.076, $p = 0.024$).[16]

4.4 WTP and Use

In this section we analyze short- and long-term use of the filter using our one-month and one-year data. The potential health gains of the filter may not be achieved if it is not used properly or cleaned regularly. Variation in use over time and across individuals with different levels of WTP could produce the patterns of impacts observed in the previous section.[17]

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 above the level of the tap (an indicator of whether filtered water was immediately available to drink); and (iii) whether water was in the clay filter pot. To aggregate the three measures in an agnostic way, we create an index by normalizing each measure to have mean 0 and standard deviation 1 and taking their average (Kling, Liebman and Katz 2007).

We perform the analysis in two ways. We first regress usage measures on the BDM bid among those who purchased the filter. Table 3 presents these results. Then, for comparability with our analysis of heterogeneous treatment effects in Section 4.3, we restrict the sample to winning households with children aged 0 to 5 and model the relationship

[16] In Appendix D we implement the local instrumental variables (LIV) estimator of Heckman et al. (2006). The pattern is similar, but confidence intervals for the BDM estimates are narrower by 40 percent on average.

[17] The BDM mechanism allows us to separately estimate the relationship between WTP and use from the causal effect of price paid on use. We show in Appendix E that in our data there is no evidence of a causal effect of price paid.

between WTP and use nonparametrically using kernel regression. Figure 3 displays these results for both the aggregate usage index and, for ease of interpretation, the indicator of whether filtered water was immediately available to drink.

In the short term, use is generally high. As shown in Panel A of Table 3, the filter is present and operational in nearly 90 percent of households that purchased, and filtered water is available to drink in more than 75 percent. There is limited variation with respect to WTP: neither the linear regressions of use on WTP nor the nonparametric kernel regressions reveal significant heterogeneity.

In the one-year data, use has fallen substantially. Filtered water is immediately available in fewer than half of households. Although the confidence intervals are wide, the kernel estimates now reveal substantial variation in use with respect to WTP. The conditional mean of filtered water being available ranges from 35 percent in households with a WTP of GHS 2 to 59 percent in households with a WTP of GHS 4 ($p = 0.036$). The usage index follows a similar pattern, with a difference of 0.29 standard deviations between those with a valuation of GHS 2 and those with a valuation of GHS 4 ($p = 0.096$). The magnitudes of these differences are large and economically meaningful.

These results are consistent with effort as an important mediator of treatment effects. In the short-term, effort is uniformly fairly high and there is evidence of benefits for most of the population. In the longer-term, effort and benefits have both fallen overall, and benefits are greatest in the population that is exerting the most effort.

4.5 Understanding the Pattern of Treatment Effects

In this section, we explore potential mechanisms behind the detrimental long-run impacts of the filter observed in the lower half of the WTP distribution. In Appendix H we provide a formal model and additional discussion. The possibility that compensatory responses to public health or environmental interventions could offset the intended benefits has been studied extensively in other contexts, most notably auto-safety improvements

(Peltzman 1975; Keeler 1994) and sexual health (Cassell et al. 2006; Lakdawalla et al. 2006). In settings closer to ours, Bennett (2012) finds that the introduction of piped water in the Philippines led to decreased investment in private sanitation, reversing the gains from cleaner water, and Gross et al. (2017) show that improved water sources in Benin led to decreases in point-of-use water quality, likely through changes in water handling practices. If households perceive the filter and other sanitation practices as substitutes in their health production function, receiving the filter will reduce other health investment. While the standard models of compensatory behavior could explain a muted treatment effect, they would not generate the negative treatment effects we observe. We speculate that in our setting, three additional factors may have combined with compensatory behavior to generate these detrimental effects.

First, upon receipt of the filter—a large shock to their health production function—households may have reoptimized, engaging in compensatory behavior. Then, in response to a gradual decrease in usage or the filter's effectiveness over time, they may have failed to reoptimize again, either due to rational inattention (Tobin 1982; Reis 2006; Da et al. 2014) or simple mistakes: households may have misperceived the benefits of maintaining the filter or using it regularly. If individuals are more attentive when they value the filter more, we would expect more failures to reoptimize among those with low WTP.

Second, building on an extensive literature regarding the intrahousehold allocation of health and nutrition (e.g., Pitt et al. 1990; Thomas 1997; Hoffmann 2009), even in households that purchased the filter, some children may not have had access to the filtered water. The filter produces a limited supply of drinking water, but this water comprises multiple goods. Most importantly, we consider children's health and better tasting water for adults. Before receiving the filter, households made health investments (such as traveling to cleaner water sources or boiling their water) that jointly produced both goods. The filter can decouple this production, allowing for compensatory behavior within household. With the filter, adults can obtain better tasting water with less investment in other sanitation activities that improve water quality for the entire household or compound. In particular, some children may not have been allowed to drink filtered water because of concerns that they might damage the fragile ceramic filter element or that, if they did, there would be insufficient "sweet tasting water" for the male head of the household. Field reports documented this behavior during piloting and throughout the study.[18] The pattern of treatment effects we observe is consistent with this intrahousehold mechanism. Those with a low value for children's health would be less likely to provide filtered water for their children and, all else equal, tend to have a lower WTP for the bundle of goods produced by the filter.

Finally, compensatory behavior can worsen the targeted outcome while improving utility if the alternative health production technology is non-convex. Peltzman (1973) made a similar observation looking at the effect of government subsidies on private expenditures for higher education. Many sanitation investments have a fixed cost component. For a concrete example in our setting, suppose a household can either obtain its water at low cost from a dirty source or at a higher cost from a cleaner source. Without the filter, the household chooses to incur the higher cost and drink relatively clean water. The filter improves the quality of the dirty water sufficiently that, if the household has the filter, it optimally chooses not to incur the cost of obtaining clean water. If filtered water with low other investment produces less health than unfiltered water with high other investment, purchasing the filter can increase utility but reduce health. As shown in Appendix H.1, households that make discrete changes to lower other health investments when they receive the filter will also tend to have a lower WTP for the filter itself.

With the strong caveat that households' use of the filter is endogenous, the pattern of results and usage over time is consistent with compensatory behavior playing a role in the negative treatment effects. Diarrhea rates are marginally higher for children in house-

[18] In response to the field reports, we added survey questions regarding children's access to filtered water, but subjects' answers proved unreliable.

holds that purchased the filter but are no longer using it after a year than in those that never purchased (0.34 vs. 0.24; $p = 0.069$). This difference is driven by those households who were using the filter after one month—and hence might rationally engage in compensatory behavior. Among this group, the incidence of children's diarrhea increases to 0.37 relative to 0.24 for those who never purchased ($p = 0.036$). Those who purchased but were not using the filter after one month—and hence would have been unlikely to engage in compensatory behavior—report outcomes similar to those who never purchased (0.22). Figure A9 displays these results.

While we would not suggest making policy solely based on the perverse impacts observed at the lower half of the WTP distribution in our setting, our results suggest that treatment effect heterogeneity and the potential for such adverse effects warrant more attention. We are not the first to highlight the potential for compensatory behavior that blunts or even reverses the gains from well-intentioned policy. By carefully localizing these effects to a particular part of the population that can be identified and targeted with standard policy tools, one may be able to improve outcomes. For example, knowing what we know now, it would be worth considering intensive follow-up with low-WTP households to measure and possibly support improvements to other sanitation and hygiene practices and to monitor intrahousehold allocation issues.

Together with Section 4.4, this analysis underscores the importance of effort, allocation, and related actions (both complements and substitutes) in determining outcomes. This contributes to an important but still nascent literature in economics exploring the role of subjects' behavior as a moderator of treatment effects (Chassang et al. 2012; Hanna et al. 2016). More broadly, it relates to an active literature in medicine that distinguishes between efficacy or explanatory trials, which determine treatment effects and mechanisms under ideal circumstances, and effectiveness or pragmatic trials, which measure impacts in clinical settings (Glasgow et al. 2003; Gartlehner et al. 2006). Our analysis also highlights the challenges in studying these mechanisms as each is less amenable to experimental variation than assignment of a program or product, such as the filter. The willingness-to-pay data generated by BDM allows us to explore these mechanisms. In our context, it also demonstrates that price can be used as a policy lever to screen out those with the lowest use and impacts. We present a more formal analysis of these policy implications in the next section.

5 Policy Counterfactuals and Valuing Health

In this section we explore the policy implications of the treatment effects estimated in the previous section. First, we analyze the filter's total benefits under different counterfactual prices to inform optimal pricing policy. In the short run, positive prices merely reduce access. However, in the longer term, positive prices screen out those with the lowest benefits and improve allocative efficiency. Second, we estimate a household's valuation of the filter's health benefits by combining our treatment effects estimates with the household's WTP for the filter. Our results imply low valuations compared with the value of health assigned by policy makers.

5.1 Policy Counterfactuals

In this section we show how the distribution of treatment effects estimated in Section 4 can be used to simulate impacts under alternative pricing policies. We consider a social planner who values disability-adjusted life years (DALYs) at B. The planner's choice variable is the sales price P. The social planner places equal weight on subsidy and private expenditure. That is, P is of interest only for its effect on allocation, not for the revenue generated.

Under these assumptions, the social planner will increase subsidies to lower the price P as long as the marginal cost per DALY is less than B. If the health benefits of the filter are constant at all prices, then the marginal cost per DALY will be constant. The filter will be fully subsidized if the marginal cost per DALY is below B, or not distributed at all if the marginal cost per DALY is above B. On the other hand, if the benefits of the filter are increasing in price, the social planner will set the price such that the marginal cost per DALY equals B. At this point, decreasing the price further will include households whose benefits cost more than B, and increasing the price will screen out households for whom the benefits cost are less than B.

We consider two different scenarios. First, we assume that the health gains from the one-month survey persist for a full year. While in practice the average treatment effects diminished over time, this provides a bound on the health gains if usage patterns could be maintained over the life of the filter. Since there is little evidence of heterogeneity in the short term, we assume these effects are constant with respect to WTP. Appendix F provides detail on the formulas used for calculating DALYs averted at each price and the marginal cost per DALY. Panel A of Table 4 displays the results. As the price increases (across columns), coverage decreases. Since we have assumed a constant treatment effect, cases reduced conditional on purchase are constant, and total cases reduced per household in the population decrease proportionally with demand. The same holds for DALYs gained conditional on purchase and total DALYs gained per household in the population. The constant treatment effects imply that the average and marginal costs per DALY gained are constant and equal to USD 369. A social planner with a value per DALY of at least USD 369 would maximize total gains by distributing the filter for free. This value falls below cost-effectiveness thresholds typically used by policy makers. Although precise thresholds are subject to debate, the commonly used WHO-CHOICE threshold for cost-effective interventions is one to three times annual per capita PPP GDP, or USD 2,997 to 8,991 for Ghana at the time of our study (Hutubessy et al. 2003).

In our second scenario, we assume that health effects initially are initially equal to our short-term estimates and then evolve smoothly over 12 months to the long-term estimates. We again assume the short-term effects are constant with respect to WTP and impose a linear functional form on the one-year effects. Panel B of Table 4 summarizes the impacts of different pricing policies in this scenario. Now, as price increases, negative-gains purchasers—those with low WTP—no longer purchase the filter, and diarrhea cases reduced conditional on purchase increase. For small positive prices, total gains in the population increase as well. Above a price of GHS 4, the decrease in coverage outweighs the increasing gain per household and total gains decline. We see a similar pattern in DALYs gained, both conditional on purchase (monotonically increasing with price) and total DALYs gained in the population (increasing, then decreasing, with a maximum at GHS 4). Finally, the marginal and average costs per DALY gained are monotonically decreasing. A policymaker with a value per DALY of at least USD 361 would optimally sell the filter at a price of GHS 4. A lower price would reduce total benefits, and a higher price would reduce coverage among those whose benefits cost less than USD 361 per DALY.

5.2 Valuing Children's Health

By combining our WTP data with our estimates of the impact of the filter on child health, we can directly estimate households' valuation of children's health. There are few well-identified revealed-preference estimates of this parameter, or of WTP for health or environmental quality more generally, in spite of its importance for optimal policy (Greenstone and Jack 2015). A notable exception is Kremer et al. (2011), in which the authors randomize water quality improvements at springs in Western Kenya and observe how much additional time households travel to collect better quality water. They then use wage data to convert this implicit valuation in terms of time to monetary valuation. Using this travel cost model, estimated mean WTP to avoid a case of children's diarrhea equals USD 0.89, which, with additional assumptions, translates to a value of a DALY of USD 23.7 and a value of a statistical life (VSL) of USD 754. A key advantage of our approach is that we observe WTP directly, rather than inferring it through travel time and an assumed value of time. We can simply calculate the household's observed WTP to avoid a case of diarrhea as the household's WTP for the filter divided by the number of cases avoided over the anticipated life of the filter.

While this quantity is simple to calculate in our setting, interpreting it as the household's underlying value of child health requires several assumptions. First, households know the effect of the filter on children's health. Second, households only value the filter's effect on children's health. That is, the household's WTP does not reflect other potential benefits of the filter, such as improved taste or prestige. Third, households only value reductions in diarrhea for children aged five and below. This assumption is made because diarrhea has more severe health consequences for young children, but it is also made due to data limitations: our pilot surveys indicated respondents were unable to accurately report diarrhea cases among older children or adults. Fourth, households are not liquidity constrained. Fifth, using the filter entails no change in convenience or time costs relative to current practices. We return to these assumptions at the end of this section.

We estimate households' WTP to avoid a case of diarrhea under two scenarios, making the same assumptions on treatment effects as in Section 5.1 above.[19] In the first scenario, we use the estimated impact from the one-month follow-up survey to project benefits over a year. This corresponds to the household believing that its own short-run use and maintenance practices as well as the filter's impact will persist over the first year. Again, we restrict the treatment effect to be constant with respect to WTP since there is little evidence of heterogeneous treatment effects in the short run. Figure 4a plots the distribution of WTP to avoid a case of children's diarrhea. The resulting median WTP is GHS 1.58, or USD 1.12. If we assume deaths from diarrhea are proportional to incidence and that households value only the reduction in mortality risk, not the reduced morbidity, we can compute the value of a statistical life using a ratio of mortality to incidence of one death per 3,216 cases of diarrhea in children under five, estimated for Ghana in 2010 (Global Burden of Disease Collaborative Network, 2017). The resulting median VSL is GHS 5,081

[19]See Appendix G for details on these calculations.

(USD 3,604). Again assuming that the reduction in DALYs is proportional to the reduction in incidence, we can apply a ratio of one DALY for each 35.3 cases of children's diarrhea (again estimated for Ghana in 2010 from Global Burden of Disease Collaborative Network, 2017) to calculate a median value of a DALY of GHS 55.77 (USD 39.56). Similar to the findings of Kremer et al. (2011), this is well below the typical cost effectiveness thresholds described in the previous subsection. As Kremer et al. (2011) argue, these valuations are consistent with Hall and Jones's (2007) estimates of high income elasticity of demand for health; however, we note that they are well below what would be implied by range for the income elasticity of VSL reported by Viscusi and Aldy (2003).

In the second scenario, we use both the short-term and one-year effects and compute the total effect of the filter over the first year as if the effect changed smoothly over the course of the year. We again assume the short-term effects are constant and impose a linear functional form with respect to the WTP on the one-year effects. Figure 4b plots the distribution of these estimates. The most striking feature of the graph is the large share of households with negative WTP to avoid children's diarrhea: the median WTP is GHS -0.20 (USD -0.14). Mechanically, this occurs because the average of the one-month and one-year treatment effects are negative for just over half of the population even though they exhibit positive WTP.

It is unlikely that households have a negative WTP for children's health. We posit two key explanations for this result related to Section 4.5's discussion of compensatory behavior. First, households may have misperceived the benefits of maintaining the filter or using it regularly. Improper usage or a failure to re-optimize compensatory behaviors over time could produce negative long-run treatment effects. If a household failed to foresee these actions, it might pay a positive amount for these negative treatment effects even if it valued health, and we would estimate a negative value for health. Second, as in Kremer et al. (2011), the calculations above are based on the assumption that the filter produces a single good: children's health. In fact, the filter produces multiple goods, for example, adults' health and better tasting water, that may also be valued by the household. A household's total WTP for the filter is the sum of its value for all of these goods. As discussed further in Appendix H, this bundling can explain why a household might rationally be willing to pay for the filter despite a negative impact on children's health.

While our empirical setting does not allow us to precisely identify the individual components of a household's valuation for the filter, by simply comparing valuations from households with and without children under age five we estimate that the valuation of the other goods produced by the filter could represent as much as 85 percent of the total willingness to pay. Incorporating this information in our estimates of the WTP to avoid a case of children's diarrhea would eliminate many of the negative valuations implied by the longer-term impacts.[20] These households may be willing to accept a reduction in children's health in exchange for the bundle of goods the filter provides. This highlights both the challenge and importance of constructing accurate WTP measures for health and environmental goods in developing countries.

6 Comparing Mechanisms

In addition to using BDM to conduct analyses of demand for the filter and its benefits, we designed our study to assess demand elicited under BDM with that elicited under TIOLI. While BDM produces more precise information than TIOLI offers at randomized prices, this benefit may be mitigated by its complexity. Furthermore, although bidding one's true maximum WTP is the dominant BDM strategy for expected utility maximizers, this does not necessarily hold for non-expected utility maximizers (Karni and Safra 1987; Horowitz 2006b).

There is an extensive literature in experimental economics studying the behavior of BDM among subjects in laboratory settings. It raises several issues. Several papers find

[20]Assigning a value to other goods produced by the filter would also reduce the mean and median estimates based on our short-term treatment effects.

that BDM-elicited valuations can be sensitive to the distribution of potential prices (Bohm et al. 1997; Mazar et al. 2014). Cason and Plott (2014) show that subjects' misunderstanding of the best response can also influence the WTP elicited by BDM. In addition, several studies explicitly compare BDM with other incentive-compatible elicitation mechanisms and find differences in elicited WTP (Rutstram 1998; Shogren et al. 2001; Noussair et al. 2004).

In spite of the large laboratory literature on BDM, little is known about its performance in field settings. We therefore designed our study to allow direct comparison of the demand estimates from BDM and TIOLI and to investigate the causes of any differences. Although both mechanisms are research tools and may not map directly to typical market interactions, TIOLI offers at randomized prices are common in applied research. They provide a useful benchmark for the signal contained in BDM offers. We present what is, to our knowledge, the first direct comparison of BDM and TIOLI in a developing-country field setting with the aim of better understanding the suitability of BDM for extracting additional information from field experiments.[21]

We organize the analysis comparing BDM and TIOLI as follows. Section 6.1 compares the demand estimates and out-of-sample predictive accuracy of both mechanisms. The BDM-based demand model has similar accuracy in predicting out-of-sample TIOLI decisions as the TIOLI model itself, indicating that the BDM bids contain substantial signal. As is common in the consumer behavior literature, there is substantial unobserved heterogeneity in demand estimates using either mechanism, which underscores the utility of measuring demand directly. Section 6.2 tests several potential explanations for the BDM-TIOLI demand gap. Our main finding is that the gap is largest among the most risk-averse subjects and negligible for the most risk tolerant. Finally, Section 6.3 offers some sugges-

[21]Subsequent to our study, Cole et al. (2016) study demand for weather insurance and an agricultural information service in Gujarat, India using both BDM and TIOLI. They find that BDM-measured demand is similar to that of TIOLI on average, although the exact relationship between the two mechanisms depends on the product offered. In a more traditional lab setting, de Meza and Reyniers (2013) compare BDM and TIOLI elicitation of willingness-to-accept for a handmade beeswax candle among college undergraduates. They find that willingness-to-accept is higher under BDM than TIOLI.

tions for implementing BDM in the field and identifies opportunities for future work to assess and improve its usefulness as a field research tool.

6.1 Comparing Demand Estimates and Predictive Accuracy

This section compares the correlates of demand obtained using each mechanism as well as the accuracy of each mechanism for predicting out-of-sample purchase behavior. In addition to providing a point of comparison between mechanisms, understanding the relationship between household characteristics and WTP can be directly useful by informing how pricing policies target particular types of households. Previous studies have found limited evidence that higher WTP for health goods in low-income countries is related to health characteristics or wealth (Ashraf, Berry and Shapiro 2010; Cohen and Dupas 2010), reflecting a common finding in the consumer behavior literature: choice is often only weakly correlated with standard consumer attributes (Browning and Carro 2007; Nevo 2011). This makes predicting individual purchase behavior difficult and underscores the usefulness of direct measurement of WTP.

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

$$\mathrm{WTP}_{ic} = \alpha_0 + X_{ic}^{'}\beta + \varepsilon_{ic}, \tag{7}$$

where X_{ic} is a vector of characteristics of interest for subject i in compound c, and ε_{ic} is an error term.

In our BDM sample, we observe WTP directly and can estimate Equation (7) via ordinary-least-squares. Columns 1 and 2 of Table 5 present these results. The BDM bid is positively related to the number of children aged five and under 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), is associated with an increase of GHS 0.55 in the BDM bid. The BDM bid is also positively related to durables ownership and education, although the latter is not significant. These relationships are consistent with hypotheses from the pricing literature (Ashraf, Berry and Shapiro 2010; Cohen and Dupas 2010). However, we note that, also consistent with that literature, the estimates are generally imprecise. Household characteristics explain very little of the variation in WTP. Moreover, as shown in Column 2, the best predictor of WTP for the filter is a household's WTP for soap, a related health product. When we control for a household's bid for soap in the BDM practice rounds, the share of variation explained by the model increases from 0.053 to 0.214.

For TIOLI subjects, WTP is an unobserved latent variable, so we estimate (7) indirectly using a discrete choice model:

$$\text{buy}_{i,p} = 1\left\{\text{WTP}_i \geq p_i\right\} = 1\left\{\text{WTP}_i - p_i \geq 0\right\} = 1\left\{\alpha_0 + X_{ic}'\beta + \varepsilon_{ic} - p_i \geq 0\right\} \tag{8}$$

where $\text{buy}_{i,p}$ is an indicator equal to 1 if respondent i agreed to buy when assigned price p_i. We estimate (8) on TIOLI subjects by probit. In the estimation, we normalize the coefficient on price (in GHS) to -1, so the estimated coefficients β are interpreted in terms of GHS and are comparable to those obtained by estimating Equation (7) directly with BDM subjects. Columns 3 and 4 of Table 5 presents these results.

When we compare the correlates of demand using each mechanism (Column 5 of Table 5), there are a few significant differences between the estimates for BDM and TIOLI. In several key cases, the BDM coefficient conforms more closely to hypothesized mechanisms from the literature and to our prior beliefs. For example, respondents that are more educated tend to express a higher WTP under BDM but are significantly less likely to accept a TIOLI offer at a given price. That said, and consistent with the aforementioned consumer behavior literature, much of the heterogeneity across subjects remains unexplained. For both mechanisms, a household's purchase decision for soap are more predictive of filter demand than the set of all other household characteristics. Appendix

I.1 describes the results of applying LASSO regression to determine the most relevant attributes to predict filter demand. Here too the WTP for soap in the practice round is the dominant feature predicting filter demand.

An alternative method of evaluating BDM is to analyze the extent to which it can predict non-BDM purchase behavior. We therefore compare both mechanisms on their ability to predict out-of-sample TIOLI decisions. Appendix I.2 details the procedure and provides additional results. In summary, we split each of the BDM and TIOLI samples into 10 roughly equally-sized parts or folds. For each fold k in the TIOLI sample, we use the remaining $k-1$ folds in each of the BDM and TIOLI samples to predict purchase behavior in the k^{th}, holdout, fold. We then calculate prediction error for each model and combine the estimates of the 10 folds. BDM and TIOLI correctly predict TIOLI behavior in the holdout samples correctly in 76.0 percent and 73.9 percent of observations, respectively, relative to a base rate of 56.2 percent. While additional work is required to link behavior under either mechanism to actual market purchase behavior, in this setting the predictive ability of BDM for TIOLI behavior is comparable to that of TIOLI itself.

6.2 Mechanism Effects

As shown in Figure 1a, demand is lower under BDM than TIOLI at each of the three TIOLI price points. This section investigates these differences. While the implied price elasticity of demand is almost identical under the two mechanisms, the difference in levels is not negligible. Demand under BDM is 18.2 percentage points lower that TIOLI at a price of 2 GHS ($p = 0.000$), 16.3 percentage points lower at 4 GHS ($p = 0.002$), and 10.0 percentage points lower at 6 GHS ($p = 0.012$).[22] The adjustment to BDM bids that would minimize the differences in demand at the three TIOLI price points equals GHS 1. Under the assumption that TIOLI reflects true WTP, this implies a BDM "mechanism effect" of GHS 1.

[22]See Appendix J.1 for full presentation of these results.

The remainder of this section investigates potential explanations for this gap. First, we examine the relationship between the BDM-TIOLI gap and risk aversion. Theory predicts no gap in elicited WTP between BDM and TIOLI when agents are expected utility (EU) maximizers. In our setting, there are multiple likely sources for deviations from EU maximization including loss aversion, ambiguity aversion, and non-standard beliefs about probability. Based on survey responses to questions on hypothetical gambles, 30.4% of our subjects exhibit loss aversion, 41.6% exhibit some degree of ambiguity aversion, and 64.6% at least one of these two. The theoretical literature on the BDM mechanism finds that, among non-EU maximizers, the optimal BDM bid can differ from the TIOLI reservation price, and this difference is likely to be increasing in risk aversion (Safra et al. 1990; Keller et al. 1993).

To test this hypothesis, in the one-year followup villages we collected standard survey measures of risk aversion using stated-preference responses to hypothetical gambles. (See Appendix B for detail.) We then divide the sample into terciles by risk aversion and estimate the gap separately for each tercile. In order to implement the comparison between BDM and TIOLI, we collapse the more precise individual WTP 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 BDM bid. We create the variables 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

$$\text{buy}_{icp} = \sum_{t=1}^{3} \alpha_p^t \text{RA}_i^t + \sum_{t=1}^{3} \beta_p^t \left(\text{RA}_i^t \times \text{BDM}_i\right) + x_{ic}'\gamma + \varepsilon_{icp}, \tag{9}$$

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 differences without controls are presented in Figure 5. The top panel plots the estimated coefficients $\hat{\beta}_2^1$, $\hat{\beta}_4^1$, $\hat{\beta}_6^1$, with 90 percent 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 5 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 they are robust to controlling for a large set of household controls (see Figure A6) and when testing multiple possible determinants of the BDM-TIOLI gap jointly (see Table A7).

Second, we examine how BDM-TIOLI gap differs with respect to other household observables, with the caveat that this is ex post hypothesizing rather than guided by theory. Here, we highlight the most interesting findings; we present the methods and full set of results in Appendix J.2. The mean BDM-TIOLI gap is 13.8 percentage points narrower for subjects with a child age 0 to 5 than for subjects without ($p = 0.002$). Furthermore, within the set of subjects with children age 0 to 5, the gap is 14.2 percentage points narrower if the subject reported a case of diarrhea among her young children in the previous two weeks ($p = 0.015$). In fact, among this latter group, the BDM-TIOLI gap is negligible (point estimate -0.009 , standard error of estimate 0.052, $p = 0.865$). This suggests that respondents with more at stake may have taken the exercise more seriously. These estimates are from single comparisons but are similar when testing multiple possible determinants of the BDM-TIOLI gap jointly (see Tables A7 and A8, with discussion in Appendix J.2).

Third, based on our piloting, we tested two hypotheses for reasons underlying a potential BDM-TIOLI gap: (a) that the TIOLI price offer could serve as an anchor; and (b) that subjects might be generally uncomfortable with the randomness involved in BDM. We included several variations of our basic BDM and TIOLI procedures as experimental sub-treatments designed to test these hypotheses. We found little evidence in support of our hypotheses from these sub-treatments. We provide details on the sub-treatments and analysis in Appendix J.3.

Fourth, evidence is not consistent with the gap being driven by lack of familiarity with the filter or by uncertainty about its benefits. As shown in Appendix J.4, we observe a BDM-TIOLI gap in demand for soap, a familiar product, during the practice rounds.

Finally, ex post regret—BDM subjects regretting their bid after the draw was realized—could be responsible for the BDM-TIOLI gap. This could arise from either misunderstanding the mechanism or non-EU preferences in which the resolution of uncertainty increases one's reference point. Immediately after the BDM price draw, we asked losing respondents if they wished they had bid more. A substantial share, 19.2 percent, said that they did, and Appendix J.5 explores this as a potential explanation of the differences between BDM and TIOLI. We note, however, that a comparable share of TIOLI subjects, 17.0 percent, attempted to bargain with surveyors even though the script emphasized there would be no bargaining.

6.3 Using BDM in the Field

Is the gap between the TIOLI acceptance rates and the demand curve calculated from BDM bids meaningful? In our setting, the differences between mechanisms are locally meaningful but small relative to the production costs of the filter. Still, for researchers interested in accurately predicting demand in a TIOLI environment, our results suggest caution in directly mapping BDM-elicited WTP to that of TIOLI. BDM is complex and quite different from typical market interactions. However, 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. There is little work in this area. 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 that tests both mechanisms as predictors of actual market demand. Our experience suggests the relationship between experimentally elicited demand and actual market behavior will likely depend on context and individual characteristics.

Our results also suggest at least two useful avenues of research into understanding the workings of BDM in field environments. The first follows from the finding that the BDM-TIOLI gap was close to zero among subjects displaying lowest risk aversion. This suggests exploring ways to frame BDM to reduce the salience of randomness and further emphasize the dominance of bidding one's true maximum WTP (Cason and Plott 2014). The additional confirmation steps we added to the normal BDM protocols were an attempt to move in this direction, creating explicit choices similar to a multiple price list exercise (Andersen et al. 2006) in the neighborhood of subjects' initial BDM bids. Further rigorous methodological work aimed at getting subjects to focus less on the randomization and more on how they value a product relative to a fixed sum of money would be valuable.

Second, in our exploratory analysis we found that the BDM-TIOLI gap was smaller for subjects with children aged five or under, and smaller still for those who reported that a child aged five or under had a case of diarrhea in the previous two weeks. We speculate that these subjects may have perceived that they had more at stake and taken the BDM task more seriously, thinking more carefully about their true maximum WTP.[23] This suggests further investigation of how carefully subjects consider the BDM exercise and how best to frame BDM to increase subjects' engagement. Of course, these factors are likely to be context- and product-specific, so there may not be general answers. We

[23] In the language of Harrison (1992), these subjects may perceive their payoff functions to be steeper below their optimum bid, and so face a greater possible penalty for a bid that does not equal their true maximum WTP.

expect that iteration between the field and the lab will be useful in understanding the mechanisms influencing how subjects form their bids and how different aspects of the BDM protocol may influence behavior.

Although more research is needed to evaluate the functioning of BDM in field settings, our results show that BDM is a promising tool for field research. Indeed, a number of recent papers have used BDM to elicit precise willingness to pay or willingness to accept in a range of settings (e.g., Hoffmann et al. 2009; Hoffmann 2009; Cole et al. 2014; Guiteras et al. 2016; Grimm et al. 2017; BenYishay et al. 2017). For researchers interested in using BDM in the field, Appendix K discusses some of the practical tradeoffs between BDM and TIOLI.

7 Conclusion

This paper has demonstrated the use of the BDM mechanism to elicit willingness to pay and estimate impacts for point-of-use water technology in rural Northern Ghana. Our results have several important implications for pricing policy. We find that willingness to pay for the filter is low, corresponding to less than 15 percent of the cost of production. If the policy goal is for most households to have access to in-home clean water technology, heavy subsidies or vastly cheaper alternatives will be necessary. However, demand does not fall abruptly as the price increases above zero. A small positive price would not dramatically reduce coverage. In fact, it would improve outcomes by screening out those unlikely to use or maintain the filter and for whom long-term treatment effects were negative.

Combining the elicited WTP and treatment effects estimates yields a low implied valuation for children's health: less than USD 40 per DALY and a VSL on the order of USD 3,600. The precise interpretation of these estimates remains subject to a number of caveats that apply broadly to efforts to estimate the WTP for health or environmental quality in developing countries. The filter, like many products, provides a bundle of goods, making it hard to assign a precise value to children's health alone. Households may not have understood the effect of water filtration on health, or perhaps they correctly projected that usage and maintenance would be imperfect and the benefits would wane over time. However, consistent with the work of Kremer et al. (2011) in Kenya, the magnitude of the implied valuation is far below those typically used by public health planners or estimated in higher income countries (Viscusi and Aldy 2003).

We also show that behavior matters for the effectiveness of the filter: the filter's benefits decrease over time and even become negative for households exerting low effort. Along with Hanna et al. (2016), this highlights the importance of considering household behavior when evaluating health and environmental technologies. Even a technically sound product can have its effects blunted by slippage in consistency or quality of use, and policymakers should not underestimate the importance of costly effort. One direction to pursue is to invest more in understanding user behavior and working to motivate and sustain behavioral change. A second is to refine existing products or develop new products that are less dependent on correct use or impose lower effort costs on the user.

As we demonstrate, embedding BDM in field experiments can also provide insights into how usage and treatment effects vary with WTP. This is a key dimension of heterogeneity both for policy analysis and for uncovering structural parameters along the lines emphasized by Heckman, Vytacil and Urzua (1999; 2005; 2006). 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 valuations—the procedure can be readily understood, even in an environment with low literacy and numeracy.

However, this added information comes with the potential cost of added complexity. Experimental mechanisms to recover valuations differ from normal market interactions, but BDM can seem particularly unusual. We are encouraged that the predictive power of

BDM estimates for TIOLI behavior is comparable to that of TIOLI itself and price elasticity estimates are similar under both mechanisms. BDM is doing more than generating precise noise. However, demand under BDM is systematically lower than TIOLI at each of the TIOLI price points, particularly among the most risk-averse households. Ultimately, the value of implementing BDM will depend on the context. Further research exploring how and when BDM can be a useful tool in field settings would be highly valuable.

References

Ahuja, A., M. Kremer, and A. P. Zwane, "Providing Safe Water: Evidence from Randomized Evaluations," *Annual Review of Resource Economics*, 2010, *2*, 237–256.

Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutström, "Elicitation Using Multiple Price List Formats," *Experimental Economics*, December 2006, *9* (4), 383–405.

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

Becker, G. M., M. H. DeGroot, and J. Marschak, "Measuring utility by a single-response sequential method," *Behavioral Science*, 1964, *9*, 226–236.

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

BenYishay, A., A. Fraker, R. P. Guiteras, G. Palloni, N. Shah, S. Shirrell, and P. Wang, "Microcredit and willingness to pay for environmental quality: Evidence from a randomized-controlled trial of finance for sanitation in rural Cambodia," *Journal of Environmental Economics and Management*, November 2017, *86*, 121–140.

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

Brinch, C. N., M. Mogstad, and M. Wiswall, "Beyond LATE with a Discrete Instrument," *Journal of Political Economy*, 2017, *125* (4), 985–1039.

Brown, J. and T. Clasen, "High Adherence Is Necessary to Realize Health Gains from Water Quality Interventions," *PLoS ONE*, May 2012, *7* (5), e36735.

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

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

Cassell, M. M., D. T. Halperin, J. D. Shelton, and D. Stanton, "HIV and risk behaviour: Risk compensation: the Achilles' heel of innovations in HIV prevention?," *BMJ: British Medical Journal*, 2006, *332* (7541), 605.

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

Clasen, T. F., K. T. Alexander, D. Sinclair, S. Boisson, R. Peletz, H. H. Chang, F. Majorin, and S. Cairncross, "Interventions to Improve Water Quality for Preventing Diarrhoea," in "Cochrane Database of Systematic Reviews," Vol. 10 October 2015.

Clopeck, K. L., "Monitoring and Evaluation of Household Water Treatment and Safe Storage Technologies: The Sustained Use of the KOSIM Ceramic Water Filter in Northern Region Ghana." MA Thesis, Massachusetts Institute of Technology 2009.

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

Cole, S. A., A. N. Fernando, D. Stein, and J. Tobacman, "Field Comparisons of Incentive-Compatible Preference Elicitation Techniques," *Mimeo, University of Delaware*, 2016.

_ , D. Stein, and J. 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.

Cole, S., X. Giné, J. Tobacman, P. Topalova, R. Townsend, and J. Vickery, "Barriers to Household Risk Management: Evidence from India," *American Economic Journal: Applied Economics*, January 2013, *5* (1), 104–35.

Da, Z., U. G. Gurun, and M. Warachka, "Frog in the Pan: Continuous Information and Momentum," *The Review of Financial Studies*, 2014, *27* (7), 2171–2218.

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

Dupas, P., "Health Behavior in Developing Countries," *Annual Review of Economics*, 2011, *3*, 425–449.

_ and E. Miguel, "Impacts and Determinants of Health Levels in Low-Income Countries," in A. V. Banerjee and E. Duflo, eds., *Handbook of Economic Field Experiments*, Vol. 2 of *Handbook of Economic Field Experiments*, North-Holland, January 2017, pp. 3–93.

Ellsberg, D., "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics*, 1961, *75* (4), 643–669.

Foster, A. D. and M. R. Rosenzweig, "Microeconomics of Technology Adoption," *Annual Review of Economics*, 2010, *2*, 395–424.

Gartlehner, G., R. A. Hansen, D. Nissman, K. N. Lohr, and T. S. Carey, "A simple and valid tool distinguished efficacy from effectiveness studies," *Journal of Clinical Epidemiology*, October 2006, *59* (10), 1040–1048.

Glasgow, R. E., E. Lichtenstein, and A. C. Marcus, "Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition," *American Journal of Public Health*, August 2003, *93*, 1261–1267.

Global Burden of Disease Collaborative Network, "Global Burden of Disease Study 2016 Results," 2017.

Greenstone, M. and B. K. Jack, "Envirodevonomics: A Research Agenda for an Emerging Field," *Journal of Economic Literature*, March 2015, *53* (1), 5–42.

Grimm, M., L. Lenz, J. Peters, and M. Sievert, "Demand for off-grid solar electricity – Experimental evidence from Rwanda," February 2017. ZEF Discussion Paper 232.

Gross, E., I. Guenther, and Y. Schipper, "Water Supply and Water Handling: Perceived Substitutes but Revealed Supplements for Safe Water," 2017.

Guiteras, R. P., D. I. Levine, S. P. Luby, T. H. Polley, K. K. e Jannat, and L. 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*, June 2016, *3* (2), 321–359.

_ , J. Levinsohn, and A. M. Mobarak, "Encouraging Sanitation Investment in the Developing World: A Cluster-Randomized Trial," *Science*, May 2015, *348* (6237), 903–906.

Hall, R. E. and C. I. Jones, "The Value of Life and the Rise in Health Spending," *The Quarterly Journal of Economics*, 2007, *122* (1), 39–72.

Hanna, R., E. Duflo, and M. Greenstone, "Up in Smoke: The Influence of Household Behavior on the Long-Run Impact of Improved Cooking Stoves," *American Economic Journal: Economic Policy*, February 2016, *8* (1), 80–114.

Harrison, G. W., "Theory and Misbehavior of First-Price Auctions: Reply," *The American Economic Review*, 1992, *82* (5), 1426–1443.

Heckman, J. J. and E. J. Vytlacil, "Local instrumental variables and latent variable models for identifying and bounding treatment effects," *Proceedings of the National Academy of Sciences*, 1999, *96* (8), 4730–4734.

_ 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 Environments," in J. Heckman and E. Leamer, eds., *Handbook of Econometrics*, Vol. 6 September 2007, chapter 71.

_ and S. Urzúa, "Comparing IV with Structural Models: What Simple IV Can and Cannot Identify," *Journal of Econometrics*, May 2010, *156* (1), 27–37.

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

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

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

Horowitz, J. 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.

Hutubessy, R., D. Chisholm, T. T.-T. Edejer, and WHO-CHOICE, “Generalized cost-effectiveness analysis for national-level priority-setting in the health sector,” *Cost Effectiveness and Resource Allocation*, 2003, *1* (8).

Imbens, G. W. and J. D. Angrist, “Identification and Estimation of Local Average Treatment Effects,” *Econometrica*, March 1994, *62* (2), 467–475.

Ito, K. and S. Zhang, “Willingness to Pay for Clean Air: Evidence from Air Purifier Markets in China,” Working Paper 22367, National Bureau of Economic Research June 2016.

Jack, B. K., “Constraints on the Adoption of Agricultural Technologies in Developing Countries,” White paper, Agricultural Technology Adoption Initiative, Boston: J-PAL (MIT) and Berkeley: CEGA (UC Berkeley) 2011.

Johnson, S. M., “Health and Water Quality Monitoring of Pure Home Water’s Ceramic Filter Dissemination in the Northern Region of Ghana.” MA Thesis, Massachusetts Institute of Technology June 2007.

Jones, D., “The Economics of Exclusion Restrictions in IV Models,” NBER Working Paper 21391 July 2015.

Karlan, D. and J. Zinman, “Observing Unobservables: Identifying Information Asymmetries With a Consumer Credit Field Experiment,” *Econometrica*, 2009, *77* (6), 1993–2008.

Karni, E. and Z. Safra, “Preference Reversal and the Observability of Preferences by Experimental Methods,” *Econometrica*, 1987, *55* (3), 675–685.

Keeler, T. E., “Highway Safety, Economic Behavior, and Driving Environment,” *The American Economic Review*, 1994, *84* (3), 684–693.

Keller, L. R., U. Segal, and T. 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, J. R., J. B. Liebman, and L. F. Katz, “Experimental Analysis of Neighborhood Effects,” *Econometrica*, 2007, *75 (1)*, 83–119.

Kowalski, A. E., “Doing More When You’re Running LATE: Applying Marginal Treatment Effect Methods to Examine Treatment Effect Heterogeneity in Experiments,” NBER Working Paper 22363 June 2016.

Kremer, M. and A. Holla, "Pricing and Access: Lessons from Randomized Evaluations in Education and Health," in W. Easterly and J. Cohen, eds., *What Works in Development: Thinking Big and Thinking Small*, Brookings Institution Press, 2009.

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

Lakdawalla, D., N. Sood, and D. Goldman, "HIV breakthroughs and risky sexual behavior," *The Quarterly Journal of Economics*, 2006, *121* (3), 1063–1102.

Li, Q. and J. S. Racine, *Nonparametric Econometrics*, Princeton University Press, 2007.

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

Miguel, E. and M. K. Gugerty, "Ethnic Diversity, Social Sanctions, and Public Goods in Kenya," *Journal of Public Economics*, December 2005, *89* (11), 2325–2368.

Miller, K. M., R. Hofstetter, H. Krohmer, and Z. J. 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.

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

Noussair, C., S. Robin, and B. 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.

Peltzman, S., "The Effect of Government Subsidies-in-Kind on Private Expenditures: The Case of Higher Education," *Journal of Political Economy*, January 1973, *81* (1), 1–27.

_ , "The Effects of Automobile Safety Regulation," *Journal of Political Economy*, 1975, *83* (4), 677–725.

Pitt, M. M., M. R. Rosenzweig, and M. N. Hassan, "Productivity, Health, and Inequality in the Intrahousehold Distribution of Food in Low-income Countries," *The American Economic Review*, 1990, pp. 1139–1156.

Plott, C. R. and K. Zeiler, "The Willingness to Pay-Willingness to Accept Gap, the Endowment Effect, Subject Misconceptions, and Experimental Procedures for Eliciting Valuations," *American Economic Review*, June 2005, *95* (3), 530–545.

Reis, R., "Inattentive Consumers," *Journal of Monetary Economics*, 2006, *53* (8), 1761–1800.

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

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

Shogren, J. F., S. Cho, C. Koo, J. List, C. Park, P. Polo, and R. Wilhelmi, "Auction mechanisms and the measurement of WTP and WTA," *Resource and Energy Economics*, 2001, *23* (2), 97–109.

Thomas, D., "Incomes, Expenditures, and Health Outcomes: Evidence on Intrahousehold Resource Allocation," *Intrahousehold Resource Allocation in Developing Countries*, 1997, pp. 142–64.

Tibshirani, R., "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, 1996, *58* (1), 267–288.

Tobin, J., "Money and finance in the macroeconomic process," *Journal of Money, Credit and Banking*, 1982, *14* (2), 171–204.

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

Viscusi, W. K. and J. E. Aldy, "The Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World," *Journal of Risk and Uncertainty*, 2003, *27* (1), 6–76.

WHO, "Global Health Estimates 2015: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2015," Technical Report, World Health Organization, Geneva 2016.

WHO and UNICEF, "Progress on Drinking Water, Sanitation and Hygiene: 2017 Update and SDG Baselines," Technical Report, Joint Monitoring Programme of World Health Organization (WHO) and the United Nations Children's Fund (UNICEF), Geneva 2017.

Wright, J., S. Gundry, and R. Conroy, "Household Drinking Water in Developing Countries: A Systematic Review of Microbiological Contamination between Source and Point-of-Use," *Tropical Medicine & International Health*, January 2004, *9* (1), 106–117.

Table 1: Sample Composition and Descriptive Statistics

	Mean			Diff.	Regressions	
	Full Sample	BDM	TIOLI	BDM-TIOLI	BDM Draw	TIOLI Price
	(1)	(2)	(3)	(4)	(5)	(6)
Number of respondents in compound (census)	3.593	3.305	3.859	-0.554*	0.236***	-0.051
	[2.323]	[1.816]	[2.683]	(0.323)	(0.079)	(0.045)
Husband lives in compound	0.794	0.792	0.796	-0.004	0.453	-0.243
	[0.404]	[0.406]	[0.403]	(0.022)	(0.367)	(0.168)
Number of children age 0-5 in household	1.135	1.069	1.196	-0.127*	0.195	0.028
	[0.978]	[0.941]	[1.008]	(0.073)	(0.159)	(0.078)
Number of children age 6-17 in household	1.303	1.389	1.224	0.165**	0.028	-0.013
	[1.282]	[1.304]	[1.258]	(0.084)	(0.129)	(0.047)
Number of children age 0-5 with diarrhea in past two weeks	0.243	0.208	0.277	-0.069*	-0.372	0.075
	[0.525]	[0.487]	[0.557]	(0.035)	(0.376)	(0.128)
Number of children age 6-17 with diarrhea in past two weeks	0.049	0.050	0.048	0.002	-0.499	0.463*
	[0.272]	[0.302]	[0.241]	(0.016)	(0.417)	(0.267)
Respondent has ever attended school	0.090	0.079	0.100	-0.021	-0.025	-0.077
	[0.286]	[0.270]	[0.301]	(0.016)	(0.515)	(0.195)
First principal component of durables ownership	0.132	0.059	0.198	-0.139	-0.046	0.005
	[1.555]	[1.512]	[1.592]	(0.126)	(0.091)	(0.056)
All-year access to improved water source	0.187	0.196	0.179	0.017	-0.126	0.119
	[0.390]	[0.397]	[0.384]	(0.038)	(0.376)	(0.252)
Currently treats water	0.115	0.109	0.120	-0.011	0.567	0.048
	[0.319]	[0.312]	[0.325]	(0.024)	(0.468)	(0.257)
E. coli count, standardized	-0.052	-0.026	-0.076	0.050	-0.102	0.038
	[0.949]	[1.012]	[0.887]	(0.089)	(0.162)	(0.120)
Turbidity, standardized	-0.065	-0.099	-0.032	-0.068	-0.008	0.224***
	[0.997]	[0.922]	[1.063]	(0.096)	(0.178)	(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. Standard errors clustered at the compound (extended family) level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Constant-Effects Instrumental Variables Estimates

Dependent Variable: Child age 0 to 5 has had diarrhea over previous two weeks

	Combined all subjects		TIOLI subjects		BDM subjects	
	(1)	(2)	(3)	(4)	(5)	(6)
A. One-month followup						
Bought Filter	-0.065*	-0.072**	-0.100*	-0.098*	-0.049	-0.058
	(0.037)	(0.035)	(0.054)	(0.051)	(0.050)	(0.043)
Mean dependent variable	0.145	0.145	0.149	0.149	0.142	0.142
Number of compounds	472	472	244	244	229	229
Number of subjects	786	786	418	418	368	368
Number of children	1244	1244	665	665	579	579
B. One-year followup						
Bought Filter	0.093	0.121*	0.148	0.220**	0.090	0.108
	(0.070)	(0.071)	(0.099)	(0.100)	(0.089)	(0.090)
Mean dependent variable	0.241	0.241	0.215	0.215	0.262	0.262
Number of compounds	247	247	121	121	126	126
Number of subjects	387	387	197	197	190	190
Number of children	539	539	266	266	273	273
Controls	No	Yes	No	Yes	No	Yes
Village FEs	No	Yes	No	Yes	No	Yes

Notes: Each column displays the results of a linear two-stage least squares regression of child diarrhea status at the child level on filter purchase, where filter purchase is instrumented by random BDM draw for BDM subjects and by randomly assigned TIOLI price for TIOLI subjects. Controls include all variables (other than BDM bid) listed in Table 1. Missing values of control 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: Relationship between Use and Willingness to Pay

	Filter present and unbroken (1)	Storage vessel contains water (2)	Clay pot contains water (3)	Usage index (4)
A. Short-term effects				
Bid (GHS)	-0.010	-0.008	-0.009	-0.022
	(0.010)	(0.012)	(0.013)	(0.021)
Mean dep. var.	0.877	0.753	0.728	-0.003
Adj. R-sqd.	0.002	-0.002	-0.002	0.002
Num. Obs.	235	235	235	235
B. One-year effects				
Bid (GHS)	0.013	0.027*	-0.013	0.018
	(0.014)	(0.014)	(0.012)	(0.021)
Mean dep. var.	0.641	0.486	0.380	0.066
Adj. R-sqd.	-0.002	0.016	-0.002	-0.003
Num. Obs.	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. Usage index is the average of the normalized values of the three individual usage measures. Usage measures are observed by the 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 4: Estimated Impacts of Pricing Policy

	Price (GHS)						
	0	1	2	3	4	5	6
Share Purchasing	1.00	0.94	0.73	0.46	0.31	0.19	0.11
A. Constant one-month effects							
Diarrhea cases averted per household (conditional on purchase)	1.43	1.43	1.43	1.43	1.43	1.43	1.43
Diarrhea cases averted per household (unconditional)	1.43	1.35	1.05	0.66	0.44	0.28	0.15
DALYs averted per household (conditional on purchase)	0.041	0.041	0.041	0.041	0.041	0.041	0.041
DALYs averted per household (unconditional)	0.041	0.038	0.030	0.019	0.013	0.008	0.004
Average social cost per DALY (USD)	369	369	369	369	369	369	369
Marginal cost per DALY (USD)		369	369	369	369	369	369
B. Average of one-month effects and one-year effects							
Diarrhea cases averted per household (conditional on purchase)	-1.09	-0.72	0.62	2.73	4.29	5.73	6.81
Diarrhea cases averted per household (unconditional)	-1.09	-0.68	0.46	1.26	1.33	1.10	0.73
DALYs averted per household (conditional on purchase)	-0.031	-0.021	0.018	0.077	0.121	0.162	0.193
DALYs averted per household (unconditional)	-0.031	-0.019	0.013	0.036	0.038	0.031	0.021
Average social cost per DALY (USD)	–	–	849	194	123	92	78
Marginal cost per DALY (USD)		–	–	–	361	128	79

Notes: In Panel A, short-term impacts on diarrhea are assumed to be constant and last for one year. Panel B assumes the average of short- and long-term impacts last for one year. In Panel B, the short-term impacts are constant and the long-term impacts are linear in willingness-to-pay. Diarrhea incidence is converted to DALYs at the rate of 35.3 cases per year to one DALY, using data from the Global Burden of Disease Collaborative Network (2017). The average social cost does not account for revenue generated from sales. The marginal cost per DALY is computed as the difference in costs between price $P - 0.5$ and price $P + 0.5$ divided by the difference in DALYs averted between price $P - 0.5$ and price $P + 0.5$. Missing entries in the average and marginal cost rows indicate that costs cannot not be computed because treatment effects are negative for average or marginal households at the prices indicated.

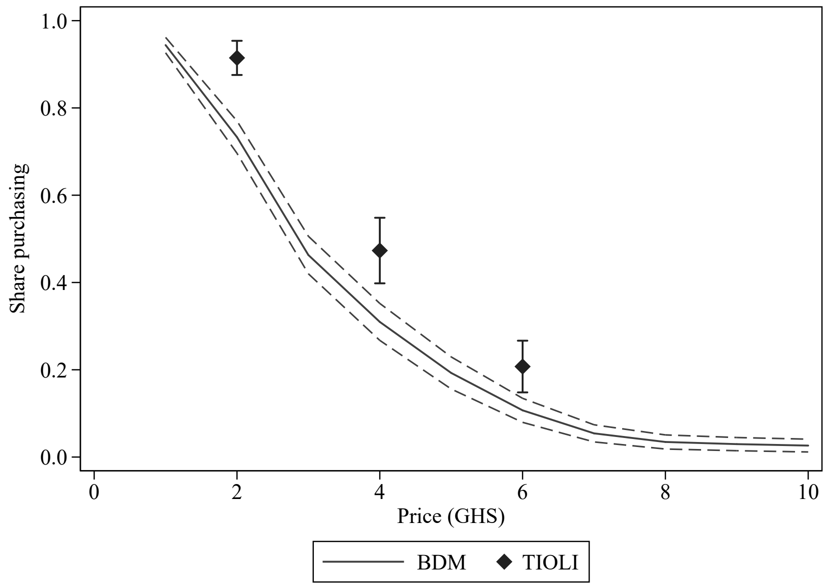
Table 5: Correlates of Willingness to Pay

	BDM		TIOLI		
	OLS		Probit		Diff. (2)-(4)
	(1)	(2)	(3)	(4)	(5)
Number of respondents in compound (census)	0.053	0.085	-0.089^{***}	-0.117^{***}	0.203^{***}
	(0.061)	(0.059)	(0.034)	(0.035)	(0.068)
Husband lives in compound	-0.005	0.157	-0.463^{*}	-0.471^{**}	0.629^{**}
	(0.249)	(0.220)	(0.244)	(0.233)	(0.318)
Number of children age 0-5 in household	0.067	0.098	-0.066	-0.053	0.151
	(0.114)	(0.098)	(0.092)	(0.093)	(0.134)
Number of children age 6-17 in household	0.018	-0.013	0.197^{**}	0.172^{**}	-0.185^{*}
	(0.068)	(0.064)	(0.080)	(0.080)	(0.102)
Number of children age 0-5 with diarrhea in past two weeks	0.550^{*}	0.387	-0.260	-0.284	0.671^{**}
	(0.290)	(0.266)	(0.170)	(0.175)	(0.315)
Number of children age 6-17 with diarrhea in past two weeks	-0.187	-0.210	-0.663^{*}	-0.592^{*}	0.382
	(0.223)	(0.228)	(0.355)	(0.343)	(0.409)
Respondent has ever attended school	0.604	0.556	-0.535^{**}	-0.542^{**}	1.098^{**}
	(0.418)	(0.410)	(0.236)	(0.239)	(0.470)
First principal component of durables ownership	0.128^{*}	0.011	0.099	0.102	-0.092
	(0.075)	(0.066)	(0.072)	(0.068)	(0.094)
All-year access to improved water source	-0.307	-0.074	-0.259	-0.220	0.146
	(0.253)	(0.231)	(0.265)	(0.257)	(0.344)
Currently treats water	0.560	0.526	0.246	0.076	0.451
	(0.378)	(0.344)	(0.270)	(0.274)	(0.435)
E. coli count, standardized	-0.123	-0.180^{*}	0.134	0.088	-0.269
	(0.111)	(0.103)	(0.161)	(0.166)	(0.194)
Turbidity, standardized	-0.190^{**}	-0.217^{**}	0.076	0.042	-0.259^{*}
	(0.087)	(0.089)	(0.123)	(0.117)	(0.146)
BDM Soap Bid (GHS)		3.527^{***}			
		(0.579)			
Purchased Soap				1.195^{***}	
				(0.261)	
R-squared	0.053	0.214			
Log-likelihood			-347.1	-321.2	
Number of households	607	607	657	656	
Number of compounds	275	275	283	282	

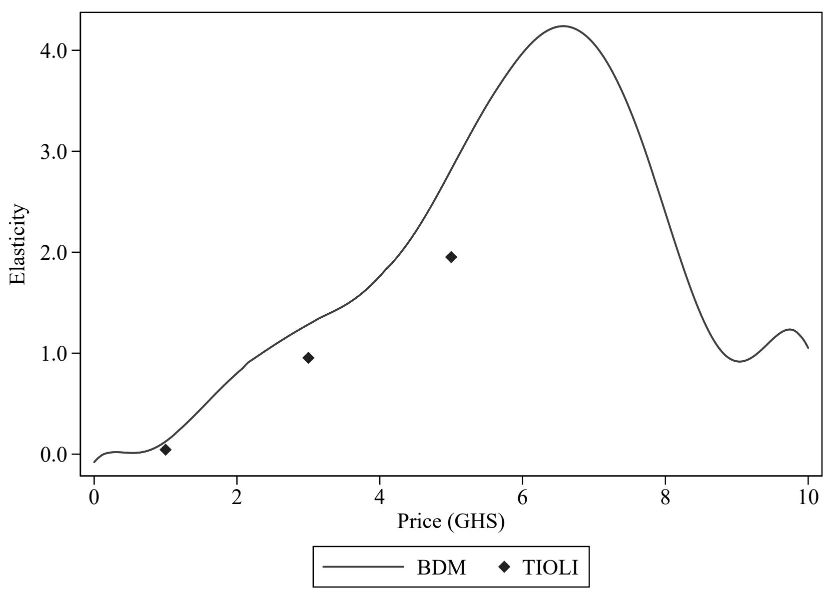
Notes: Columns (1) and (2) display coefficients from a linear regression of directly reported willingness to pay (the BDM bid) on baseline characteristics. Columns (3) and (4) report coefficients from probit models, where the dependent variable is the TIOLI purchase decision. As discussed in the text, 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 and are comparable to the OLS estimates from the BDM subjects. Missing values of the independent variables are set to 0, and dummy variables are included to indicate missing values. Column (5) reports differences in the estimated coefficients between BDM (Column (2)) and TIOLI (Column (4)), with standard errors calculated via SUR. 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



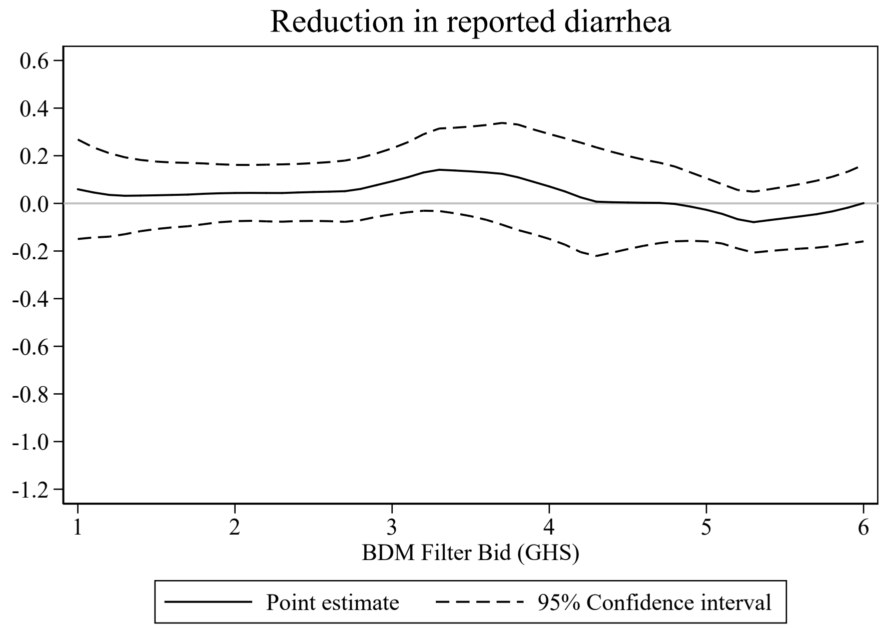
(b) Price Elasticity of Demand



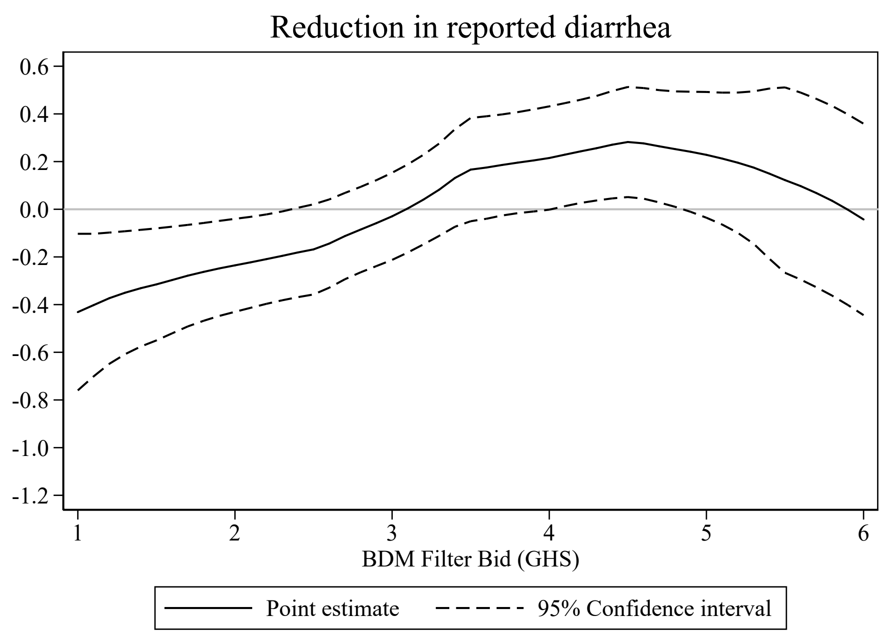
Notes: The top panel plots the BDM demand curve, with a 90% confidence band, and take-it-or-leave-it (TIOLI) demand at three price points (2, 4 and 6 GHS), with 90% confidence intervals. The BDM demand curve indicates the share of respondents with a BDM filter bid greater than or equal to the indicated 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: Kernel IV Estimates of Treatment Effects

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



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

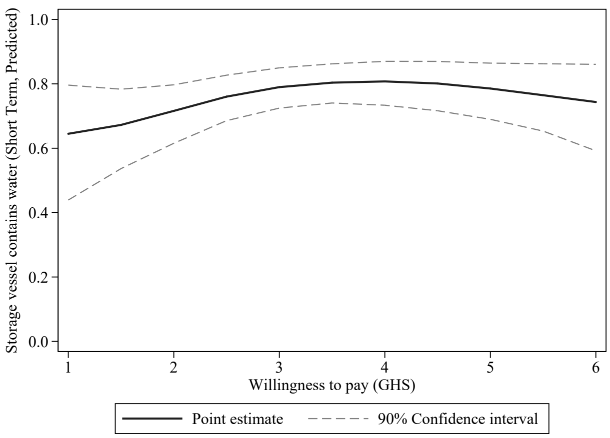


Notes: These graphs present estimated treatment effects (reduction in diarrhea among children age 0 to 5) as a function of willingness-to-pay (WTP). Estimates are by linear two-stage least squares at WTP = 1.0, 1.1, . . . , 6.0, weighting observations by their distance from the evaluation point (Epanechnikov kernel, bandwidth by Silverman's rule of thumb). 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 4.3 for details, and Figures A4 and A5 for ancillary statistics (sample sizes and instrument strength) and first-stage results.

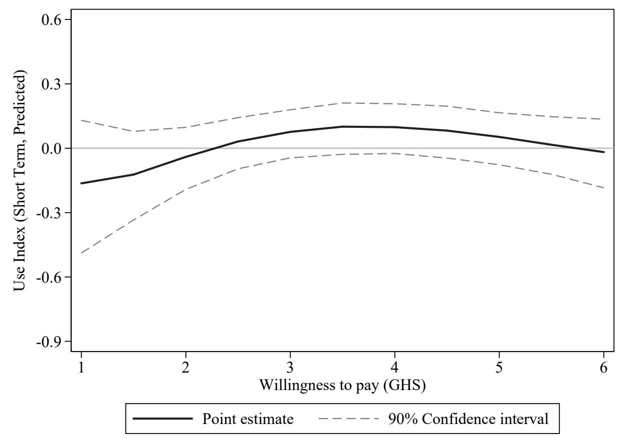
Figure 3: Relationship between Use and Willingness to Pay
BDM Purchasers with Children 0 to 5

(a) One-Month Follow Up

Storage Vessel Contains Water

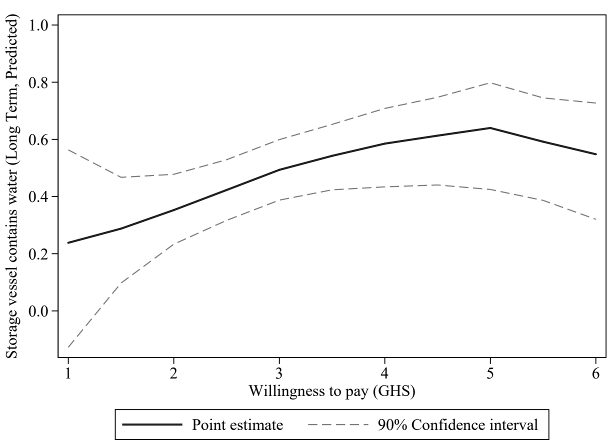


Use Index

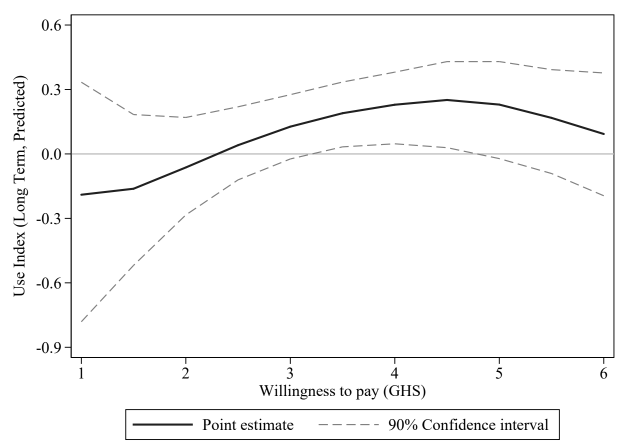


(b) One-Year Follow Up

Storage Vessel Contains Water



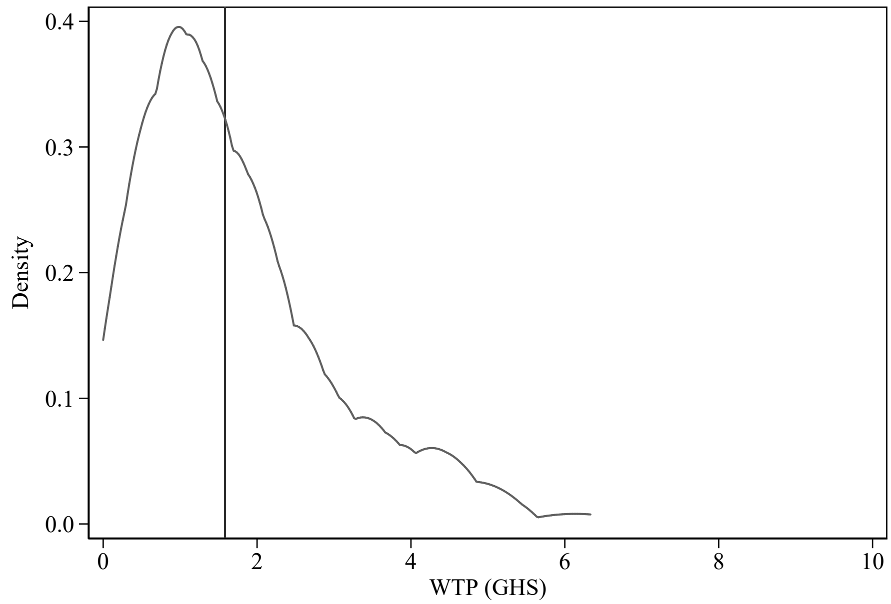
Use Index



Notes: These figures show predicted values from a kernel regression (local polynomial of degree 1) for measures of usage on the household's willingness-to-pay (WTP), as stated in the BDM sale. The left figures display an indicator for whether the safe storage container contained water at or above the level of the spigot. The right figures display an index of use measures comprising indicators for whether the filter was observed in the compound, whether the ceramic pot contained water, and whether the safe storage container contained water at or above the level of the spigot. These measures are standardized and combined following 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).

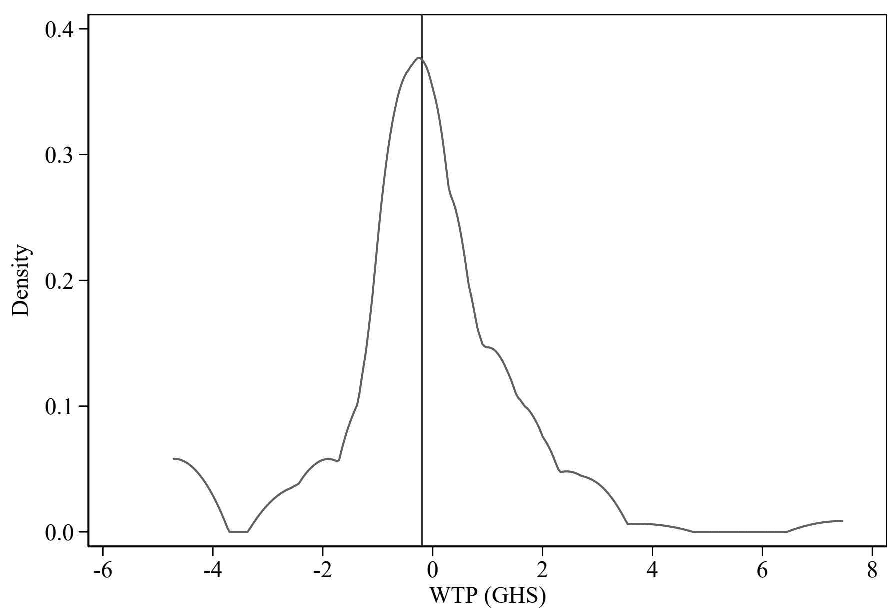
Figure 4: WTP to Avoid a Case of Children's Diarrhea

(a) One-Month Treatment Effect



Median WTP of GHS 1.58 indicated by vertical line.

(b) Average Treatment Effect over One Year

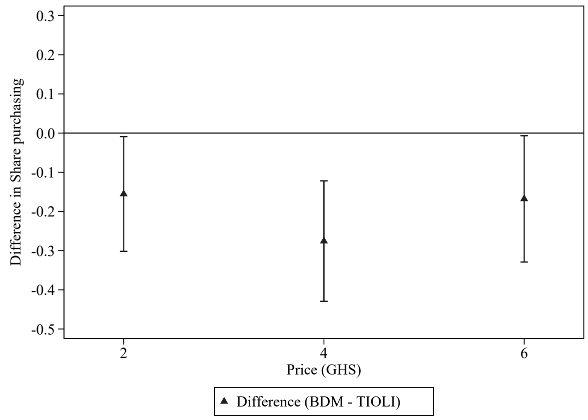


Median WTP of GHS -0.20 indicated by vertical line.

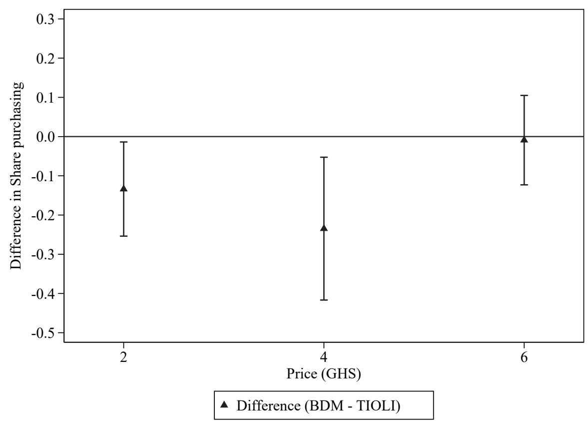
Notes: These figures present distributions of the WTP to avoid a case of diarrhea based on BDM bids and the treatment effects estimated in Section 4. In the top panel, short-term impacts on diarrhea are assumed to be constant and last for one year. The bottom panel assumes the average of short- and long-term impacts last for one year. In the bottom panel, the short-term impacts are constant and the long-term impacts are linear in willingness-to-pay.

Figure 5: BDM–TIOLI gap by tercile of risk aversion

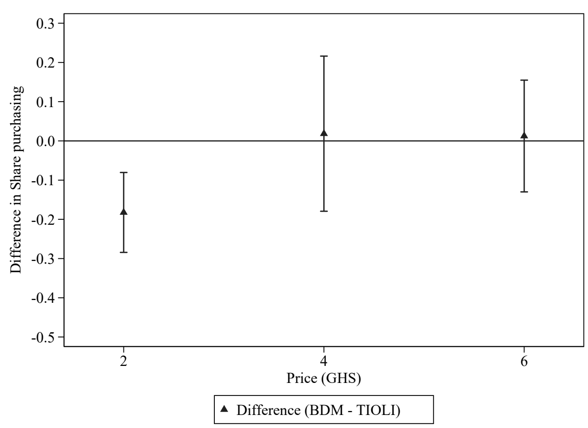
(a) Tercile 1 (most risk-averse)



(b) Tercile 2



(c) Tercile 3 (least risk-averse)



Notes: These figures plot estimated differences, with with 90 percent confidence intervals, 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 A6 for robustness checks with additional controls.