Maryland Population Research Center

WORKING PAPER

Microcredit and Willingness to Pay for Environmental Quality: Evidence from a Randomized-Controlled Trial of Finance for Sanitation in Rural Cambodia

PWP-MPRC-2016-003

January 2016

Authors:

Ariel BenYishay
William and Mary

Andrew Fraker
IDinsight

Raymond P. Guiteras
University of Maryland

Giordano Palloni
IFPRI

Neil Buddy Shah
IDinsight

Stuart Shirrell
IDinsight

Paul Wang
IDinsight

www.popcenter.umd.edu

Microcredit and Willingness to Pay for Environmental Quality: Evidence from a Randomized-Controlled Trial of Finance for Sanitation in Rural Cambodia*

Ariel BenYishay
William and Mary

Andrew Fraker
IDinsight

Raymond Guiteras†
University of Maryland

Giordano Palloni
IFPRI

Neil Buddy Shah
IDinsight

Stuart Shirrell
IDinsight

Paul Wang
IDinsight

January 2016

Abstract

Low willingness to pay (WTP) for environmental quality in developing countries is a key research question in environmental economics. One explanation is that missing credit markets may suppress WTP for environmental improvements that require large up-front investments. We test the impact of microloans on WTP for hygienic latrines via a randomized controlled trial in 30 villages in rural Cambodia. We find that microcredit dramatically raises WTP for improved latrines, with 60% of households in the Financing arm willing to purchase at an unsubsidized price, relative to 25% in the Non-financing arm. Effects on latrine installation are positive but muted by several factors, including a negative peer effect: randomly induced purchases by neighbors reduce a household's probability of installing its own latrine. On methodological grounds, this paper shows that a "decision-focused evaluation" can be integrated into academic analysis to provide insight into questions of general interest.

Keywords: Willingness to pay, sanitation, microcredit, Becker-DeGroot-Marschak, randomization inference

*We received helpful comments from Maureen Cropper, Doug Miller, Brian Quistorff and Chris Udry. This paper's pre-analysis plan is posted at the AEA RCT Registry, `https://www.socialscienceregistry.org/trials/37`. We discuss differences with the pre-specified analysis in Appendix A. We are grateful to IDinsight (particularly Benjamin Brockman, Eva Ghirmai, and Esther Wang), iDE Cambodia (particularly Karen Genzink, Chris Nicoletti, Ou Savoeun, Matt Seitz, and Toeur Veasna) and VisionFund Cambodia (Preap Piseth and Lim Sotheary) for their collaboration. This research was supported by the Bill and Melinda Gates Foundation through a grant to iDE. Neither the funders nor our implementation partners had any role in the analysis, writing of the paper, or decision to submit for publication. The study protocol was approved by the National Ethics Committee for Health Research, Ministry of Health, Kingdom of Cambodia. ABY, RG and GP dedicate this paper to the memory of Wallace E. Oates.

†Corresponding author. Mail: Department of Economics, 3114 Tydings Hall, University of Maryland, College Park, MD 20742 USA; email: guiteras@econ.umd.edu; tel: 301-405-3480; fax: 301-405-3542.

1 Introduction

Environmental conditions are typically worse in poor countries along key dimensions such as clean air and clean water (WHO 2014a, b, 2015), and the burden of environmental disease similarly falls disproportionally on poor countries (Prüss-Üstün and Corvalán 2006). However, willingness to pay (WTP) for environmental quality is typically low in developing countries, and understanding the reasons why is a key research question at the intersection of environmental and development economics (Greenstone and Jack 2015).

Poor sanitation in developing countries is an important example of this general phenomenon. Globally, 2.5 billion people live without access to improved sanitation, with one billion of these people practicing open defecation (WHO and UNICEF 2014). Inadequate sanitation is believed to cause 280,000 deaths per year (Prüss-Ustün et al. 2014), contribute to serious health problems such as chronic diarrhea and tropical enteropathy (Dangour et al. 2013; Lin et al. 2013), and may diminish human capital through impacts on stunting and cognitive capacity (Spears 2013; Spears and Lamba 2013). Although simple, relatively affordable solutions such as low-cost pour-flush latrines exist and major policy initiatives have promoted their adoption, growth in latrine coverage and reduction in open defecation have been slow in many parts of the developing world.

Several explanations for this puzzle have been proposed, including lack of information on health benefits, peer effects and social influence, supply-side failures, and difficulties in coordinating in the face of externalities and complementarities (Pattanayak and Pfaff 2009; Pattanayak et al. 2009; Trémolet 2012; Perez et al. 2012; Gertler et al. 2015; Guiteras et al. 2015b). In this paper, we focus on one aspect of a household's decision to purchase and install a latrine: latrines require a large up-front investment whose benefits are realized over time. This is a common characteristic of many actions required to improve environmental conditions and health, and could limit investment for any of several reasons (e.g., lack of consumer credit, high discount rates, or present bias) and suggests that interventions to fill missing capital markets may increase investment and improve welfare. However, there is little evidence on the effectiveness of credit-based interventions at affecting WTP for environmental quality (Greenstone and Jack 2015), and recent research showing limited impacts of micro-credit interventions on income and welfare has lead to increased skepticism of the value of micro-finance (Banerjee 2013; Banerjee et al. 2015).

This paper reports the results of what is, to our knowledge, the first randomized-controlled trial of the effect of micro-loans on WTP for improved sanitation. In a representative sample of 30 villages in rural Cambodia, our NGO partner conducted group information sessions and sales meetings to market a low-cost, hygienic latrine. In 15 randomly selected villages, households were offered the option of financing their purchase with a loan from a local micro-finance organization. In the remaining 15 villages, sales were made on a standard lump-sum, cash-on-delivery basis. To maximize the precision of our estimates and broaden the set of research questions we are able to address, we used the Becker-DeGroot-Marschak (BDM) mechanism to obtain precise measures of WTP for each household (Becker et al. 1964).

We find that the offer of a micro-loan dramatically increases WTP. Mean WTP in the Financing arm is \$53.9, as compared to \$29.9 in the Non-financing arm. This large increase in WTP occurs across quantiles and among both the poorest and relatively less-poor households.[1] At the approximate break-even price of \$40, only 25% of households agree to purchase the latrine without financing, but more than 60% of households offered financing are willing to pay this full, unsubsidized price. Because NGOs and social enterprises face large village-level fixed costs in marketing and delivery, these increases in WTP can increase the cost-effectiveness of interventions, even net of the cost of providing financing, by amortizing these fixed costs over a greater number of sales.

The impact of finance on latrine coverage, while positive, is not as large as the effect on WTP. Latrine installation rates 1.5-2 years after the sale were low in both arms, muting the effect of finance. The primary barrier to installation was the high cost of households' desired latrine superstructure (walls and roof), for which financing was not offered. A secondary

[1] 25^{th} percentile: \$34.2 with financing vs. \$20 without; median: \$50.8 with financing vs. \$30 without; 75^{th} percentile: \$76.4 with financing vs. \$40 without. We discuss results by poverty subgroup in Section 3.3.

barrier, which we identify using the quasi-random variation in latrine purchases produced by the BDM mechanism, is a negative social spillover: exogenous increases in neighbors' purchases lead to lower installation rates and, ultimately, latrine coverage. This is consistent with several potential mechanisms: shared use of the latrines creating negative strategic complementarities in latrine installation – i.e. the private return to installation is decreasing in neighbors' latrine ownership – as well as with strategic substitutability in health investments.

This paper provides a model for the integration of "decision-focused evaluations" with academic analysis of general-interest questions. A decision-focused evaluation is one demanded by an implementer, designed to rapidly inform a specific policy or programmatic decision, and carried out in the specific context of interest and within in the implementer's usual operating and decision-making structures (Shah et al. 2015).[2] The randomized evaluation in this paper was designed to inform our implementation partner's decision of whether it should scale up microfinance loans for latrines in rural Cambodia. The evaluation took 3.5 months to complete from inception to reporting results, cost less than 60,000 USD, and provided a clear, actionable recommendation to the implementer. Given the decision-focus of this evaluation, the survey instruments used were designed to to minimize the time and cost of the evaluation, focusing only on collecting data that would be necessary for guiding iDE's decision. While this focus limits our ability to investigate mechanisms behind the effects we observe, we are still able to provide rigorous evidence on an important question of broad interest.

The paper is organized as follows: in Section 2, we describe our setting and experimental design, including our study sample, intervention, and data collection. In Section 3, we show the dramatic effect of finance on demand. In Section 4, we show that the effect on coverage was less and explore reasons for this gap. In Section 5, we discuss the results, including the implications for cost-effectiveness, and conclude.

[2]See Shah et al. (2015) for an extended discussion of the difference between decision-focused evaluations and "knowledge-focused evaluations," those designed by researchers to address academic questions.

2 Experimental Setting and Design

We conducted our experiment in rural Kampong Thom province, Cambodia, a region with generally low access to improved sanitation: as of 2012, 31% of rural residents had access to a hygienic latrine. At baseline, 71% of respondents in our sample primarily defecate in the open, with 90% of children under 5 in these households doing so. These children experience diarrhea regularly, with 39% of children in sample households having diarrhea in the 7 days preceding the survey.

Formal financing is not uncommon, with 41% of households having held a loan from a bank or microfinance institutions in the preceding year. Consumer loans from formal sources are very rare; most loans are for productive assets only. Informal credit is more prevalent (61% of households), but average loan holdings are relatively small.

We partnered with two institutions to implement the sanitation marketing and the microfinance loans. iDE Cambodia (iDE) has conducted sanitation marketing and training of sanitation suppliers in rural Cambodia since 2007 and currently is active in sanitation in eight provinces across Cambodia. Microloans were provided by VisionFund Cambodia (VFC), a microfinance institution established in 1994 that has served more than 140,000 clients with a total loan portfolio over US$37M.[3] iDE had worked in Kampong Thom province, but not in any of the sample villages, while VisionFund was active in all.

As a demand-driven, decision-focused evaluation, the study design was tailored to the implementers' budgetary, timeline, and operational constraints and the intervention mimicked the implementers' standard procedures as closely as possible. However, all data were collected by employees of IDinsight, independently of the implementers.

[3] In 2010, iDE received the World Toilet Organization Hall of Fame Award for its Sanitation Marketing program and resulting Easy Latrine. In 2011, VFC received the Platinum Award from MIX (Microfinance Information eXchange).

2.1 Sample Frame and Randomization

The study sample consists of 1,500 households from 30 villages in Kampong Thom province, selected using a multi-stage random sampling process. First, from a list of all villages in Kampong Thom, villages above the 95th percentile and below the 5th percentile with respect to population and share of households classified as "IDPoor"[4] were dropped to facilitate the selection of a representative sample. From the resulting list, samples of 30 villages were randomly selected without replacement and with a probability of selection weighted by village size. At the end of each draw, if the selected villages had significantly different sizes, poverty levels, or latrine coverage rates (measured at the province level) from the overall population averages then the draw was discarded. This procedure was repeated until 100 non-discarded samples were drawn. Of these 100 qualified samples, one was randomly selected as binding. From this sample of 30 villages, 15 were randomly assigned to receive the Financing treatment.

In each village, a census was taken to obtain the names of the head of household and spouse and IDPoor status of the household, and to identify whether each household owned a latrine. Fifty households were invited to participate in a group information session and sales meeting. These 50 households were randomly selected from the village population without a latrine, after stratifying on IDPoor status so that 30% of the selected households (i.e., 15 in each village) were classified as IDPoor. In four villages, fewer than 50 households did not have a latrine. In these four villages, all non-latrine-owning households were invited, yielding a total sample size of 1,383 households. On average, 76% of invited households attended the sales meeting. The household could be represented by any household member older than 18 with the authority to make large purchase decisions for that household. Field staff then followed up with non-attending households to conduct the same information and

[4]IDPoor 1 and 2 are the Government of Cambodia's official poverty lines: IDPoor 1 identifies "poorest or destitute" households; IDPoor 2 identifies "poor" households. IDPoor status is initially assigned based on a set of poverty indicators, then adjusted based on a consultation with village representatives to identify households with "special circumstances" (Ministry of Planning 2008). "IDPoor" includes both IDPoor 1 and IDPoor 2.

sales session, typically within one day of the initial session in the village. Ultimately, 1,380 of 1,383 invited households (99.8%) participated.

As described below, the intervention involved randomization of latrine prices at the household level. This randomization was unstratified and thus produced village-level variation in the average price draw.

2.2 Group Information Session

The core of the intervention was the group information session and sales meeting. Invited households gathered in a common location (e.g. school, village pagoda, village chief's house). Sales staff led a 45-60 minute interactive session emphasizing the health and convenience benefits of having a latrine and its status as an aspirational good. The latrine offered for sale included three concrete rings, each 80 cm in diameter, a concrete pan, a concrete slab with a porcelain bowl that fits into the pan, and a PVC pipe that connects the three rings to the concrete pan. The approximate cost of this set of parts was 160,000 KHR (USD 40).[5] Participants were not given information on the cost. The latrine was marketed as being easy to self-install, requiring approximately one day's labor to dig a 1.5 x 1m cylindrical pit to house the three concrete rings and a separate mound to keep the basin. No additional material was required to install the latrine after delivery other than a shovel and water to mix the mortar. The sale did not include installation or a superstructure.[6]

[5]The Cambodia Riel is pegged to the dollar. It fluctuates within a band of 3,900 KHR/USD to 4,100 KHR/USD. We use an exchange rate of 4,000 KHR/US$ throughout this document.

[6]Approximately one person-day of low-skilled labor was required for installation, which typically would cost no more than USD 5. The costs of a superstructure are often substantially higher than those of the latrine or its installation. There are three basic categories of superstructure: (a) bamboo / thatch, made from locally gathered materials; (b) tin, which can be self-constructed using materials typically costing less than $5; (c) concrete, which requires hiring a mason and costs at least USD 100. Options (a) and (b) typically require less than half a day's labor. Survey data from other villages in Kampong Thom finds an average cost of approximately USD 200, roughly 5 times the installed cost of a latrine itself. Anecdotally, households value both the durability of the concrete structure and the prestige it brings and are often reluctant to install a latrine without a fairly elaborate superstructure. This appears to have depressed installation rates in this study, as we discuss in Section 4.

2.3 Financing Treatment and Sale

At the end of the information session, attendees were offered an opportunity to purchase a latrine. The sales offer was made using the Becker-DeGroot-Marschak (BDM) mechanism (Becker et al. 1964). In the Non-Financing (cash-on-delivery) arm, households were commonly anchored on an initial price of 300,000 KHR (USD 75), with prices subsequently reduced in 10,000 KHR (USD 2.5) increments until the participant was willing to accept that price. Prices were then increased in smaller increments (2,000 KHR) until the participant was no longer willing to pay. This mechanism identified the maximum WTP, or "bid", for the latrine. The enumerator then allowed the household to choose from a set of sealed envelopes marked only with the participant's ID, each containing a randomly chosen price. The distribution of prices was 80,000 KHR, 120,000 KHR, 160,000 KHR, and 200,000 KHR (USD 20, 30, 40 and 50) with probability weights 1/7, 2/7, 2/7, 2/7, respectively.[7,8] If the price inside the chosen envelope was less than or equal to the subject's bid, then the subject purchased the latrine at the randomly determined price. If the price in the envelope was greater than the subject's bid, then the subject could not purchase the latrine. The subject was not allowed to change her bid after the price was revealed. For expected utility maximizers, the subject's best strategy is to bid her maximum WTP.[9] The bids were given

[7] The lowest price was down-weighted for budgetary reasons. Households were not informed of the distribution of prices. Although in principle the distribution of prices should not matter for an optimal bid, experimental data suggests it can (Urbancic 2011; Mazar et al. 2014). The distribution used here was chosen for simplicity.

[8] In the Financing arm, the prompts and price draws for the Financing arm were in terms of the first monthly payment. The initial anchor was 30,000 KHR (USD 7.5) per month. Prices were reduced by 2,000 KHR ($0.5) per month, then increased by 1,000 KHR (USD 0.25) per month. The price draws were 8,900 KHR, 13,400 KHR, 17,800 KHR, and 22,300 KHR, which represent total payments over the year approximately equal in NPV to the price draws in the Non-financing arm. As in the Non-financing arm, the lowest price received 1/7 probability weight and the others 2/7.

[9] See Shogren (2005) for a survey of applications of BDM in environmental economics. Horowitz (2006) points out that violations of expected utility can lead BDM bids to diverge from maximum WTP. In a study of WTP for ceramic water filters in Northern Ghana, Berry et al. (2015) find a gap of approximately USD 1 between WTP elicited through BDM and through take-it-or-leave-it (TIOLI) offers at randomized prices. For our application, we require (1) that the degree to which individuals' preferences differ from expected utility maximization is uncorrelated with the treatment, which in expectation will hold under randomization, and (2) that the effects of non-expected utility are similar for bidding behavior under the two treatments. Under these assumptions, while we might be skeptical of the *levels* of WTP reported by participants, the *difference* in reported WTP between treatments is still informative.

and the price revealed in private. All subjects also took part in a practice round in which participants bid for a token item (a box of cookies worth approximately KHR 2,000) to ensure they understood the procedure.

The study of credit constraints involved two treatment arms: Lump Sum (control) and Financing (treatment), which were randomly assigned at the village level (as discussed above). In Lump Sum villages, the household was required to pay the full agreed price upon delivery, which would occur within 10 days. In Financing villages, the household was offered a loan from VFC, which could be repaid over a term of up to 12 months. The loans were group liability with monthly interest rates of approximately 2.8%.[10] The sales mechanism was the same in Lump Sum and Financing villages, with two exceptions. First, in Financing villages, the loan option was explained before the bidding so that subjects could take this information into account when deciding their maximum WTP. Second, as indicated above, in Lump Sum villages, bids were made in terms of the full payment amount, while in Financing villages bids were made in terms of the monthly installment payments.[11] Finally, in Financing villages, typically within 24 hours of the group sales meeting, winning households met with a VFC underwriter, who used a basic battery of questions on the customer's age, income and assets to determine his eligibility for a loan. The underwriter then determined whether or not to extend the loan.[12]

To avoid non-random selection into attendance at the sales meeting, the treatment status

[10] Cambodian law prohibits flat interest rates on microfinance loans, so loans were "declining balance." That is, each month the client would pay 1/12 of the principal, plus accumulated interest on the outstanding principal. This is standard practice in VFC's local operations, and was familiar to most participants. The prices in the envelopes reflect the first month's payment, which would then decline over time.

[11] Because the monthly payments fall in declining balance loans, the first month's payment served as the bid price in the Financing arm. The field team took care to explain this feature to participants to minimize the possibility that participants who were unfamiliar with declining balances misinterpret the payment amount as constant over the 12-month term and potentially under-bidding relative to their true WTP. When we convert the stream of payments to net present value for comparability with bids in the lump-sum treatment, we account for this declining payment. We view this as conservative, since it would understate the perceived NPV of a household that did not understand that payments would decline.

[12] In retrospect, it likely would have been preferable to have this assessment occur before the household participated in BDM, so that the household's stated WTP would be constrained by the amount it could borrow. In our main results, we code WTP of rejected households without adjustment, and in Section 3.4 we provide robustness checks where we code rejected households' WTP as zero or 50% of their bid.

of villages was not announced until the meeting itself. In fact, in Finance villages, no mention was made of the possibility of finance until the sales meeting, and in Lump Sum villages, the possibility of a loan was not mentioned at all. Households who were invited to the group meeting but did not attend were visited by enumerators within one day and offered an opportunity to participate at home. If they agreed, the information session and sales exercise were conducted in similar fashion to the group meeting.[13]

2.4 Data Collection

We collected data using four instruments. First, we administered a census to all households in each village, obtaining the name of the head of household and spouse, whether the household owned a latrine, and whether the household was classified as IDPoor. Second, a baseline survey was conducted with all invited, consenting households. The survey covered latrine type and conditions, defecation practices, knowledge of latrine components and costs, household demographics, informal borrowing of rice or money from other people, housing quality, income sources, land holdings and agricultural production.

Third, we also obtained data from the sales exercise. For household h in village v, BDM provides data on willingness to pay (WTP_{hv}), the price offer (Draw_{hv}), and whether the household won the latrine ($\mathrm{Won}_{hv} = 1\{\mathrm{WTP}_{hv} \geq \mathrm{Draw}_{hv}\}$). We also record whether the household actually purchased the latrine (Bought_{hv}). If the household won the latrine but reneged, that is recorded as a refusal: $\mathrm{Refuse}_{hv} = (\mathrm{Won}_{hv} = 1) \bigcap (\mathrm{Bought}_{hv} = 0)$. For households that purchase the latrine, the price paid is the price offer, Draw_{hv}. The price paid is not defined for households that do not purchase. We also document whether households that lost attempted to bargain for the latrine and whether ex-post they wish they had bid more. For Financing villages, we convert to the declining balance sequence of payments to

[13] Because information travels quickly in villages, it is likely that at least some of these households learned if finance was available, so while selection into not attending the meeting is random relative to treatment, there is the possibility of non-random selection into agreeing to participate the next day. However, as noted in Section 2.1, only 3 of the 1,383 households ultimately declined to participate (0.2%, two Financing, one Non-financing), so the potential for bias is low. We code these three households as having zero WTP.

NPV using the interest rate of 2.8% per month as a proxy for VFC's cost of funds.

Finally, we conducted a follow-up survey to assess installation and use of the latrine via both self-reports and direct observation. The survey revisited all participating households approximately 18-24 months after the initial sales offers. Enumerators observed whether latrines (from any source) were installed, as well as conditions of installation, presence of a superstructure, and indicators of regular use and maintenance.

2.5 Summary Statistics

Table 1 presents baseline summary statistics and measures of balance for the sample. Column 1 presents means for the entire sample with standard deviations shown below in parentheses. Columns 2 and 3 present the means and standard deviations for Non-financing and Financing households, again with standard deviations shown in parentheses below each mean. Column 4 presents the difference in means between the Non-financing and Financing groups with standard errors presented in brackets below each difference. Finally, Column 5 presents the normalized difference $\left(\bar{X}_1 - \bar{X}_0\right)/\sqrt{(s_0^2 + s_1^2)}$ between the two means (Imbens and Wooldridge 2009).

81% of respondents are female, and nearly 50% live in a household with a child five years old or younger (25% live in household with child two years old or younger). By design, just under 30% of households are IDPoor, and mean household monthly income is just over $120. Many households had been exposed to microfinance prior to the study: 41% had taken out a loan from a formal source in the previous year.

Open defecation is extremely common among sample households: 70% of all individuals and 90% of children under the age of five primarily defecated in the open over the fifteen days preceding the survey. Despite this, households are clearly familiar with sanitation options: nearly 95% had previously considered purchasing a latrine in the past.

While the Financing and Non-financing groups appear generally well balanced in terms of baseline characteristics, there are a few significant differences. Episodes of diarrhea over the week preceding the survey are significantly lower in the Financing group relative to the control group, both among all individuals and among children age five and younger. The BDM latrine offer price is, on average, \$2 higher for Financing households than non financing households. To the extent that these baseline characteristics are predictive of successful latrine purchase through the BDM procedure, we would expect them to bias our results against finding any significant impact of Financing; that is, higher latrine offer prices should decrease latrine purchases in the Financing group relative to the control group. Similarly, the lower frequency of diarrhoeal episodes in Financing households may suggest they have smaller expected health returns to improved sanitation. Section 3.4 explores whether our main results are sensitive to these baseline differences.

3 Effect of Finance on Demand

3.1 Main Results

We are interested in the effect of finance on the (inverse) demand curve, i.e., $s\left(\mathrm{WTP}_{hv} \geq p\right)$, the share of households willing to purchase at a set of prices $p_L, \ldots, p_H$. Figure 1a plots $s\left(\mathrm{WTP}_{hv} \geq p\right)$ at each price $p = \{5, 10, \ldots, 100\}$ separately by treatment group (Non-Financing and Financing). Households in both groups purchased latrines at relatively high rates ($> 80\%$) when the price is below 20 USD. Demand is quite elastic over prices in the [20,40] range (or 50%-100% of market prices), especially in the Non-Financing group where only 27.9% of households would purchase the latrine at 40 USD.

In Figure 1b, we plot the estimated treatment effect of finance, with 95% confidence intervals constructed using both randomization inference and standard regression methods. (We discuss randomization inference in Section 3.2 and in Appendix B.) By randomization, this estimated treatment effect is the observed difference between the two demand curves in Figure 1a. We weight each household equally but the effects are similar if we weight villages equally. We find large treatment effects that initially increase with price, reaching a maximum that exceeds 40 pp at 45 USD. Beyond this price, demand falls relatively quickly even in Financing villages, with treatment effects diminishing (although remaining greater than zero even at 100 USD). The lower bound for treatment effects exceeds 0 for all $p \geq 20$. At the approximate break-even or unsubsidized price of 40 USD, treatment effects are 33.8 pp (RI p-value = 0.00), more than doubling the share of households purchasing latrines relative to the Non-Financing group.

One reason that financing may lead to higher purchase rates may be that households expect a substantial likelihood of defaulting. However, using administrative data from VisionFund, we track repayment rates for the ensuing year after the initial offer and find 100% repayment of the latrine loans.

3.2 Estimation and Inference

By virtue of randomization, obtaining point estimates for treatment effects is simple: at each price p, we can take the difference in shares purchasing, i.e.,

$$\hat{\beta}(p) = s_F\left(\mathrm{WTP}_{hv_F} \geq p\right) - s_{NF}\left(\mathrm{WTP}_{hv_{NF}} \geq p\right). \quad (1)$$

Alternatively, we can use the estimated coefficient from a regression of $1\{\mathrm{WTP}_{hv} \geq p\}$ on an indicator for whether household h's village was in the Financing treatment group, plus a constant.

Computing p-values and confidence intervals is not as straightforward. Because the treatment was randomized at the village level, and determinants of demand are likely correlated within village, inference must be made robust to clustering within village. However, because we have only 30 villages, the standard cluster-robust regression standard errors may be unreliable (Cameron et al. 2008). Furthermore, the more robust small-G alternative of the wild bootstrap applies to linear models (Cameron and Miller 2015).

Instead, we employ randomization inference both to obtain p-values and to construct confidence intervals by inverting the relationship between effect and p-value. While randomization inference for p-values has grown in popularity in recent years,[14] accurate confidence intervals are important for policymakers because policy decisions usually depend on the magnitude of an effect, not just whether an effect exists, and use of randomization inference to construct confidence intervals is rare in economics.[15] The intuition for randomization inference dates back to Fisher (1935): the researcher specifies a sharp null hypothesis, imposes this null hypothesis on the data, generates the distribution of the test statistic of interest under this null over all or many combinations of treatment assignments, and then compares the actual, observed value of the test statistic to its generated distribution to estimate how extreme the observed value is when the null is true (Rosenbaum 2010; Imbens and Rubin 2015). This procedure provides p-values directly when the null hypothesis is one of no effect. To obtain confidence intervals of level α, we test a series of null hypotheses (in our application, effects of $-0.40, -0.39, \ldots, +0.80$) and construct the confidence interval as all values that are not rejected at the α level. We provide an extended discussion and formal exposition in Appendix B, and our code is available to interested researchers upon request.

For comparison, we also compare to "analytical" CIs computed using Huber-Eicker-White (HEW) heteroscedasticity-robust standard errors. We collapse the household-level data to the village-level share of households purchasing at price p, so, with one observation per village, clustering is not needed. However, this creates the possibility of small-sample bias in the HEW standard errors, so, following Angrist and Pischke (2009), we take the maximum of i.i.d, HEW, HC_2, HC_3, where HC_2 and HC_3 are finite-sample corrections to HEW standard errors. As shown in Figure 1b, the randomization inference confidence intervals are only slightly wider than the "analytical" CIs.

[14] For example, Bloom et al. (2006), Cohen and Dupas (2010), Bloom et al. (2013), and Gertler et al. (2014) obtain p-values from randomization inference, but construct confidence intervals using regression-based methods.

[15] Applications in statistics and political science include Ho and Imai (2006), Small et al. (2008), Hansen and Bowers (2009). To our knowledge, the only economics papers that use randomization inference to construct confidence intervals are Barrios et al. (2012) and Quistorff (2015), both observational studies.

3.3 Subgroup Analysis

One question of particular policy interest is whether the effect of finance differs by a household's initial poverty status.[16] It is reasonable to think that poorer households' WTP will be relatively more responsive to access to financing. Poorer households have fewer financial assets (e.g., savings that can be drawn against to pay for a large lump-sum expense) and fewer physical assets (either for liquidation or to serve as collateral), and they may already be at sufficiently low levels of consumption that cutting consumption to finance an investment in a durable good carries a high utility cost. On the other hand, perhaps even with the availability of finance, poor households may find other consumption needs to be higher priorities, whereas relatively better-off households may find that the financing option makes purchasing the latrine more attractive. Figure 2 presents estimated treatment effects by IDPoor subgroup, showing that WTP increased similarly in both subgroups. In fact, the demand curves in Non-Financing villages for poor and non-poor households are also remarkably similar. (See Figure A1 in the Appendix.) This is somewhat puzzling, since we expect sanitation to be a normal good. Our initial hypothesis was that IDPoor status may have been manipulated, or was a poor proxy for relative poverty within these communities. However, this does not appear to be the explanation: we constructed an alternative measure of household socio-economic status using asset measures from our baseline survey and found (a) that our measure is correlated with IDPoor status and (b) levels of demand and effects on demand were also generally similar which we used this alternative variable to classify households as rich or poor. One possible explanation is that the sample frame consisted of households without latrines at baseline, so the less-poor households are implicitly selected for having relatively low interest in sanitation in spite of their socio-economic status.

[16] This subgroup analysis was pre-specified as Hypothesis 2 in the Pre-Analysis Plan. As pre-specified, we group both IDPoor 1 ("poorest or destitute") and IDPoor 2 (merely "poor") households. Given the sampling strategy described in Section 2.1, in which 30% of households were IDPoor (but not stratified by IDPoor 1 vs. IDPoor 2), our precision for estimating effects on IDPoor 1 vs. 2 is limited. The point estimates indicate that finance increased demand among IDPoor 1 households more than among IDPoor 2 households (not reported, available on request).

3.4 Robustness Checks

3.4.1 Cancelled Orders and Loan Rejections

There were a small number of households who won a latrine during BDM but then cancelled their order upon latrine delivery. These "cancelling" households make up 2.5% of winners, but are disproportionately found in the Non-financing group: they represent 4.4% of winners in the Non-financing group and just 0.71% of winners the Financing group. One concern with respect to cancelling households is that their stated WTP in the BDM procedure may not represent their true WTP for the latrine. Given that these households end up electing to cancel the transaction at their stated WTP, it is possible that they overstated their true WTP. Similarly, 7.9% of winning households in the financing arm had their loan rejected by VFC, so their BDM bid likely overstated what they truly were willing and *able* to pay.

To test whether our results are robust to these considerations, Figure 3 re-estimates the differences in the share of households purchasing the latrines under different assumptions about the true WTP for households who end up cancelling their order and households who are not approved by VFC.

We begin by making the most extreme assumption possible: Figure 3a assumes that all households who cancelled their purchase or who were rejected by VFC for a loan have a true WTP of $0. While the difference in the share of households purchasing the latrine at prices between $20 and $100 appears slightly smaller, it remains statistically significant and quite large in magnitude. At the market price of $40, the estimated difference in the share of households purchasing the latrine is still 30 pp when assuming true WTP is zero for these households.

Figure 3b makes a less extreme assumption about true WTP for these households: that their true WTP is 50% of their stated WTP. Given that our results were largely unaffected when assuming true WTP was 0% of stated WTP, the estimated differences in the share of households purchasing the latrine at each price are unsurprisingly close to our main estimates in Figure 1b. At the market price, the estimated treatment effect of Financing drops to 31.5pp, just slightly lower than the 33.8pp difference we find when using stated WTP.

The vast majority of loan rejections occurred in two villages where approximately 25% of households were rejected, as compared to less than 3% of households in the other 13 financing villages. Anecdotally, it seems VFC had experienced high rates of partial or full default during previous lending activities in these villages and was especially cautious about extending loans. It does not appear that WTP was unusually high or low in these two villages. As shown in Figure 4, which displays our main results both when retaining the two high rejection rate villages and when dropping them from the sample, WTP in these two villages was not noticeably different from the other villages in the financing treatment.

3.4.2 Baseline Controls

A third potential concern with our main results is that they may be, in part, driven by the baseline differences shown in Table 1. To account for this possibility, Figure 5 displays the estimated effect of financing on the share of households purchasing the sanitary latrine at each $5 increment between $0 and $100, controlling for covariates. In Figure 5a, we control for the variables with significant differences in Table 1: BDM offer price indicators ($30,$40,$50) (with $20 being the omitted category), and the mean household number of diarrhoeal episodes experienced during the week prior to the survey.[17] In Figure 5b, we control for all baseline characteristics reported in Table 1. In both cases, including controls leads to only minor differences in the estimates, and the main result – economically important and statistically significant differences in demand between the financing and no financing groups at all prices between USD 20 and USD 100 – is unchanged.

[17]Note that in Table 1 there are significant mean differences both in the total number of cases of diarrhea and in the number of cases in children age five and younger, conditional on having such children in the household. Since less than half of the households in the sample have a child in this age range, we use just the first of these two variables as a control, but results are largely unchanged if we include a dummy for the presence of a child in the age range and code the number of such cases in households without a child in the age range a zero rather than missing.

4 Effects on Latrine Installation

The previous section provides strong evidence that finance increases demand for sanitation. However, increasing initial purchases of latrines is necessary but not sufficient for improvements in environmental quality: toilets cannot reduce fecal loads in the environment unless they are installed and used. We conducted a followup survey 18-22 months after the initial sale to measure installation rates and collect objective indicators of latrine maintenance and use. At followup, we found that roughly 30-40% of households that had purchased a latrine during the sales exercise had installed it.[18] Although this is comparable with rates of increased coverage from other sanitation interventions,[19] it is somewhat surprising that only 30-40% of households that had purchased a latrine (as opposed to simply being encouraged to install one or given one for free) would have installed it nearly two years later. In this section, we first describe installation rates as a function of WTP and Financing treatment, then provide evidence the cost of the superstructure as a barriers to installation, and finally explore the role of peer effects in depressing installation.

4.1 Finance, WTP, and Installation Rates

Figure 6a plots, at each indicated price, the installation rate among all households that purchased the latrine and had WTP greater than or equal to that price. That is, the figure answers the question, "If an NGO offered latrines for sale at a given price, what share of latrines purchased at that price would be installed?" The figure is restricted to prices where 10% or more of the households purchased a latrine ($50 or less for Non-financing; $85 or less for Financing). There are three salient messages from the figure. First, installation

[18]The latrine parts were not stamped with identifying marks, so we cannot say with certainty that an installed latrine is the same as the one purchased in our sale. Here, we code a winning household as having installed its latrine if any improved latrine is present. In our followup survey, we collected detailed information on the type of latrine installed, if any, and the results are almost identical if we restrict the definition of "installed" to pour flush latrines with concrete rings lining the pit, consistent with the type of latrine being offered in our sale and not generally available in the area otherwise.

[19]For example, Clasen et al. (2014), in an RCT studying the Total Sanitation Campaign in Orissa, India, report a program effect of +28 percentage points in the share of households with a functional latrine.

rates are below 40% except among the Non-Financing households with WTP above \$40. Second, installation rates at any given level of WTP are slightly higher among Non-Financing households than Financing, although these differences are not statistically significant (see Figure 6b). Third, there is little evidence of screening effects: installation rates are roughly constant as a function of WTP, although there are slight differences in the right tails of the WTP distribution: installation rates are slightly increasing among Non-Financing households with relatively high WTP and slightly decreasing among Financing households with relatively high WTP.[20]

Figure 6 shows installation rates conditional on purchase, but it is also useful to consider unconditional installation rates. That is, if latrines were offered for sale at a given price, what share of households would purchase and install a latrine? Because of the randomness of BDM, this unconditional installation rate has to be calculated indirectly. We would like to calculate

$$s\left(\text{Install}\right) = s\left(\text{Install} \mid \text{Purchase}\right) s\left(\text{Purchase}\right). \tag{2}$$

However, the purchase decision, $s\left(\text{Purchase}\right)$, has a random element since it depends in part on the BDM draw. To estimate the share that would purchase and install a latrine if the fixed price were p, we substitute $s\left(\text{Bid} \geq p\right)$ for $s\left(\text{Purchase}\right)$. That is, we use the installation rate that would have occurred among participants if a fixed price of p had been offered (or if everyone bidding p or more had won).

This synthetic installation rate is plotted in Figure 7a as a function of price, separately for Financing and Non-Financing villages. Figure 7a answers the question, "If an NGO offered latrines for sale at a given price, what share of households would purchase and install a latrine?" Note that the installation rate is now *higher* for Financing villages than Non-Financing at most prices, seemingly contradicting Figure 6. However, the two figures are in fact consistent – as shown in Figure 6, installation rates are higher in Non-Financing villages

[20]We supplement this descriptive analysis of screening effects with a more rigorous regression analysis in Appendix C.

holding WTP constant, but WTP is much higher in Financing villages (see Figure 1), so overall coverage rates are generally higher. This difference is statistically significant at all prices above $40, as shown in Figure 7b.

What explains these installation rates? First, the cost of the superstructure (walls, roof, and door) appears to have been an important barrier. Although is is possible to build an inexpensive superstructure using locally gathered materials (bamboo or thatch) or tin (typically less than $10), households exhibit a strong stated preference towards much more elaborate, expensive concrete structures costing $200 or more.[21] In informal discussions during the followup, households that had purchased a latrine but not yet installed it commonly stated that they intended to install the latrine eventually but had not yet saved enough for their desired superstructure. This suggests two approaches that might be effective at increasing installation rates: first, encouraging households to construct acceptable interim superstructures while they gather funds for the high-end superstructure they desire; second, by financing the superstructure in addition to the latrine itself. The former approach could be implemented by a subsidy, e.g., forgiving part of the loan if the latrine is installed. Cash-on-hand constraints do appear to matter to some extent: using the random variation in price paid generated by the BDM draw, we find that each $10 reduction in the price paid (holding WTP constant) by the household increases the probability of installation at followup by approximately 4 percentage points. See Appendix D for the details of this analysis.

4.2 Peer Effects and Installation

The second possible explanation for low installation comes from a negative peer effect. We can use the randomization embodied in the BDM mechanism to test for peer effects in installation. Peer effects are typically difficult to identify, because observed correlation in behavior could be the result of a true, causal peer effect or simply homophily: peers tend

[21] This aspiration for very high-end, expensive facilities resembles that described in Northern India by Coffey et al. (2014), who report that the minimally acceptable latrine described by their respondents would cost, on average, $350 to construct.

to have similar tastes, so we might expect their behaviors to be similar even absent a true, causal effect (Manski 1993). Because the price draw in BDM is random, there will be some random variation in latrine purchase rates in households' peer groups. Furthermore, since BDM provides data on the household's and peer households' WTP, we can control for these and use only random variation in purchase rates conditional on own and peer WTP.

We define each household's peer group as the other participating households in that household's village. Given the small size and rural nature of the villages in our sample, it seems likely that the resulting village groups are not far from the true peer groups relevant for latrine ownership, installation, and use.[22]

Our goal is to estimate the causal effect of latrine purchases by a household's peer group on that household's propensity to install a latrine. That is, we wish to estimate

$$1\left\{\text{Install}_{h,v}\right\} = \beta_0 + \beta_1 s\left(\text{Buy}_{\sim h,v}\right) + \varepsilon_{h,v}, \tag{3}$$

where $1\left\{\text{Install}_{h,v}\right\}$ is an indicator for whether household h in village v has installed a latrine at the time of the followup survey, and $s\left(\text{Buy}_{\sim h,v}\right)$ represents the share of other households (excluding household h) in village v who purchased a latrine. However, OLS estimation of Equation (3) may produce biased estimates of the causal parameter of interest β_1. The most plausible source of bias would be if households' preference for improved sanitation tended to be correlated within village, so in villages where WTP was greater and, therefore, $s\left(\text{Buy}_{\sim h,v}\right)$ tended to be higher, households were also more likely install a latrine they had purchased (i.e., $\varepsilon_{h,v}$ was higher on average).

To overcome this identification problem, we can use the randomness provided by the price draw in BDM. The simplest strategy would be to instrument for $s\left(\text{Buy}_{\sim h,v}\right)$ with $\overline{\text{Draw}}_{\sim h,v}$, the average price draw among household h's peer group. However, we can increase power by identifying the heterogeneous effects of price draws among households with varying WTP.

[22]Funding limitations prevented us from obtaining the detailed social network data necessary to construct more refined peer groups.

To do so, we focus on the randomly induced difference between actual latrine purchases and those predicted by households' WTP. Households with high WTP would have been predicted to purchase latrines under most price draws, while the opposite is true for low WTP households. The effects of price draws are thus conditional on WTP.

We can thus construct the household-specific unexpected purchases among its peer group, defined as $\tilde{s}\left(\text{Buy}_{\sim h,v}\right) = \sum_{j\sim h}\left[\left(\text{Buy}_{j,v}\right) - \left(\hat{\text{Buy}}_{j,v}\right)\right]$, where $\hat{\text{Buy}}_{j,v}$ is household j's probability of purchasing given its WTP. We use the theoretical distribution of price draws and the elicited WTP for each household to calculate their ex-ante probability of winning the latrine. As described in Section 2.3, latrine prices were drawn from a discrete distribution of \$20, \$30, \$40, \$50, with probabilities 1/7, 2/7, 2/7, 2/7. respectively.[23] Therefore, any household with WTP < 20 had a 0% chance to win the latrine, any household with $20 \leq$ WTP < 30 had a 1/7 chance to win, households with $30 \leq$ WTP < 40 had a 3/7 chance to win, households with $40 \leq$ WTP < 50 had a 5/7 chance to win, and those with $50 \leq$ WTP had a 100% chance to win. We can then calculate expected latrine ownership counts at the village level by summing these probabilities at the village level.

Table 2 displays summary statistics for unexpected latrine ownership, $\tilde{s}\left(\text{Buy}_{\sim h,v}\right)$, as well as the observed latrine installation rates at endline. On average, households are slightly less likely to have won a latrine through BDM than their WTP and the ex-ante distribution of potential draw prices would predict (mean and median of -0.01), and the range of observed values is -0.13 to 0.09.

Although we can simply estimate the effects of peers' unexpected ownership on latrine installation, we want to maximize precision by controlling for a household's own predictors of installation, including own latrine ownership. We do so by controlling for the randomly induced latrine ownership based on a household's own price draw and WTP. That is, we construct the own-household equivalents of our peer measures as $\hat{\text{Buy}}_{h,v} = \Pr\left(\text{Draw} \leq \text{Bid}_{h,v}\right)$ and $\tilde{\text{Buy}}_{h,v} = \text{Buy}_{h,v} - \hat{\text{Buy}}_{h,v}$. The latter term, $\tilde{\text{Buy}}_{h,v}$, represents the unexpected or random

[23] The draws in the Financing treatment were in terms of monthly payments corresponding to total payments with net present value equal to \$20, \$30, \$40, \$50.

component of latrine ownership.

Using these variables, we estimate

$$1\left\{\text{Install}_{h,v}\right\} = \beta_0 + \beta_1 \tilde{s}\left(\text{Buy}_{\sim h,v}\right) + \beta_2 \widetilde{\text{Buy}}_{h,v} + \gamma' x_{h,v} + \varepsilon_{h,v}, \tag{4}$$

by OLS, with results reported in Table 3. We condition on an indicator for the village Financing treatment, and standard errors are clustered at the village level.

As expected, own ownership has a positive and significant effect on the likelihood that a household installed a latrine; the point estimate suggests a 15 percentage point increase in the probability of latrine installation. The fact that the estimate is not 1 is indicative of the relatively low installation rates in our sample. More interestingly, the coefficient on the village leave out mean unexpected latrine share is negative and marginally significant ($p < .10$). This suggests that, after removing the effect of WTP and conditional on own ownership, having more neighbors purchase a latrine actually reduces the likelihood that a household installs a latrine. The point estimate indicates that shifting a household from a village where there are no unexpected latrines to a village where all other households have unexpected latrines reduces the likelihood the household will install a latrine by 70 percentage points. Or, scaling the point estimate to a more reasonable level, shifting a household from one village to another village with an unexpected latrine share that is one standard deviation (.045) higher is expected to reduce the likelihood that household installs a latrine by 3.2 percentage points.

We consider this a reduced form estimation in which we identify effects of unexpected peer ownership. One may naturally want to assess the effects of *any* peer ownership (rather than only unexpected purchases); that is, to estimate effects of $s\left(\text{Buy}_{\sim h,v}\right)$ rather than $\tilde{s}\left(\text{Buy}_{\sim h,v}\right)$. We thus also estimate an instrumental variables (IV) specification in which we instrument for $s\left(\text{Buy}_{\sim h,v}\right)$ with $\tilde{s}\left(\text{Buy}_{\sim h,v}\right)$ and $\text{Buy}_{h,v}$ with $\widetilde{\text{Buy}}_{h,v}$. The results, shown in Table 4, include point estimates that are precisely estimated and statistically indistinguishable from the OLS results. Once again, peer ownership dramatically reduces a household's installation of its own latrine. The first stage regressions show that unexpected peer ownership significantly predicts peer ownership (but, comfortingly, not own ownership).

Finally, Table 5 replicates the OLS specification from Table 3 but includes interactions between the Financing indicator and own and peer unexpected latrine ownership. Neither of the interactions are statistically significantly different from zero.

5 Discussion and Conclusion

This paper shows that providing finance can dramatically increase willingness to pay for improved latrines. This finding provides evidence on the previously under-studied question of whether imperfect credit markets may be in part responsible for low WTP for environmental quality in developing countries (Greenstone and Jack 2015). Furthermore, while micro-credit to date has not been demonstrated to provide significant benefits in terms of increasing income for clients (Banerjee 2013; Banerjee et al. 2015), this paper suggests that there may be important other ways that micro-finance can enhance welfare.[24]

This initial investment – in this context, the purchase of a latrine – is a necessary but not sufficient condition for improving environmental conditions and health. We find that, despite the large increase in WTP for the initial investment, installation rates are low even 18-24 months after purchase. An interesting question for further research is why households exhibit preferences for expensive, elaborate superstructures and what strategies or policies might encourage completing the investment in such situations.

The dramatic increase in latrine sales in financing villages has important implications for cost-effectiveness. As with many interventions, a large share of costs are village-level fixed costs, such as the time of sales agents and their transportation costs to remote villages. Since sales agents made over four times as many sales per village meeting when loans were offered,

[24] Other examples include micro-finance (loans or savings) increasing demand for health goods (Devoto et al. 2012; Tarozzi et al. 2014; Guiteras et al. 2015a) and micro-savings increasing business growth (Dupas and Robinson 2013).

the fixed cost of their time and transportation was amortized over many more latrines sold. Of course, offering finance carries a cost, but because VisionFund was already operating in program villages, the marginal cost of managing and collecting on loans was low. Using conservative assumptions on the costs of sales and marketing, and the marginal cost of providing finance, we calculate that offering finance can reduce program costs per latrine sold by up to 70%. For example, a direct-sales intervention, in which a team of 8 full-time sales agents travel from village to village offering latrines at USD 50 without financing, would incur operational costs (sales and marketing) of USD 19 per latrine sold. Providing financing would add to operational costs by requiring a loan officer travel to process and collect loans (note we are making the conservative assumption that an additional MFI employee would be necessary, which may not be true if the MFI already has a robust presence), but even net of this financing cost, the increase in demand through financing reduces total operational costs per latrine sold to approximately 6 USD.[25]

Finally, this paper provides a model for learning from a "decision-focused evaluation," i.e., one designed to answer a specific programmatic question for an implementer rather than a general academic question. However, this does come at a cost, in that we are limited in our ability to explain mechanisms underlying the large effect of credit on demand. In particular, either credit constraints or impatience could explain this reduced-form result, and further research is needed to understand the relative contributions of each.

[25] Detailed assumptions and calculation available on request.

References

Joshua D. Angrist and Jörn-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, 2009.

Abhijit V. Banerjee. Microcredit Under the Microscope: What Have We Learned in the Past Two Decades, and What Do We Need to Know? *Annual Review of Economics*, 5(1): 487–519, August 2013. doi: 10.1146/annurev-economics-082912-110220.

Abhijit V. Banerjee, Dean Karlan, and Jonathan Zinman. Six Randomized Evaluations of Microcredit: Introduction and Further Steps. *American Economic Journal: Applied Economics*, 7(1):1–21, January 2015. doi: 10.1257/app.20140287.

Thomas Barrios, Rebecca Diamond, Guido W. Imbens, and Michal Kolesár. Clustering, Spatial Correlations, and Randomization Inference. *Journal of the American Statistical Association*, 107(498):578–591, June 2012. doi: 10.1080/01621459.2012.682524.

Gordon M. Becker, Morris H. DeGroot, and Jacob Marschak. Measuring utility by a single-response sequential method. *Behavioral Science*, 9:226–236, 1964. doi: 10.1002/bs.3830090304.

James Berry, Gregory M. Fischer, and Raymond P. Guiteras. Eliciting and Utilizing Willingness to Pay: Evidence from Field Trials in Northern Ghana. MPRC Population Working Paper PWP-MPRC-2015-017, December 2015.

Erik Bloom, Indu Bhushan, David Clingingsmith, Rathavuth Hong, Elizabeth King, Michael Kremer, Benjamin Loevinsohn, and J. Brad Schwartz. Contracting for health: Evidence from Cambodia. 2006.

N. Bloom, B. Eifert, A. Mahajan, D. McKenzie, and J. Roberts. Does Management Matter? Evidence from India. *The Quarterly Journal of Economics*, 128(1):1–51, February 2013. doi: 10.1093/qje/qjs044.

A. Colin Cameron and Douglas L. Miller. A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources*, 50(2):317–372, March 2015. doi: 10.3368/jhr.50.2.317.

A. Colin Cameron, Jonah B Gelbach, and Douglas L Miller. Bootstrap-Based Improvements for Inference with Clustered Errors. *Review of Economics and Statistics*, 90(3):414–427, 2008. doi: 10.1162/rest.90.3.414.

Thomas Clasen, Sophie Boisson, Parimita Routray, Belen Torondel, Melissa Bell, Oliver Cumming, Jeroen Ensink, Matthew Freeman, Marion Jenkins, Mitsunori Odagiri, Subhajyoti Ray, Antara Sinha, Mrutyunjay Suar, and Wolf-Peter Schmidt. Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: a cluster-randomised trial. *The Lancet Global Health*, 2(11):e645–e653, November 2014. doi: 10.1016/S2214-109X(14)70307-9.

Diane Coffey, Aashish Gupta, Payal Hathi, Nidhi Khurana, Dean Spears, Nikhil Srivastav, and Sangita Vyas. Revealed Preference for Open Defecation. *Economic and Political Weekly*, 49(38):43–55, September 2014.

Jessica Cohen and Pascaline Dupas. Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment. *Quarterly Journal of Economics*, 125(1): 1–45, February 2010. doi: 10.1162/qjec.2010.125.1.1.

Alan D. Dangour, Louise Watson, Oliver Cumming, Sophie Boisson, Yan Che, Yael Velleman, Sue Cavill, Elizabeth Allen, and Ricardo Uauy. Interventions to improve water quality and supply, sanitation and hygiene practices, and their effects on the nutritional status of children. In The Cochrane Collaboration, editor, *Cochrane Database of Systematic Reviews*. John Wiley & Sons, Ltd, Chichester, UK, August 2013.

Florencia Devoto, Esther Duflo, Pascaline Dupas, William Parienté, and Vincent Pons. Happiness on Tap: Piped Water Adoption in Urban Morocco. *American Economic Journal: Economic Policy*, 4(4):68–99, November 2012. doi: 10.1257/pol.4.4.68.

Pascaline Dupas and Jonathan Robinson. Savings Constraints and Microenterprise Development: Evidence from a Field Experiment in Kenya. *American Economic Journal: Applied Economics*, 5(1):163–192, January 2013. doi: 10.1257/app.5.1.163.

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

Paul Gertler, Manisha Shah, Maria Laura Alzua, Lisa Cameron, Sebastian Martinez, and Sumeet Patil. How Does Health Promotion Work? Evidence From The Dirty Business of Eliminating Open Defecation. Working Paper 20997, National Bureau of Economic Research, March 2015. URL `http://www.nber.org/papers/w20997`.

Paul J. Gertler, James Heckman, Rodrigo Pinto, Arianna Zanolini, Christel Vermeersch, Susan Walker, Susan M. Chang, and Sally Grantham-McGregor. Labor market returns to an early childhood stimulation intervention in Jamaica. *Science*, 344(6187):998–1001, May 2014. doi: 10.1126/science.1251178.

Michael Greenstone and B. Kelsey Jack. Envirodevonomics: A Research Agenda for an Emerging Field. *Journal of Economic Literature*, 53(1):5–42, March 2015. doi: 10.1257/jel.53.1.5.

Raymond P. Guiteras, David I. Levine, Thomas H. Polley, and Brian Quistorff. Credit Constraints, Discounting and Investment in Health: Evidence from Micropayments for Clean Water in Dhaka. Working paper, university of maryland, December 2015a.

Raymond P. Guiteras, James Levinsohn, and Ahmed Mushfiq Mobarak. Encouraging sanitation investment in the developing world: A cluster-randomized trial. *Science*, 348(6237): 903–906, May 2015b. doi: 10.1126/science.aaa0491.

Ben B. Hansen and Jake Bowers. Attributing Effects to a Cluster-Randomized Get-Out-the-Vote Campaign. *Journal of the American Statistical Association*, 104(487):873–885, September 2009. doi: 10.1198/jasa.2009.ap06589.

Daniel E. Ho and Kosuke Imai. Randomization Inference With Natural Experiments: An Analysis of Ballot Effects in the 2003 California Recall Election. *Journal of the American Statistical Association*, 101(475):888–900, September 2006. doi: 10.1198/016214505000001258.

John K. Horowitz. The Becker-DeGroot-Marschak mechanism is not necessarily incentive compatible, even for non-random goods. *Economics Letters*, 93(1):6–11, 2006.

Guido W. Imbens and Donald B. Rubin. *Causal Inference in Statistics, Social, and Biomedical Sciences An Introduction*. Cambridge University Press, 2015.

Guido W. Imbens and Jeffrey M. Wooldridge. Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature*, 47(1):5–86, February 2009. doi: 10.1257/jel.47.1.5.

Patrick Klein and Andres Santos. A Score-Based Approach to Wild Bootstrap Inference. *Journal of Econometric Methods*, 1(1):23–41, August 2012.

Audrie Lin, Benjamin F. Arnold, Sadia Afreen, Rie Goto, Tarique Mohammad Nurul Huda, Rashidul Haque, Rubhana Raqib, Leanne Unicomb, Tahmeed Ahmed, John M. Colford, and Stephen P. Luby. Household environmental conditions are associated with enteropathy and impaired growth in rural bangladesh. *The American Journal of Tropical Medicine and Hygiene*, 89(1):130–137, July 2013. doi: 10.4269/ajtmh.12-0629.

Charles F. Manski. Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies*, 60(3):531, July 1993. doi: 10.2307/2298123.

Nina Mazar, Botond Koszegi, and Dan Ariely. True Context-dependent Preferences? The Causes of Market-dependent Valuations. *Journal of Behavioral Decision Making*, 27(3): 200–208, July 2014. doi: 10.1002/bdm.1794.

Ministry of Planning. Implementation Manual on the Procedures for Identification of Poor Households, October 2008.

Subhrendu K. Pattanayak and Alexander Pfaff. Behavior, Environment, and Health in Developing Countries: Evaluation and Valuation. *Annual Review of Resource Economics*, 1(1):183–217, 2009. doi: 10.1146/annurev.resource.050708.144053.

Subhrendu K. Pattanayak, Jui-Chen Yang, Katherine L. Dickinson, Christine Poulos, Sumeet R. Patil, Ranjan K. Mallick, Jonathan L. Blitstein, and Purujit Praharaj. Shame or subsidy revisited: social mobilization for sanitation in Orissa, India. *Bulletin of the World Health Organization*, 87(8):580–587, August 2009. doi: 10.2471/BLT.08.057422.

Eduardo Perez, Jason Cardosi, Yolande Coombes, Jacqueline Devine, Amy Grossman, Craig Kullmann, C Ajith Kumar, Nilanjana Mukherjee, Manu Prakash, and Amin Robiarto. What Does It Take to Scale Up Rural Sanitation? Technical report, World Bank, 2012.

Annette Prüss-Üstün and Carlos Corvalán. *Preventing disease through healthy environments: Towards an estimate of the environmental burden of disease*. World Health Organization, Geneva, 2006.

Annette Prüss-Ustün, Jamie Bartram, Thomas Clasen, John M. Colford, Oliver Cumming, Valerie Curtis, Sophie Bonjour, Alan D. Dangour, Jennifer De France, Lorna Fewtrell, Matthew C. Freeman, Bruce Gordon, Paul R. Hunter, Richard B. Johnston, Colin Mathers, Daniel Mäusezahl, Kate Medlicott, Maria Neira, Meredith Stocks, Jennyfer Wolf, and Sandy Cairncross. Burden of disease from inadequate water, sanitation and hygiene in low- and middle-income settings: a retrospective analysis of data from 145 countries. *Tropical Medicine & International Health*, 19(8):894–905, August 2014. doi: 10.1111/tmi.12329.

Brian Quistorff. Capitalitis? Effects of the 1960 Brazilian Capital Relocation. *University of Maryland Working Paper*, March 2015. doi: 10.2139/ssrn.2588620.

Paul R. Rosenbaum. *Design of Observational Studies*. Springer Series in Statistics. Springer New York, New York, NY, 2010.

Neil Buddy Shah, Paul Wang, Andrew Fraker, and Daniel Gastfriend. Evaluations with impact: decision-focused impact evaluation as a practical policymaking tool. *3ie Working Paper*, (25), 2015. URL `http://www.3ieimpact.org/media/filer_public/2015/10/01/wp25-evaluations_with_impact.pdf`.

Jason F. Shogren. Experimental Methods and Valuation. In Karl-Goran Maler and Jeffrey R. Vincent, editor, *Handbook of Environmental Economics*, volume Volume 2 of *Valuing Environmental Changes*, pages 969–1027. Elsevier, 2005.

Dylan S. Small, Thomas R. Ten Have, and Paul R. Rosenbaum. Randomization Inference in a Group–Randomized Trial of Treatments for Depression. *Journal of the American Statistical Association*, 103(481):271–279, March 2008. doi: 10.1198/016214507000000897.

Dean Spears. *How Much International Variation in Child Height Can Sanitation Explain?* Policy Research Working Papers. The World Bank, January 2013.

Dean Spears and Sneha Lamba. Effects of Early-Life Exposure to Sanitation on Childhood Cognitive Skills: Evidence from India's Total Sanitation Campaign. World Bank Policy Research Working Paper 6659, October 2013.

Alessandro Tarozzi, Aprajit Mahajan, Brian Blackburn, Dan Kopf, Lakshmi Krishnan, and Joanne Yoong. Micro-Loans, Insecticide-Treated Bednets, and Malaria: Evidence from a Randomized Controlled Trial in Orissa, India. *American Economic Review*, 104(7): 1909–1941, July 2014. doi: 10.1257/aer.104.7.1909.

Sophie Trémolet. Sanitation Markets: Using economics to improve the delivery of services along the sanitation value chain. Technical report, SHARE, London, UK, December 2012.

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

WHO. Exposure to ambient air pollution. Technical report, World Health Organization Global Health Observatory, 2014a.

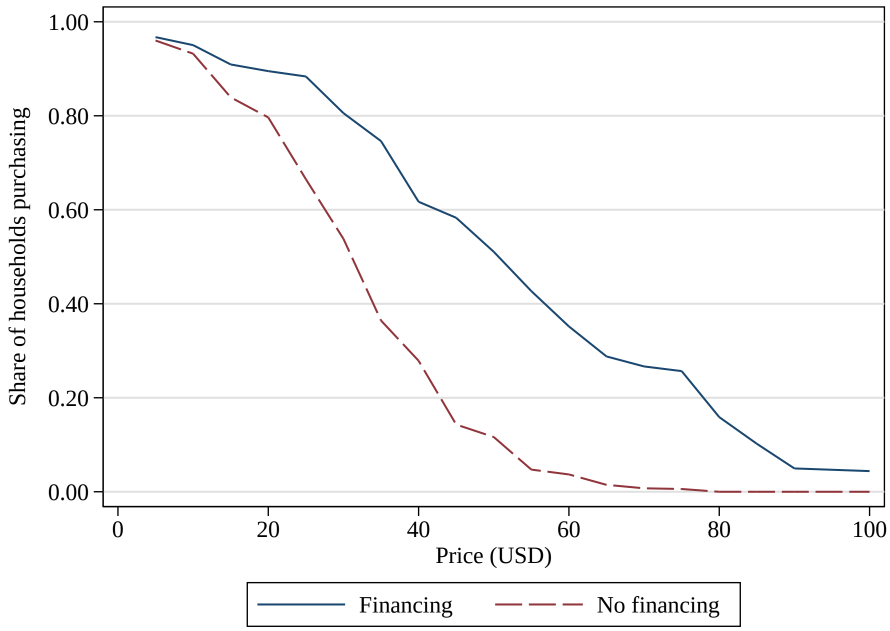
WHO. Exposure to household air pollution. Technical report, World Health Organization Global Health Observatory, 2014b.

WHO. Water, sanitation and hygiene. Technical report, World Health Organization Global Health Observatory, 2015.

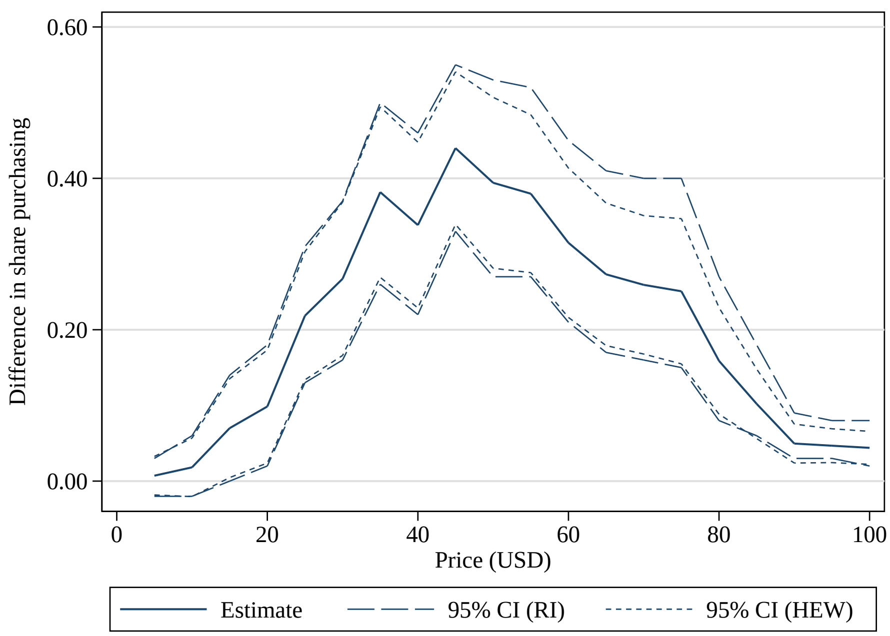
WHO and UNICEF. Progress on Drinking Water and Sanitation - 2014 Update. Technical report, 2014.

Figure 1: Levels of Demand and Effect of Finance

(a) Shares Purchasing by Treatment Arm



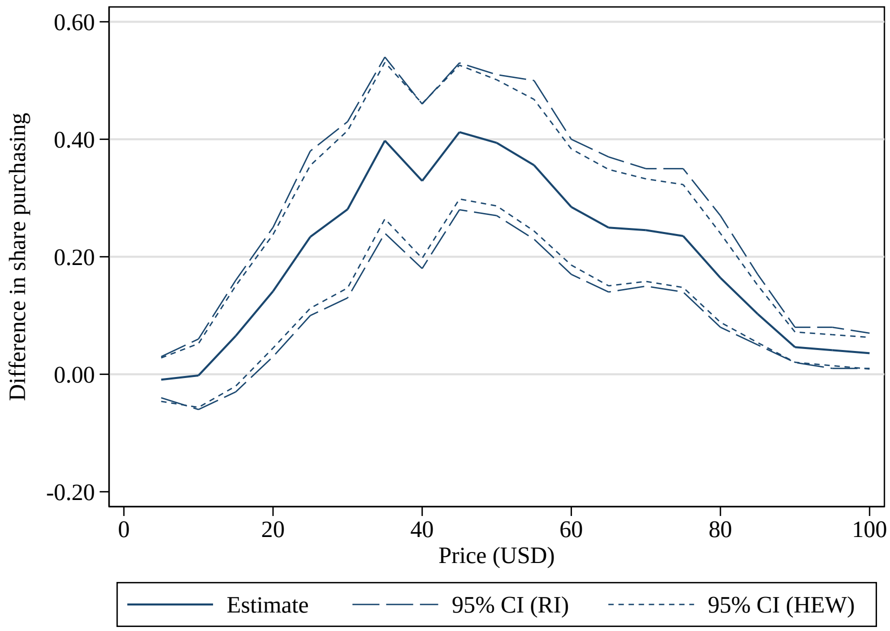
(b) Treatment Effect of Finance



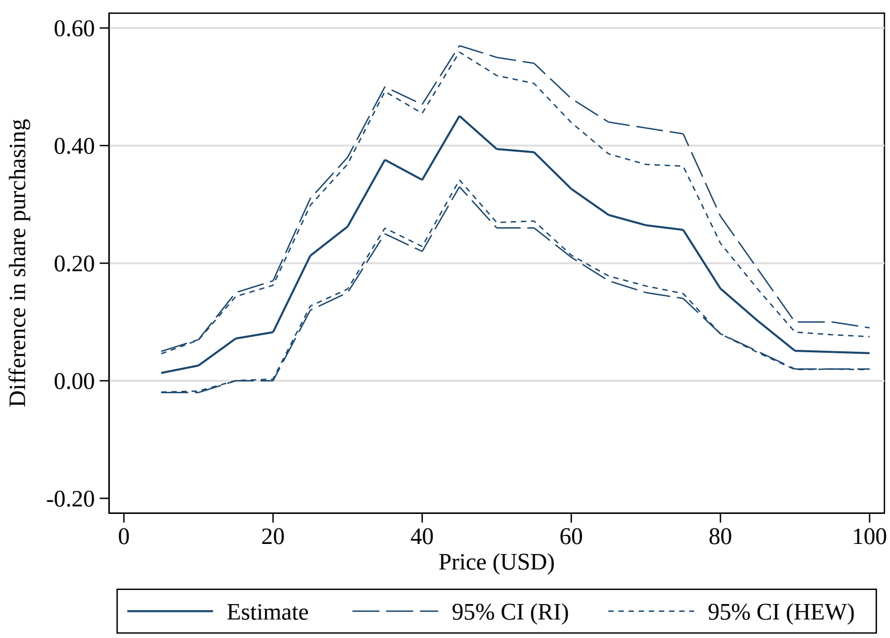
Notes: The top panel shows the shares of households with willingness to pay greater than or equal to each price, by Financing and Non-Financing treatment. The bottom panel shows the treatment effect of finance, i.e. the estimated difference between the Financing and Non-financing treatments in the share of households with willingness to pay greater than or equal to each price, with 95% confidence intervals. This difference is estimated at $5, 10, . . . , 100. The RI confidence interval (long dashes) is computed by randomization inference, while the HEW confidence interval is the usual heteroscedasticity-robust standard error. Observations are at the village level, with each village weighted by the number of participating households.

Figure 2: Effect of Finance – IDPoor vs. non-IDPoor

(a) Effect of Finance on IDPoor



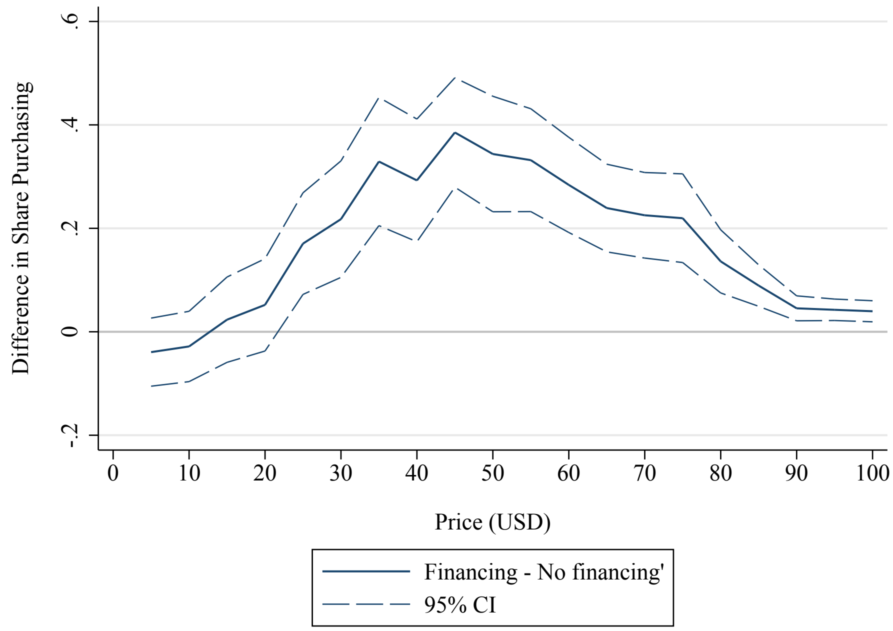
(b) Effect of Finance on non-IDPoor



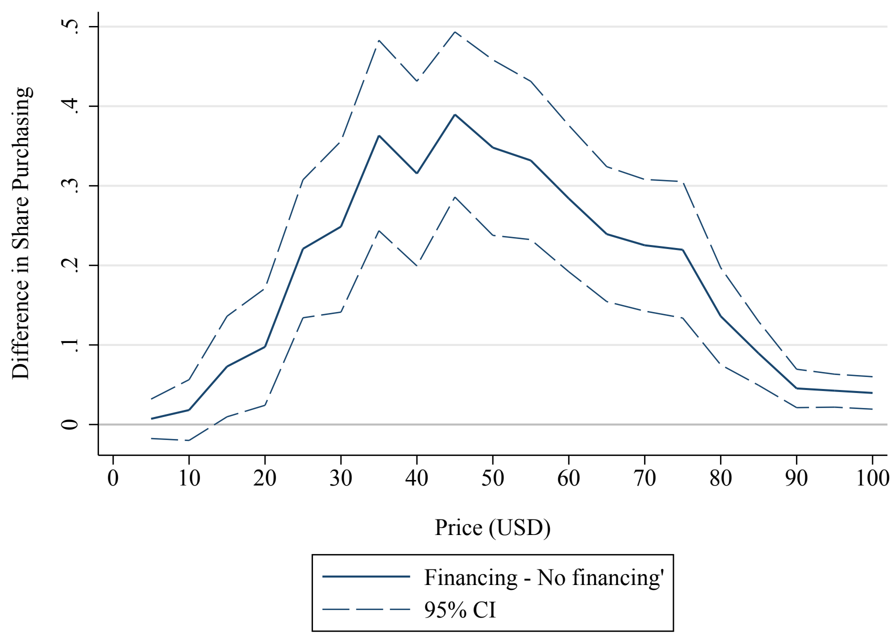
Notes: These figures show the estimated treatment effect of finance on IDPoor households (top panel) and non-IDPoor households (bottom panel). Each figure shows the estimated difference between the Financing and Non-financing treatments in the share of households with willingness to pay greater than or equal to each price, with 95% confidence intervals. This difference is estimated at \$5, 10, . . . , 100. The RI confidence interval (long dashes) is computed by randomization inference, while the HEW confidence interval is the usual heteroscedasticity-robust standard error. Observations are at the village level, with each village weighted by the number of participating households.

Figure 3: Different Assumptions about True WTP
Households Cancelling Order or Rejected for Loan

(a) Imputing WTP to 0



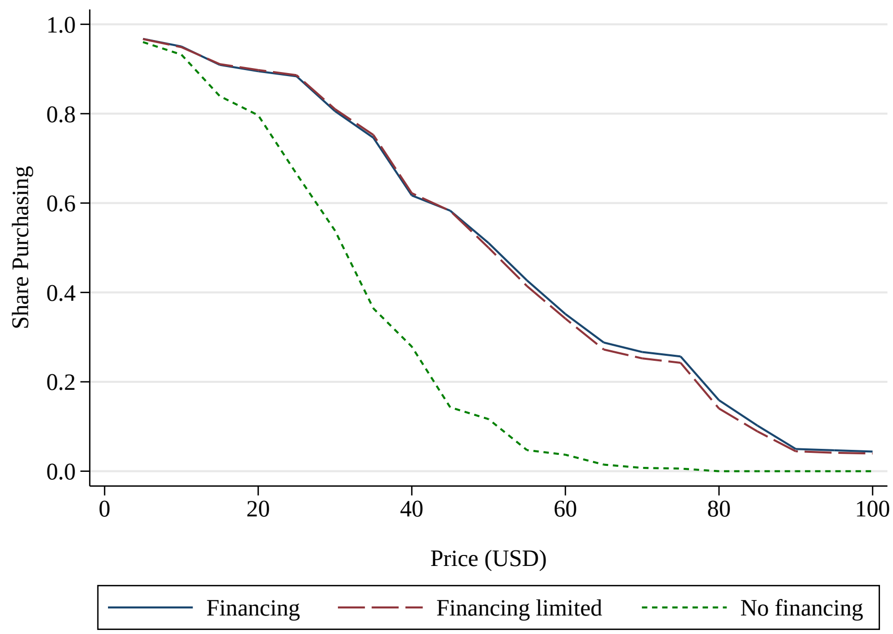
(b) Imputing WTP to 50% of Stated WTP



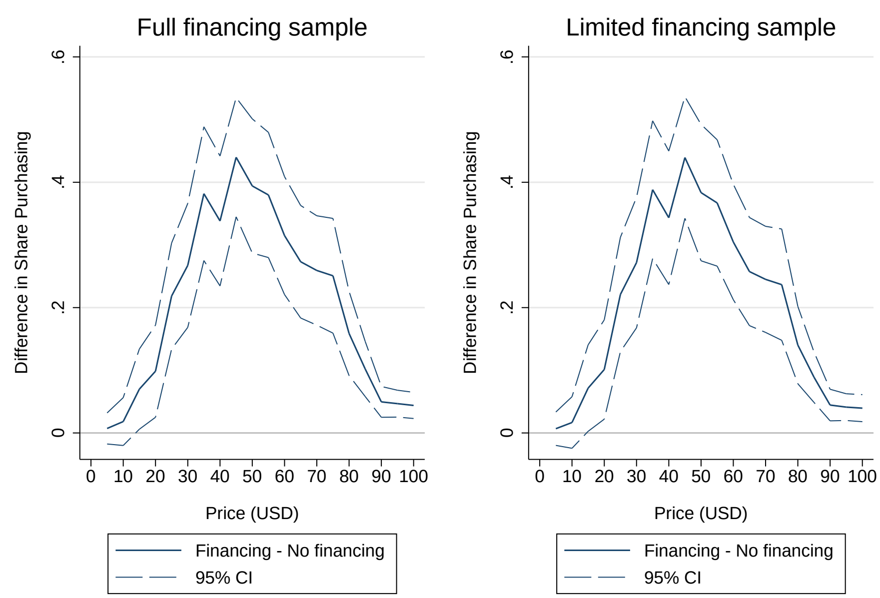
Notes: Figures show the robustness of the main results to two different assumptions about the true willingness to pay (WTP) of households who either cancelled their purchase upon latrine delivery or were not approved for a loan by the lender. The top panel assumes true WTP is 0 for all cancelling or rejected household while the bottom panel assumes true WTP is 50% of stated WTP. Both panels show the estimated difference between the Financing and Non-financing treatments in the share of households with WTP greater than or equal to each price, with 95% confidence intervals Confidence intervals are based on heteroskedasticity-robust standard errors from a village level regression weighted by the number of participating households.

Figure 4: Dropping Villages with High Loan Rejection Rates

(a) Shares Purchasing by Treatment Arm



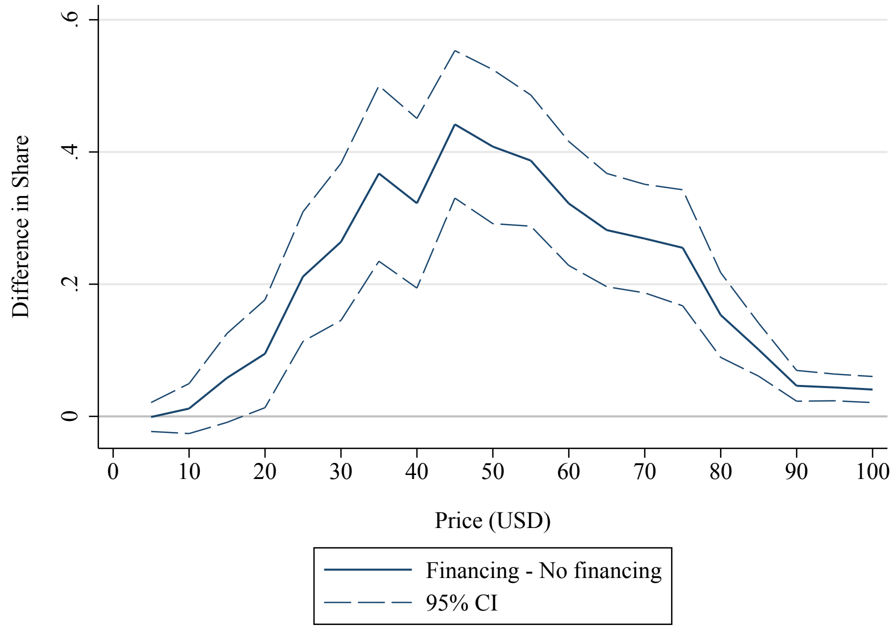
(b) Treatment Effect of Finance



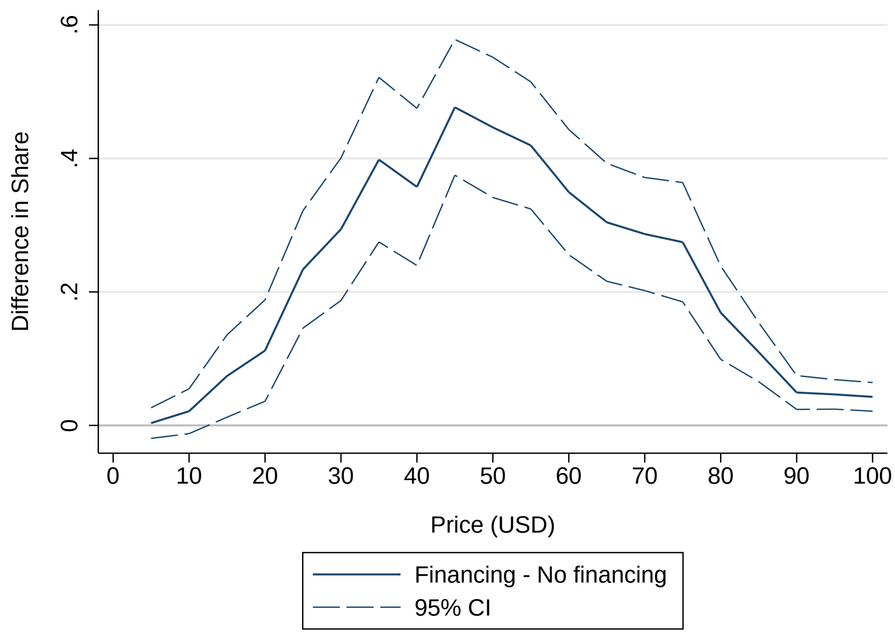
Notes: Figures show the robustness of the main results to dropping two villages in the Financing treatment group where loan rejection rates were high. The top panel displays the share of households with willingness to pay (WTP) greater than or equal to each price by treatment arm. Financing treatment shares are shown when including all villages (Financing) and when dropping the aforementioned villages (Financing limited). The bottom panel shows the estimated difference between the Financing and Non-financing treatments in the share of households with WTP greater than or equal to each price, with 95% confidence intervals when including (left) and not including (right) the high loan rejection rate villages. Confidence intervals are based on heteroskedasticity-robust standard errors from a village level regression weighted by the number of participating households.

Figure 5: Including Controls for Unbalanced Baseline Characteristics

(a) Control for Imbalanced Baseline Variables



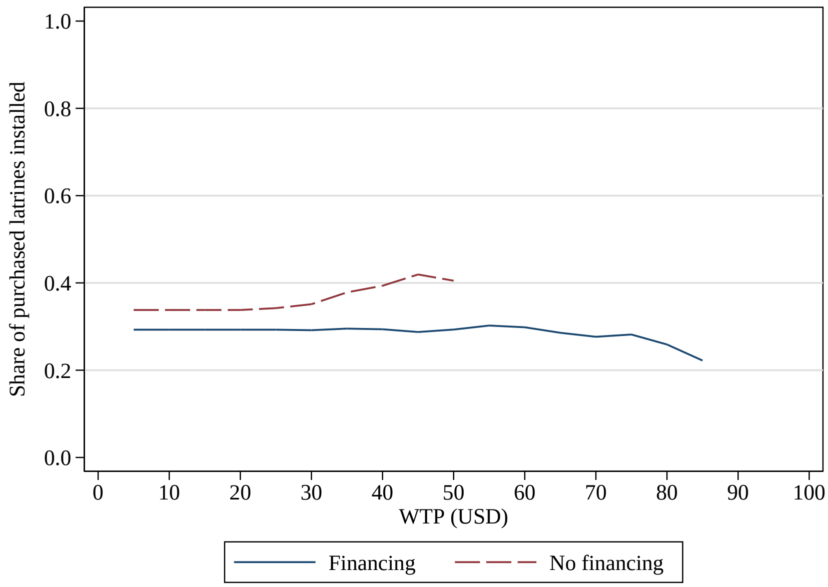
(b) All Baseline Controls



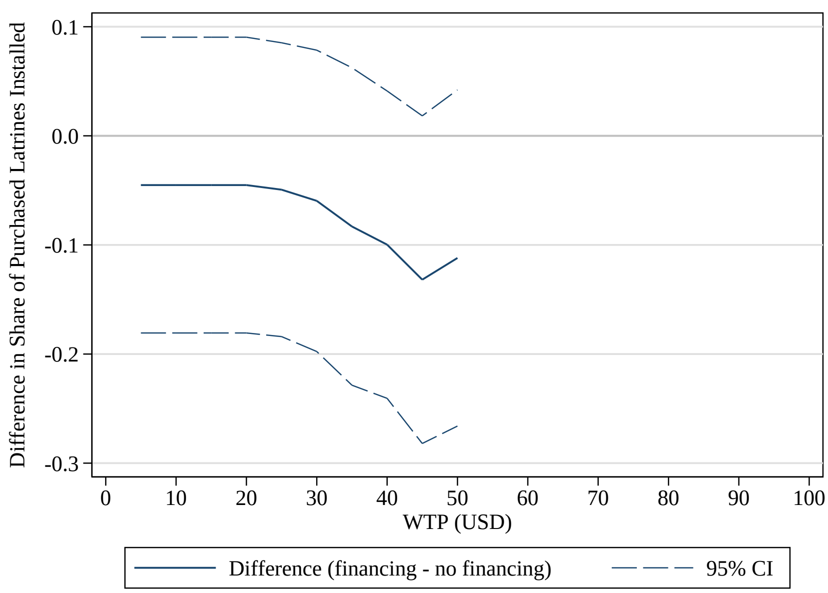
Notes: these figures show the robustness of the main results to controlling for baseline characteristics. Both panels present the coefficient and 95% confidence interval on the Financing dummy from an individual-level linear probability model where the dependent variable is an indicator for whether the household would purchase the latrine at each price. The top panel includes controls for variables from Table 1 that were imbalanced at baseline (BDM draw price dummy variables and the household average number of diarrhoeal episodes experienced over the week preceding the survey). The bottom panel controls for all baseline variables listed in Table 1. Confidence intervals are based standard errors clustered at the village level. Confidence intervals based on 999 cluster-robust bootstrap repetitions are nearly identical.

Figure 6: Latrine Installation Rates by WTP and Financing Treatment Conditional on Purchase

(a) Share of Purchased Latrines Installed



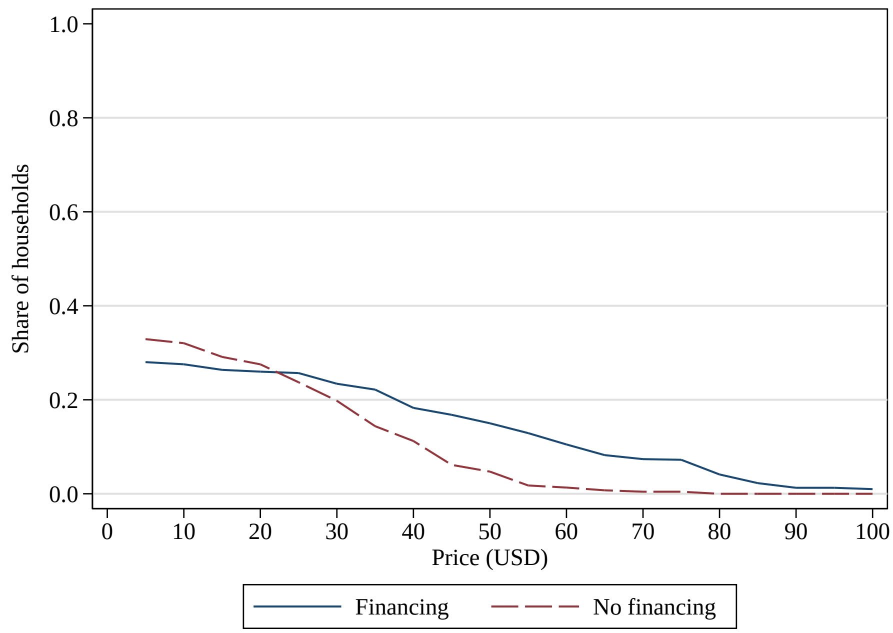
(b) Difference in Share of Latrines Installed – Financing vs. Non-Financing



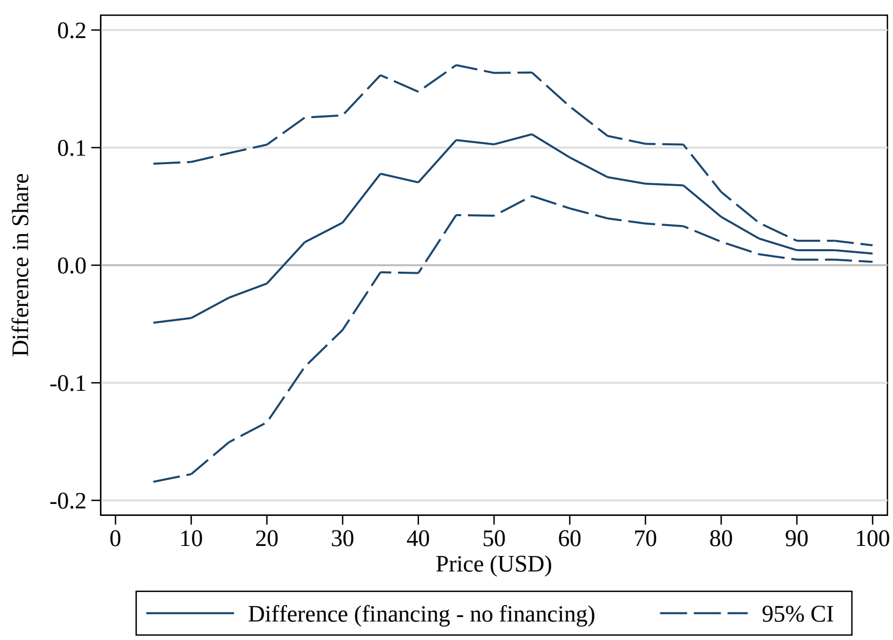
Notes: The top figure plots installation rates for households who purchase a latrine and have WTP greater than or equal to indicated price, separately by Financing and Non-financing treatment. The bottom figure plots the estimated difference in installation rates between households in the Financing and Non-financing treatments. The top figure is restricted to prices where 10% or more of the households in that treatment agreed to purchase a latrine (Non-financing: $50; Financing: $85), and the bottom figure is restricted to prices where 10% or more of the households in both treatments agreed to purchase a latrine ($50).

Figure 7: Latrine Installation Rates by Price and Financing Treatment
Unconditional

(a) Share of Households Purchasing and Installing Latrine



(b) Difference in Share of Households Purchasing and Installing Latrine



Notes: The top figure plots the share of households that, at each price, would purchase a latrine at that price and subsequently install it, separately by Financing and Non-financing treatment. The bottom panel plots the the estimated difference, with a 95% confidence interval, between households in the Financing and Non-financing treatments.

Table 1: Summary Statistics and Balance

	All households (1)	Non financing (2)	Financing (3)	Diff. (4)	Norm. diff. (5)
Household characteristics:					
Female respondent	0.811	0.806	0.816	0.010	0.018
	(0.391)	(0.396)	(0.388)	[0.027]	
Household Size	4.382	4.294	4.467	0.173	0.069
	(1.781)	(1.783)	(1.776)	[0.170]	
Number of women in household	2.243	2.178	2.305	0.126	0.079
	(1.139)	(1.125)	(1.149)	[0.090]	
Any children under age five	0.453	0.473	0.433	-0.040	-0.056
	(0.498)	(0.500)	(0.496)	[0.028]	
Any children under age two	0.252	0.253	0.251	-0.002	-0.003
	(0.434)	(0.435)	(0.434)	[0.028]	
# members who earn income	1.693	1.781	1.607	-0.175	-0.106
	(1.165)	(1.229)	(1.093)	[0.130]	
Total monthly household income (USD)	122.815	134.800	111.029	-23.772	-0.039
	(431.070)	(560.843)	(243.357)	[37.409]	
Household owns livestock	0.825	0.826	0.825	-0.001	-0.002
	(0.380)	(0.380)	(0.380)	[0.040]	
Household grows crops	0.833	0.865	0.806	-0.058	-0.111
	(0.373)	(0.342)	(0.396)	[0.054]	
Any formal loan in past year	0.414	0.405	0.423	0.018	0.026
	(0.493)	(0.491)	(0.494)	[0.051]	
Any informal loan in past year	0.655	0.643	0.667	0.024	0.036
	(0.475)	(0.480)	(0.472)	[0.032]	
Any current formal savings	0.017	0.019	0.015	-0.004	-0.021
	(0.131)	(0.137)	(0.123)	[0.009]	
Any current informal savings	0.841	0.795	0.886	0.091	0.178
	(0.366)	(0.404)	(0.318)	[0.069]	
ID poor household	0.275	0.273	0.277	0.004	0.006
	(0.447)	(0.446)	(0.448)	[0.031]	
Likelihood <2 USD a day	25.922	24.833	26.970	2.138	0.070
	(21.698)	(21.373)	(21.971)	[1.806]	
Primarily defecate in the open	0.703	0.688	0.718	0.030	0.046
	(0.457)	(0.464)	(0.450)	[0.074]	
# of diarrhoeal episodes	0.232	0.274	0.201	-0.072***	-0.180
	(0.285)	(0.296)	(0.273)	[0.020]	
Children <=5 defecate in the open	0.904	0.894	0.914	0.021	0.050
	(0.295)	(0.309)	(0.280)	[0.035]	
Children <=5 # of diarrhoeal episodes	0.390	0.472	0.323	-0.149***	-0.229
	(0.465)	(0.476)	(0.446)	[0.039]	
Has considered latrine purchase	0.946	0.942	0.950	0.008	0.025
	(0.226)	(0.234)	(0.218)	[0.014]	
Latrine offer price (USD)	35.764	34.734	36.752	2.018***	0.142
	(10.062)	(11.012)	(8.956)	[0.617]	

Notes: Table displays summary statistics for the whole sample (Column 1) and by treatment arm (Columns 2 and 3). Column 4 displays the difference between the mean in the Non-financing and Financing arms while Column 5 displays the normalized difference between the two means $\left(\bar{X}_1 - \bar{X}_0\right) / \sqrt{(s_0^2 + s_1^2)}$. Standard deviations appear in parentheses while standard errors appear in brackets. *** p<=0.01, ** p<=0.05, * p<=0.10. Child age cutoffs are inclusive of the cutoff age. The number of individuals who contribute income and total household income include non-resident members. The likelihood that the household lives on less than $2 per day is defined using the 2011 Progress out of Poverty Index (PPI). For variables in the PPI that were not included in the baseline survey we impute the mean value from the 2008 census in Cambodia for rural households in Kampong Thom province.

Table 2: Unexpected latrine ownership and latrine installation
Summary statistics

	mean	sd	min	p50	max
Fraction of HH with Unexpected Latrines	-0.01	0.04	-0.13	-0.01	0.09
Fraction HH installed latrine	0.22	0.13	0.00	0.22	0.54

Notes: Summary statistics for full sample of 1,363 households. *HH with unexpected latrines* is the difference between a household's actual latrine ownership status and its expected ownership, calculated based on the probability that its price draw falls below its stated WTP. *Fraction HH installed latrines* is an indicator of installation of an improved pit latrine at follow-up.

Table 3: Unexpected Village Latrine Ownership and Household Installation

	Installation Rate (OLS)
HH Won - Probability Win	0.159***
	(0.036)
Village fraction Won-fraction predicted Win	-0.702*
	(0.406)
Constant	0.198***
	(0.036)
Observations	1363

Notes: *** denotes $p < 0.01$, ** $p < 0.05$, * $p < 0.01$. Estimates reflect full sample of 1,363 households. *HH Won - Probability Win* is the difference between a household's actual latrine ownership status and its expected ownership, calculated based on the probability that its price draw falls below its stated WTP. *Village fraction won - fraction predicted win* is the mean of *HH Won - Probability Win* across a respondent's village, leaving out the respondent's own values. Standard errors clustered by village.

Table 4: Unexpected Village Latrine Ownership and Household Installation: IV

	First Stage Own	First Stage Village	Installation Rate
HH Won - Probability Win	1.013***	-0.004	
	(0.016)	(0.004)	
Village fraction Won-fraction predicted Win	-0.175	0.820***	
	(0.153)	(0.155)	
Household won latrine			0.147***
			(0.036)
Village fraction HH Won			-1.084*
			(0.562)
Constant	0.099	0.096	0.186
	(0.063)	(0.062)	(0.124)
First Stage Wald rK F-stat			13.325

Notes: *** denotes $p < 0.01$, ** $p < 0.05$, * $p < 0.01$. Estimates reflect full sample of 1,363 households. *HH Won - Probability Win* is the difference between a household's actual latrine ownership status and its expected ownership, calculated based on the probability that its price draw falls below its stated WTP. *Village fraction won - fraction predicted win* is the mean of *HH Won - Probability Win* across a respondent's village, leaving out the respondent's own values. *Household won latrine* denotes whether the household purchased a latrine through the BDM sales offer. *Village fraction HH won* is the mean of the *Household won latrine* across a respondent's village, leaving out the respondent's own value. Standard errors clustered by village.

Table 5: Unexpected Village Latrine Ownership and Household Installation: Financing Interaction

	Installation Rate (OLS)
HH Won - Probability Win	0.149***
	(0.045)
(HH Won - Probability Win)*Financing	0.029
	(0.073)
Village fraction Won-fraction predicted Win	-0.619
	(0.498)
Village fraction Won-fraction predicted Win*Financing	-0.245
	(0.814)
Constant	0.198***
	(0.035)
Observations	1363

Notes: *** denotes $p < 0.01$, ** $p < 0.05$, * $p < 0.01$. Estimates reflect full sample of 1,363 households. *HH Won - Probability Win* is the difference between a household's actual latrine ownership status and its expected ownership, calculated based on the probability that its price draw falls below its stated WTP. *Village fraction won - fraction predicted win* is the mean of *HH Won - Probability Won* across a respondent's village, leaving out the respondent's own values. *Financing* is an indicator of Financing treatment status at the village level. Standard errors clustered by village.

A Comparison with Pre-Analysis Plan

In this paper's discussion of effects on WTP, we emphasize impacts on the demand (the probability of purchasing at a set of given prices), rather than the effects on quantiles of WTP, as emphasized in the pre-analysis plan (PAP). These two objects are essentially isomorphic, since $q_{\mathrm{WTP}}(\tau)$ is defined by the relation $\Pr(\mathrm{WTP} \leq q_{\mathrm{WTP}}(\tau)) = \tau$. Since the two contain essentially the same information, we chose to present effects on demand rather than on quantiles for transparency and ease of exposition. Quantile effects are available from the authors upon request.

Second, the PAP's description of estimation of effects on demand uses logit estimation (see Equations (2) and (3) in that document), while this paper uses simple differences in proportions, as described in Section 3.1. Since our main estimates are simple comparisons by randomized treatment, the estimates are identical, and we found our randomization inference procedure simpler to implement than the available alternatives for small-G cluster-robust estimates for nonlinear models (e.g., Klein and Santos 2012).

Third, the overall low installation rates precludes the fully nonparametric analysis screening and sunk-cost effects described under Hypotheses 3 and 4, in particular Equations (8) and (9) in the PAP. In Section 4, we provide descriptive evidence, and in Appendices C and D we pursue the simpler regression strategy proposed as Equation (7) of the PAP.

Fourth, our funding was only sufficient to conduct basic followup surveys on installation and use, so we do not have the detailed measures on within-village peer groups that would be needed to investigate Hypotheses 5 and 6 of the PAP. Our analysis in Section 4.2 was the closest we could come to the analysis specified in the pre-analysis plan given our available resources and data.

B Randomization Inference

For simplicity, we will refer to Financing villages as treatment villages ($T_v = 1$) and Non-Financing villages ($T_v = 0$) as controls. We begin by aggregating our WTP data to the village level, obtaining an outcome variable $y_v(p)$ that reflects the share of households in village v who would purchase the latrine at price p. Using this village-price level dataset, we implement the following procedure:

Given a price p, we want to compute a $1 - \alpha$ confidence interval for the treatment effect at that price. Let

$$\begin{aligned} \hat{\beta}(p) &= \overline{y}_1 - \overline{y}_0 \\ &= \sum_{v:T_v=1} w_v y_v(p) - \sum_{v:T_v=0} w_v y_v(p) \end{aligned} \tag{5}$$

be the observed treatment effect, where $\overline{y}_1$ represents the mean share of households in treatment villages purchasing at the given price, and $\overline{y}_0$ is the share in control villages. (w_v indicates that we are weighting each village by the number of households.) We test this observed treatment effect against a set of sharp null hypotheses $\{H_0^m : \beta_0 = m\}, m = -0.40, -0.39, \ldots, +0.79, +0.80$. To obtain confidence intervals, we implement these sharp nulls and use the distribution of placebo treatment effects under these nulls, as follows.

For each value m of the null hypothesis, we generate potential outcomes in the treated and untreated state under the null $\beta_0 = m$. Potential outcomes when treated, y_{v1}, are equal to the observed share of households purchasing at the given price in treatment villages (i.e., $y_{v1} = s_v$ if $T_v = 1$). For control villages, this is the observed share plus the value of the null ($y_{v1} = s_v + m$ if $T_v = 0$). Potential outcomes when not treated, y_{v0}, are the opposite of the above: in treatment villages, the potential outcome is the observed share minus the value of the null ($y_{v0} = s_v - m$ if $T_v = 1$); in control villages, the potential outcome is the observed share ($y_{v0} = s_v$ if $T_v = 0$). See Tables A1-A2 for an illustrative example.

Next, we implement 100,000 repetitions of the following placebo experiment.[26] In each repetition r, we generate a random treatment assignment vector v_r, in which we randomly assign 15 villages to $T_v^r = 1$ and 15 villages to $T_v^r = 0$. We then assign each village the potential outcome corresponding to its (placebo) random assignment. That is, if village i is randomized to $T_v^r = 1$, it is assigned outcome $y_v^r = y_{v1}$, and if it is randomized to $y_v^r = y_{v0}$, it receives $T_v^r = 0$. We then compute the observed treatment effect for the repetition as $\hat{\beta}^r = \overline{y}_1^r - \overline{y}_0^r$ (as in Equation 5 above), the difference in the mean shares of households purchasing across the randomly assigned placebo treatment and control. Table A3 continues the illustrative example of Tables A1-A2.

We then collect the observed treatment effects $\left\{\hat{\beta}^r\right\} = \left\{\hat{\beta}^1, \ldots, \hat{\beta}^R\right\}$ over all repetitions. This simulates the distribution of estimates we would expect to see if the null hypothesis were true, since it is true in our simulation. We then assess the plausibility of this null hypothesis by observing where the actual observed value $\hat{\beta}$ falls in this distribution. For a two-sided test of the hypothesis $\beta_0 = m$, we calculate the share of repetitions r with $\left|\hat{\beta}^r\right| > |m|$. This gives us an empirical p-value for the null $H_0^m : \beta_0 = m$.

Now for each H_0^m, we have an associated p-value $p(m)$, i.e. $p(-0.40)$, $p(-0.39)$, ..., $p(-0.79)$, $p(-0.80)$. The $1-\alpha$ confidence interval for $\hat{\beta}$ is the set of m with p-values greater than $\alpha/2$. That is, the lower bound of the CI is $m_{\mathrm{LB}} = \min\{m : p(m) > \alpha/2\}$ and the upper bound is $m_{\mathrm{UB}} = \max\{m : p(m) > \alpha/2\}$. This is the set of values for β that are not rejected with 95% confidence. Figure A2 illustrates the method.

C Screening Effect of WTP

In this section, we provide additional evidence on the relationship between WTP and installation that we describe in Section 4.1. We begin with a semiparametric approach in which

[26]Computing true exact p-values would require enumerating all $\binom{30}{15} \approx 1.551 \cdot 10^8$ possible assignment vectors. With R random draws, the maximum standard error of the estimated p-value is $1/2\sqrt{R}$ (Imbens and Rubin 2015), so with $R = 100,000$, the maximum standard error is 0.0016, which is tolerable.

we categorize the sample of BDM winners into bins by WTP. First, we define a set of five WTP bins

$$\mathbb{B} = \{\text{WTP} \leq 20,\ 20 < \text{WTP} \leq 30,\ 30 < \text{WTP} \leq 40,\ 40 < \text{WTP} \leq 50,\ \ \text{WTP} > 50\}$$

and the set of four offer price indicators

$$\mathbb{P} = \{20,\ 30,\ 40,\ 50\}$$

corresponding to the four possible BDM draws. We estimate a linear probability model

$$I_{hv} = \sum_{b \in \mathbb{B}} \alpha_b 1\left\{\text{WTP}_{hv} \in \text{bin}_b\right\} + \sum_{p \in \mathbb{P}} \beta_p 1\left\{\text{Draw}_{hv} = p\right\} + \theta T_v + \gamma' X_{hv} + \varepsilon_{hv}, \tag{6}$$

where I_{hv} is an indicator for whether household h in village v had installed its latrine by endline (18-22 months after the purchase), $1\left\{\text{WTP}_{hv} \in \text{bin}_b\right\}$ is an indicator for the household's BDM bid falling in the corresponding bin, $1\left\{\text{Draw}_{hv} = p\right\}$ is an indicator for the household's BDM draw, T_v is an indicator for the treatment status of the household's village, X_{hv} is a set of baseline controls, and ε_{hv} is an error term. We omit the indicator for whether the household has a WTP $>$ \$50. The coefficients of interest are the α_b, which should be interpreted as the percentage point change in the likelihood that a household has installed their latrine at endline relative to the average household with a stated WTP greater than \$50. The sample consists of all winners, and standard errors are clustered at the village level. Figure A3 displays the point estimates and 95% confidence intervals for each WTP category indicator.

Households in higher WTP categories appear slightly more likely to install the latrine they purchase, though the relationship is not monotonic. Further, at the 5% level, we are only able to reject that the installation rate for households in the 20<WTP≤30 is different from the installation rate for households in the WTP>50 category; this suggests there is at best weak evidence of a screening effect among households in the sample as a whole.

However, since the relationship between the offer price and the share of households purchasing was quite different across treatment arms, one might worry that by pooling across arms Figure A3 may ignore important heterogeneity by treatment status. Furthermore, screening could operate differently in the two treatments. If it is primarily poverty or credit constraints, rather than low interest in health or lack of intent to use the product that limits WTP, then financing could lead to better targeting of scarce subsidy resources, i.e. allocating goods to those most likely to use them. On the other hand, households that are willing to make large, immediate sacrifices to make a lump-sum purchase could be more likely to use the item in question.

To explore this heterogeneity,[27] Figure A4 replicates the analysis discussed above for Figure A3, estimating Equation 6 separately for each treatment arm.

Despite a lack of precision, there are several interesting differences between the figures for the No Financing and the Financing treatment arms. First, the offer of financing had such a large impact on WTP that there are no households who won a latrine through BDM who bid an amount less than or equal to \$20. Second, ignoring the high but imprecisely estimated installation rate among the Non-financing "WTP $\leq$ \$20" group, the installation rate is increasing quite sharply in WTP in the Non-financing arm.[28] While the installation rate is also positively related to WTP in the Financing arm, the gradient appears substantially flatter.

While Figures A3 and A4 are fairly non-parametric tests for screening, inference based on them is limited by their relative lack of precision; we are unable to rule out the possibility of a small negative screening effect, no screening effect, or a substantial positive screening effect. Therefore, Table A4 presents a more parametric test for screening. We estimate

[27] This analysis of heterogeneity by treatment was pre-specified as Hypothesis 4B in the pre-analysis plan, and our analysis follows Equation (8) in that document. The overall low installation rates prevent us from conducting the more ambitious nonparametric analysis of Equation (9) in the pre-analysis plan.

[28] There were only 14 winning households in the "lump sum, WTP $\leq$ \$20" category, making the estimate for that category very imprecise. There were no winning households in the "Financing, WTP $\leq$ \$20" category, so that category does not appear in the figure.

$$I_{hv} = \alpha \mathrm{WTP}_{hv} + \sum_{p \in \mathbb{P}} \beta_p 1\left\{\mathrm{Draw}_{hv} = p\right\} + \theta T_v + \gamma' X_{hv} + \varepsilon_{hv}, \tag{7}$$

which closely resembles Equation 6 except we restrict the effect of WTP_{hv} to be linear. In exchange for this restriction, we gain considerable power to detect screening effects. Columns 1, 3, and 5 display the coefficients on WTP from specifications with no additional baseline controls for all households, households in lump sum villages only, and households in Financing villages only, respectively. Columns 2, 4, and 6 include baseline controls for household socio-economic status, demographics and open defecation practice. (Detailed definitions provided in the notes for Table A4.)

The estimates in Columns 1 and 2 indicate there is little relationship between WTP and latrine installation when pooling across both arms. The point estimates are small (-0.002), statistically insignificant, and unaffected by the inclusion of baseline controls. Within the Financing treatment (columns 5 and 6), the negative point estimates are marginally statistically significant, but economically small. The point estimate of -0.004 implies that a winning household at the 75th percentile of WTP ($\approx$ \$75) is 1.6 percentage points less likely to install its latrine than a winning household at the 25th percentile of WTP ($\approx$ \$35). In contrast, within the Non-financing treatment (columns 3 and 4), the point estimates are positive, an order of magnitude larger, and their statistical significance is more robust. The point estimate of 0.060 in column 3 implies that a winning household at the 75th percentile of WTP (\$40) is 12 percentage points less likely to install its latrine than a winning household at the 25th percentile of WTP (\$20), which is an economically meaningful difference. Including baseline controls slightly attenuates this relationship, but the effect of baseline willingness to pay remains positive, substantial, and statistically significant at the 10% level.

The difference between the screening figures, which appear to display a weak positive screening effect, and the negative regression-based estimates presented for the Financing arm occurs because the within-bin relationships between WTP and latrine installation are negative in the two most densely populated bins (\$30 to \$40 and \$40 to \$50). Further, the pooled regression-based screening estimates are closer to the Financing estimates because there are nearly twice as many BDM winners in the Financing arm as in the Non-financing arm.

Overall, these regression results support the interpretation of the descriptive evidence presented in Section 4.1: there is evidence of modest positive selection on WTP in the Non-financing treatment only.

D Causal Effect of Price Paid

In this section, we test for a causal effect of price paid on the probability of subsequent installation. One useful feature of BDM is that the price draw provides random variation in price paid among winners, conditional on WTP. That is, BDM provides the following experiment: two households with identical WTP, both of which buy the filter, but one purchases at a relatively low price and one at a relatively high price. Because the difference in price paid comes from the (random) BDM draw, it is uncorrelated with other determinants of use, once we have conditioned on WTP.

In our context, the theoretical relationship between price paid and subsequent use is ambiguous. Previous studies that attempt to measure the causal effect of price paid have typically done so with less costly goods (i.e. liquid chlorine, bednets), and the proposed mechanism has been primarily psychological, in which the "sunk cost fallacy" means that the act of paying more causes the purchaser to attach greater value to the product. The improved latrines being sold through the intervention studied in this paper are considerably more costly, and, as discussed above, households often desire an expensive complementary investment in a superstructure. These two factors increase the relevance of a second possible mechanism, a wealth effect: variation in price paid leads to variation in available funds for the complementary investment. That is, households who pay more for the latrine parts may have fewer remaining resources to install the latrine or purchase a superstructure, resulting in a negative relationship between the transaction price and latrine installation. Given the anecdotal information discussed above–the mean reported cost of superstructure in other rural Kampong Thom villages was 200 USD, roughly 170% of the mean monthly household income for our sample–important income effects seem likely in this context. While BDM does provide an experiment for estimating the net causal effect of price paid on subsequent installation, it does not allow us to separate these two mechanisms.

Our investigation of the causal effect of price paid largely mirrors the tests for screening conducted in Appendix C. We estimate Equation 6, but the coefficients of interest are the β_p, which tell us the probability of use associated with each price paid (20, 30, 40), controlling for category of WTP (omitting the draw price of 50). Figure A5 displays the resulting estimates.

Though quite noisy, the point estimates suggest there is a negative relationship between the transaction price and the share of households who had installed a latrine at baseline. All three point estimates are positive, indicating that households in the displayed (and lower) price draw categories were more likely to install the latrine than households in the highest ($50) price category. This would indicate that the negative wealth effect resulting from a higher price draw dominates any sunk cost effect, yielding a net negative causal effect of price paid. However, the confidence intervals are wide enough to include a range of positive and negative gradients and we fail to reject the null of no difference except when comparing the lowest price ($20) to the highest price ($50).

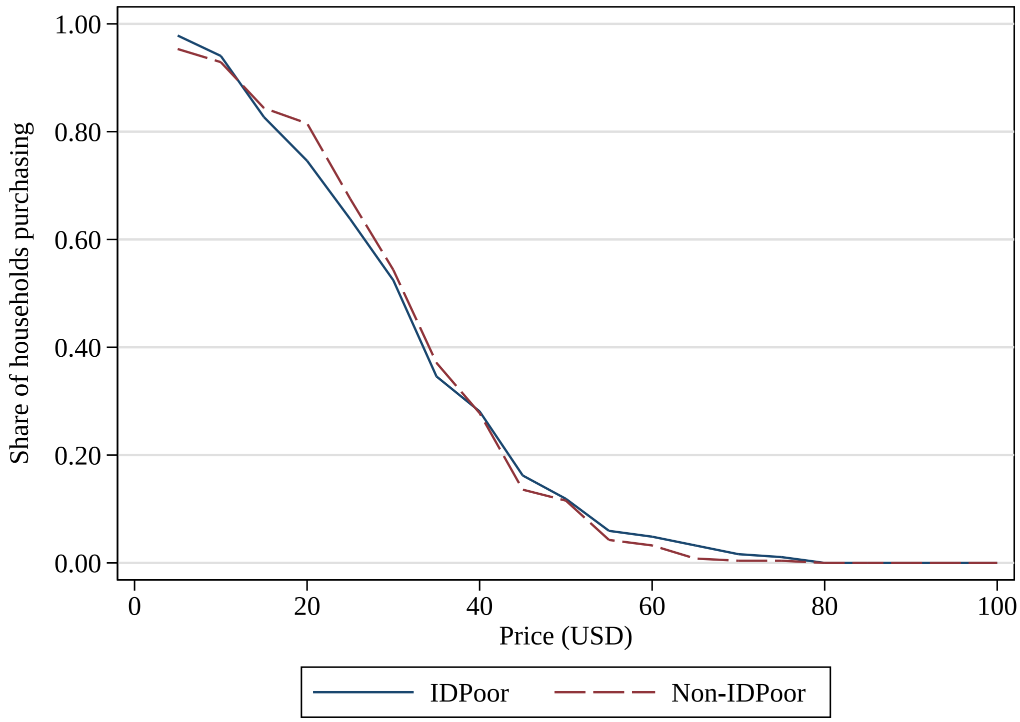
Figure A6 replicates the analysis of Figure A5 separately for each treatment arm. In both groups latrine installation at baseline is decreasing in transaction price, though as above, in all but one case–$20 compared to $50 in the Financing arm–the point estimates are too noisy to statistically reject the null of no difference between prices. To increase the statistical power of the tests for a causal effect of price paid, Table A5 assumes a linear relationship between the transaction price and latrine installation. The columns are ordered in the same manner as Table A4. We estimate

$$I_{hv} = \sum_{b \in \mathbb{B}} \alpha_b 1 \left\{ \mathrm{WTP}_{hv} \in \mathrm{bin}_b \right\} + \beta \mathrm{Draw}_{hv} + \theta T_v + \gamma' X_{hv} + \varepsilon_{hv}, \tag{8}$$

which resembles Equation 7 except we impose a linear functional form on price paid, which is the effect of interest here, and control nonparametrically for category of WTP.

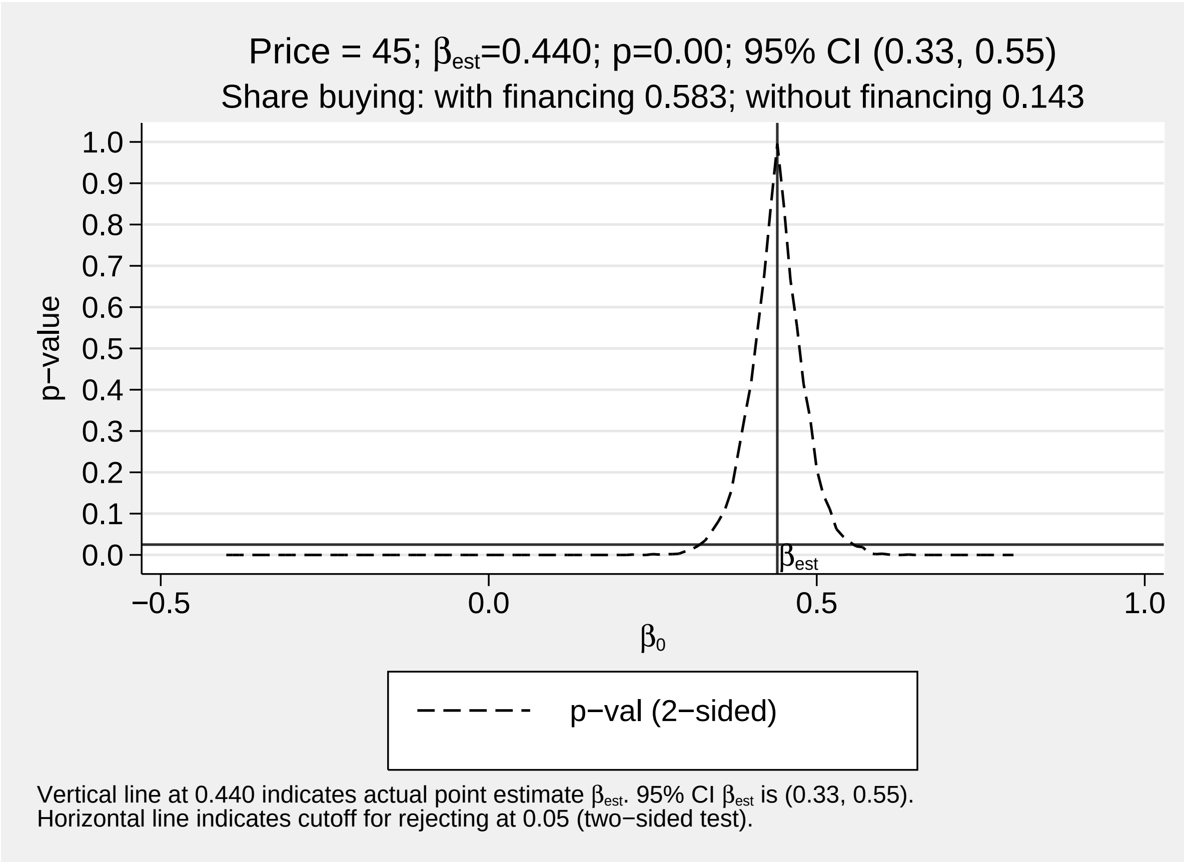
As suggested by Figures A5 and A6, Table A5 indicates there is a negative relationship between the transaction price and latrine installation at baseline. The results do not appear to vary by Financing treatment and including controls in the pooled specification has little effect on the point estimates or statistical significance. The pooled models suggest that a \$10 increase in purchase price reduces the likelihood that a household has installed their latrine at baseline by roughly four percentage points, a result that is significant at the 10% level. The estimates are generally similar between the two treatment groups.

Figure A1: Demand in Non-Finance Arm – IDPoor vs. non-IDPoor



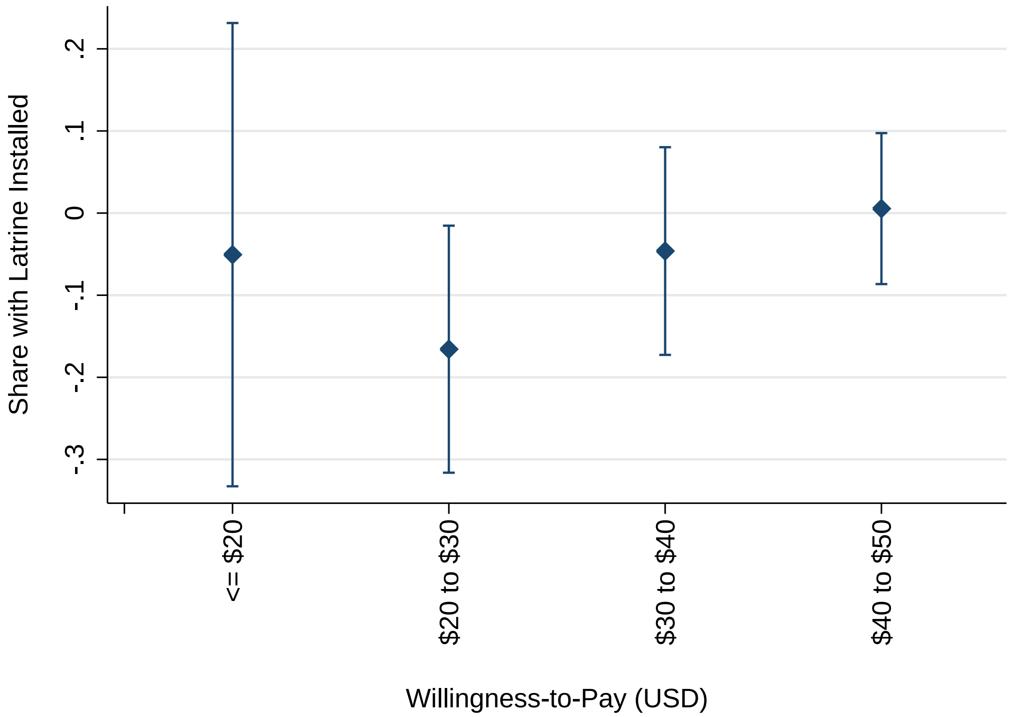
Notes: This figure shows the share of households in Non-financing villages with willingness to pay greater than or equal to each price, separately for IDPoor and non-IDPoor households. The share is calculated at at \$5, 10, . . . , 100.

Figure A2: Confidence interval by randomization inference



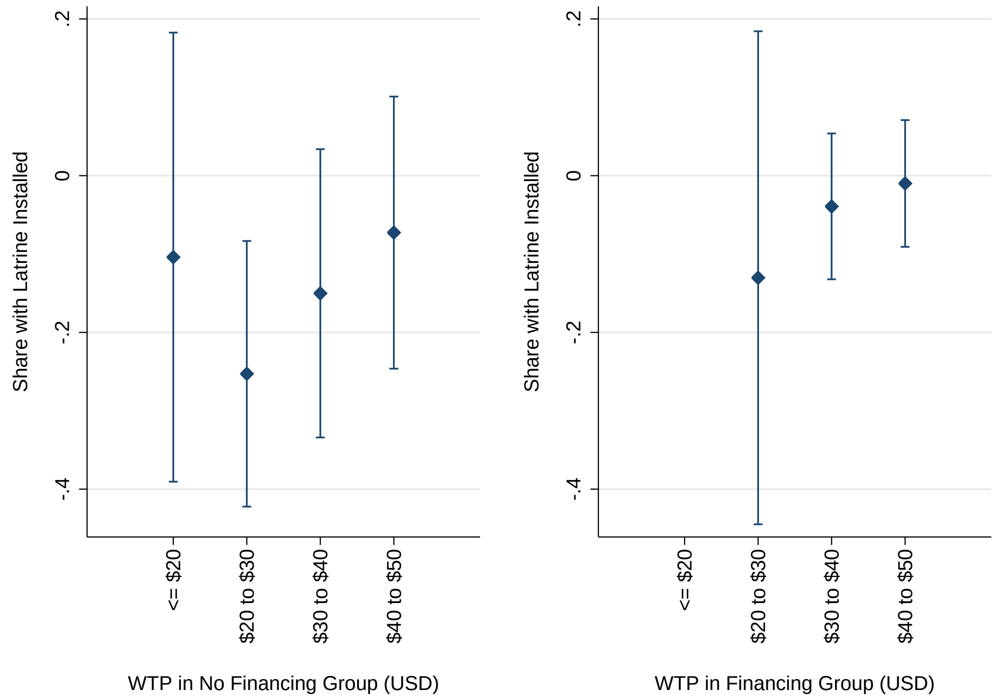
Notes: This figure illustrates the method of constructing a confidence interval via randomization inference. The vertical line indicates the point estimate, 0.440, for the price of 45. The horizontal line at 0.025 indicates the cutoff p-value for rejecting the hypothesis $\beta = \beta_0$ with 95% confidence, i.e. $p < 0.025$ for a two-sided test. The dashed line plots, for each value of the sharp null hypothesis $\beta_0 = -0.40, -0.39, \ldots, +0.79, +0.80$, the randomization inference p-value for the test $\beta = \beta_0$. The 95% confidence interval consists of the set of β_0 that are not rejected, i.e. all β_0 for which the dashed line is at or above the horizontal line. For this price, this is the set (0.33, 0.55).

Figure A3: Willingness to pay and Latrine Installation



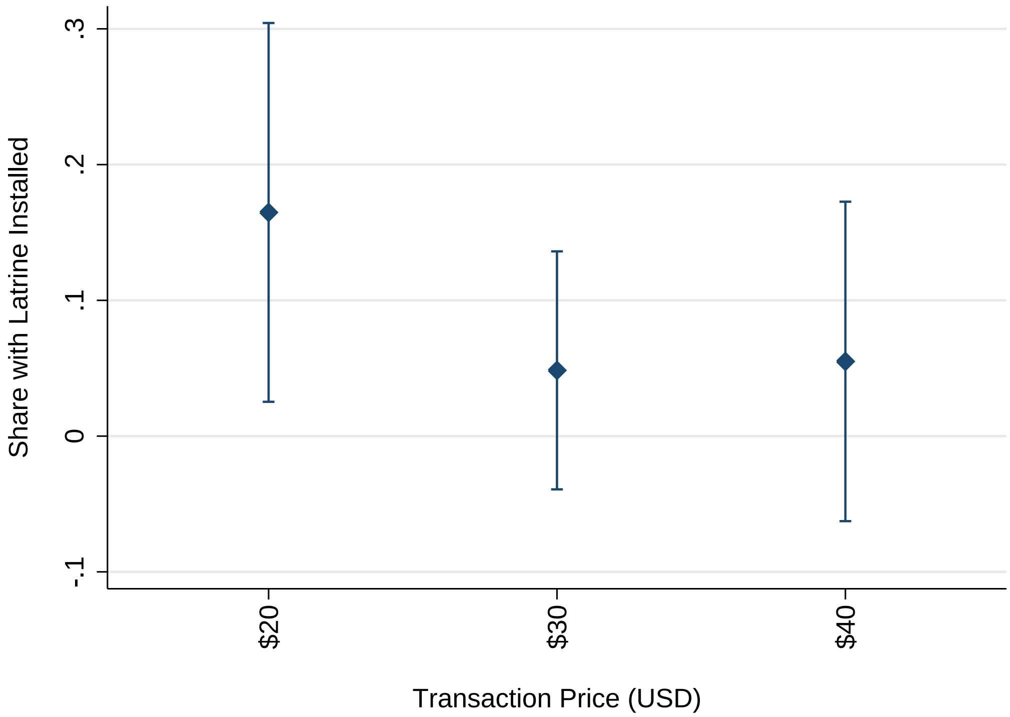
Notes: Figure presents the coefficients on the willingness to pay (WTP) category each household falls into based on a linear probability model with an indicator for whether the household had installed a latrine at endline (18 to 22 months after the sales offer) as the dependent variable. The omitted category is WTP greater than 50. The sample is limited to BDM winners (i.e. those who bid an amount at least as large as the random BDM draw price) regardless of treatment status. Dummies for the BDM draw price faced by each household are included as controls. Error bars reflect 95% confidence intervals based on heteroskedasticity-robust standard errors clustered at the village level.

Figure A4: Willingness to pay and Latrine Installation by Treatment Arm



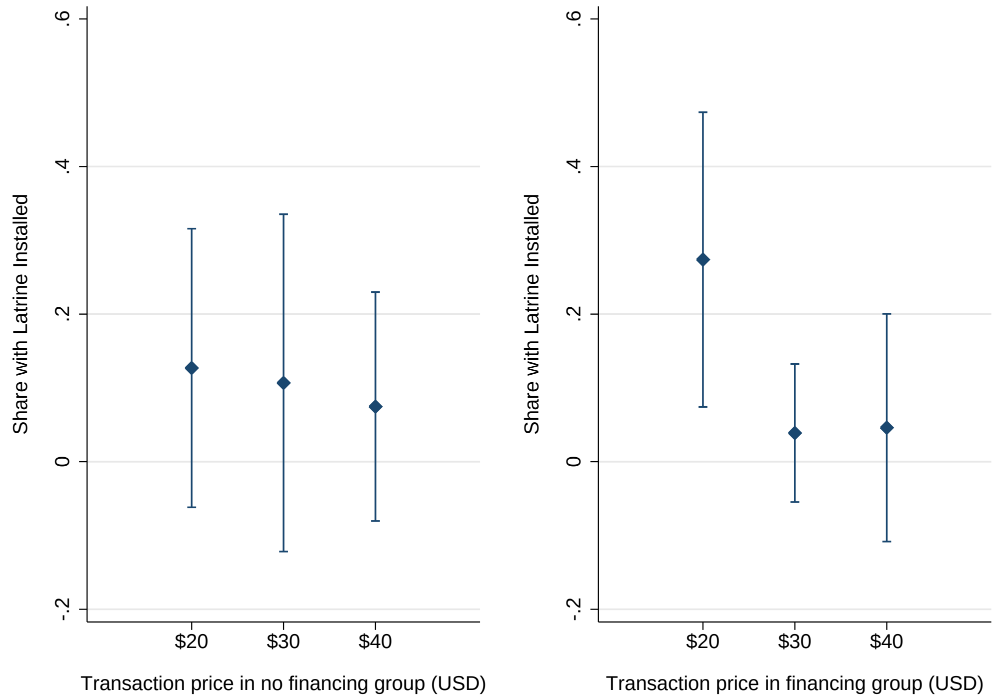
Notes: Figure presents the coefficients on the willingness to pay (WTP) category each household falls into based on a linear probability model with an indicator for whether the household had installed a latrine at endline (18 to 22 months after the sales offer) as the dependent variable. The omitted category is WTP greater than 50. The left panel includes only households in the Non-financing treatment arm while the right panel includes only households in the Financing treatment arm. The sample is limited to BDM winners (i.e. those who bid an amount at least as large as the random BDM draw price) regardless of treatment status. Dummies for the BDM draw price faced by each household are included as controls. Error bars reflect 95% confidence intervals based on heteroskedasticity-robust standard errors clustered at the village level.

Figure A5: Transaction Price (USD) and Latrine Installation



Notes: Figure presents the coefficients on the transaction price or BDM draw price faced by each household from a linear probability model with an indicator for whether the household had installed a latrine at endline (18 to 22 months after the sales offer) as the dependent variable. The omitted category is a draw price of 50. The sample is limited to BDM winners (i.e. those who bid an amount at least as large as the random BDM draw price) regardless of treatment status. We include dummies for WTP category ($\leq$20,20 to 30, 30 to 40, 40 to 50,>50). Error bars reflect 95% confidence intervals based on heteroskedasticity-robust standard errors clustered at the village level.

Figure A6: Transaction Price (USD) and Latrine Installation by Treatment Arm



Notes: Figure presents the coefficients on the transaction price or BDM draw price faced by each household from a linear probability model with an indicator for whether the household had installed a latrine at endline (18 to 22 months after the sales offer) as the dependent variable. The omitted category is a draw price of 50. The left panel includes only households in the Non-financing treatment arm while the right panel includes only households in the Financing treatment arm. The sample is limited to BDM winners (i.e. those who bid an amount at least as large as the random BDM draw price) regardless of treatment status. We include dummies for WTP category ($\leq$20,20 to 30, 30 to 40, 40 to 50,>50). Error bars reflect 95% confidence intervals based on heteroskedasticity-robust standard errors clustered at the village level.

Table A1: Randomization Inference –
Actual Treatment Assignment, Observed Outcome and Potential Outcomes

Village	T_v	s_v	y_{v1}	y_{v0}
1	1	0.5	0.5	?
2	1	0.6	0.6	?
$\vdots$	$\vdots$			
15	1	0.4	0.4	?
16	0	0.3	?	0.3
17	0	0.1	?	0.1
$\vdots$	$\vdots$			
30	0	0.2	?	0.2

Notes: we only only observe the potential outcome corresponding to the actual treatment status. The potential outcome corresponding to the other, counterfactual treatment status is unobserved and unknown.

Table A2: Randomization Inference – Hypothesis $\beta_0 = 0.2$
Actual Treatment Assignment, Observed Outcome and Potential Outcomes

Village	T_v	s_i	y_{i1}	y_{i0}
1	1	0.5	0.5	$0.5 - 0.2 = 0.3$
2	1	0.6	0.6	$0.6 - 0.2 = 0.4$
$\vdots$	$\vdots$			
15	1	0.4	0.4	$0.4 - 0.2 = 0.2$
16	0	0.3	$0.3 + 0.2 = 0.5$	0.3
17	0	0.1	$0.1 + 0.2 = 0.3$	0.1
$\vdots$	$\vdots$			
30	0	0.2	$0.2 + 0.2 = 0.4$	0.2

Notes: under a sharp null hypothesis (here, $\beta_0 = 0.2$), we can specify the potential outcome for each unit in both the actual and counterfactual treatment state.

Table A3: Randomization Inference – Hypothesis $\beta_0 = 0.2$; Placebo Randomized Treatment
Actual Treatment Assignment, Observed Outcome, Potential Outcomes, Observed Outcomes With Placebo Randomized Treatment

Village	T_v	s_v	y_{v1}	y_{v0}	T_v^r	y_v^r
1	1	0.5	0.5	$0.5 - 0.2 = 0.3$	0	0.3
2	1	0.6	0.6	$0.6 - 0.2 = 0.4$	1	0.6
$\vdots$	$\vdots$					
15	1	0.4	0.4	$0.4 - 0.2 = 0.2$	0	0.2
16	0	0.3	$0.3 + 0.2 = 0.5$	0.3	0	0.3
17	0	0.1	$0.1 + 0.2 = 0.3$	0.1	1	0.3
$\vdots$	$\vdots$					
30	0	0.2	$0.2 + 0.2 = 0.4$	0.2	1	0.4

Notes: in each repetition r, we generate a placebo random assignment and assign each unit the potential outcome corresponding to its placebo assignment. Then we calculate the observed treatment effect, $\hat{\beta}^r = \overline{y}_1^r - \overline{y}_0^r$. This is repeated R times, simulating the distribution of estimates we would expect to see when the null hypothesis is true.

Table A4: Willingness to pay and Latrine Installation

	All Villages		Lump Sum Villages Only		Financing Villages Only	
	(1)	(2)	(3)	(4)	(5)	(6)
Willingness-to-pay (10 USD)	-0.002	-0.002	0.060**	0.046*	-0.004*	-0.003
	(0.003)	(0.002)	(0.025)	(0.022)	(0.002)	(0.002)
Observations	738	709	284	255	454	454
Draw price fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Baseline controls	No	Yes	No	Yes	No	Yes
Fraction installed	0.314	0.316	0.342	0.349	0.297	0.297

Notes: Each column of the table displays the estimated coefficient on the willingness to pay (in $10 increments) for each household from a Linear Probability Model with an indicator for whether the household had installed a latrine at endline (18 to 22 months after baseline) as the dependent variable. All columns include dummy variables for price draw faced by the household in the BDM procedure ($p \in \{20, 30, 40, 50\}$). Columns 1, 3, and 5 include no other baseline controls, except a treatment indicator in Column 1. Columns 2, 4, and 6 include an indicator for whether households are IDPoor, the household progress out of poverty index score (based on the 2011 scorecard), an indicator for whether the household does any farming, an indicator for whether the household owns any livestock, an indicator for whether the household has any children under the age of two, the number of children in the household under the age of five, and an average across household members of whether open defecation was the most common method of defecation over the fifteen days preceding the survey. The sample is limited to househoulds who won a latrine through the BDM procedure, treating households who were not approved for a loan or who cancelled their order as BDM winners. Standard errors are clustered at the village level.

Table A5: Transaction Price (USD) and Latrine Installation

	All Villages		Lump Sum Villages Only		Financing Villages Only	
	(1)	(2)	(3)	(4)	(5)	(6)
Transaction Price (10 USD)	-0.040*	-0.039*	-0.034	-0.019	-0.041	-0.039
	(0.023)	(0.023)	(0.031)	(0.037)	(0.033)	(0.032)
Constant	0.464***	0.562***	0.453**	0.763***	0.481***	0.526***
	(0.149)	(0.140)	(0.162)	(0.139)	(0.063)	(0.121)
Observations	738	709	284	255	454	454
WTP bin fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Baseline controls	No	Yes	No	Yes	No	Yes
Fraction installed	0.314	0.316	0.342	0.349	0.297	0.297

Notes: each column of the table displays the estimated coefficient on the draw price (in $10 increments) faced by each household from a Linear Probability Model with an indicator for whether the household had installed a latrine at endline (18 to 22 months after baseline) as the dependent variable. All columns include dummy variables for the household's WTP category (≤ 20, $20-30$, $30-40$, $40-50$, > 50). Columns 1 and 2 include all winning households, regardless of treatment status, columns 3 and 4 are limited to households in lump sum villages only, and columns 5 and 6 are limited to households in Financing villages only. Columns 1, 3, and 5 include no other baseline controls (except a treatment indicator in Column 1). Columns 2, 4, and 6 include an indicator for whether households are IDPoor, the household progress out of poverty index scrore (based on the 2011 scorecard), an indicator for whether the household does any farming, an indicator for whether the household owns any livestock, an indicator for whether the household has any children under the age of two, the number of children in the household under the age of five, and an average across household members of whether open defecation was the most common method of defecation over the fifteen days preceding the survey. The sample is limited to househoulds who won a latrine through the BDM procedure, treating households who were not approved for a loan or who cancelled their order as BDM winners. Standard errors are clustered at the village level.