Pricing People into the Market: Targeting through Mechanism Design

Terence R. Johnson*
University of Notre Dame
and
Molly Lipscomb
University of Virginia

December 6, 2019

Abstract

Subsidy programs are typically accompanied by large costs due to the difficulty of screening those who should receive the program from those who would have purchased the good anyway. We design and implement a platform intended to increase the take-up of improved sanitation services by targeting the poorest households for subsidies and using purchases by the wealthy households to increase available subsidies to the poor. We develop a theoretical model designed to isolate the key factors of concern in designing the pricing system. The field project then proceeds in two stages: we first create a demand model based on market data and a demand elicitation experiment, and use the model to predict prices that will maximize take-up subject to an expected budget

*Johnson is assistant professor of economics at the University of Notre Dame. Lipscomb is associate professor of economics and public policy at University of Virginia's Batten School and the Department of Economics. We thank Shoshana Griffith, Adrien Pawlik and Adama Sankoudouma for excellent research assistance. We are grateful to the Bill and Melinda Gates Foundation for funding this project and research. We are grateful for helpful comments from Vivek Bhattacharya, Taryn Dinkelman, Pascaline Dupas, Joe Kaboski, Isaac Mbiti, Dilip Mookherjee, Laura Schechter, and conference participants at the Gates Foundation Workshop, Midwest International Development Conference, NBER Development Summer meetings, NYU, NYU Abu Dhabi, the North Carolina TREE seminar, University of Notre Dame, University of Stockholm, and William and Mary.

constraint. We then test the modeled prices on a new sample of households. The treatment led to an increase in market share of mechanical desludging of 4.4 percentage points. The decreased probability of purchasing a manual desludging among those with the largest subsidies was 7.6-8.2 percentage points leading to a market share increase of mechanical desludging of 7.9-9.6 percentage points in that group. The health impacts among neighborhoods with many poor households were large: a 10% increase in the number of poor households in a treatment neighborhood meant that there was a 2.2 percentage point larger decrease in diarrhea. We compare the outcomes of the pricing treatment with alternative targeting methods and pricing structures and show that the pricing treatment outperforms proxy means testing, auctions with perfect pass-through of costs, and straight subsidies on the basis of take-up and/or budget sustainability.

JEL Classification Codes: L1, D4, C7, D8
Keywords: Targeting, Subsidies, Mechanism Design

1 Introduction

Mechanism design offers an engineering approach (Roth, 2002) to economics that has revolutionized the study of many important markets[1]. A natural application of the theory is the targeting problem, in which a budget-constrained policy maker or NGO seeks to distribute aid to households who are privately informed about their willingness-to-pay and propensity to purchase a good on their own (Alatas et al., 2012, 2016), particularly when there are consumption externalities (Guiteras et al., 2015). This problem is ubiquitous, particularly in a development context, and includes the introduction of new technologies (Berry et al., 2015; Dupas et al., 2016), distribution of loans or grants to entrepreneurs (McKenzie and Sansone, 2017; Hussam et al., 2017), and the structure of payments for environmental services (Jayachandran et al., 2017). First, we provide a theory of targeting and show how the optimal mechanism uses observable information to set prices that strategically distribute aid

[1] For example, telecommunications spectrum (Milgrom and Segal, 2019), matching medical residents to hospitals (Roth and Peranson, 1999), kidney exchange (Roth et al., 2004), and assigning children to public school in Boston and New York City(Abdulkadiroglu and Sonmez, 2003)

and cross-subsidize poor households with profits made from wealthier ones. Second, we gather the data necessary to apply this solution to a particular problem: the removal of human fecal sludge in peri-urban neighborhoods of Ouagadougou. Third, we use a randomized controlled trial to evaluate the impact of this approach, comparing a treatment group of households with access to our market to a control group that was excluded. Our main results show that the intervention increased the market share of sanitary removal services by 4.4 percentage points, and that gains were concentrated among the poorest, most subsidized households, for whom market share increased by 9.6 percentage points. We conclude by estimating a model of demand for services provided both in the market and through the intervention, allowing us to illustrate how our approach compares to more standard ones, including auctions, proxy-means testing, and price ceilings.

Our setting is the handling and disposal of human fecal sludge in residential compounds in Burkina Faso, which is a challenge in urban sanitation throughout the developing world. Household sludge pits can be emptied mechanically, where a crew of two to four workers uses a truck-mounted vacuum to remove the sludge from a pit, or manually, where pits are cleaned out by family members or hired workers using trowels and buckets. Mechanical desludging minimizes exposure to fecal sludge and ensures its removal from the immediate neighborhood, while manual desludging typically ends with the disposal of the sludge in the street near the dwelling, resulting in negative externalities for nearby households. ABITABOUTTHESEARCHPROBLEM Rates of diarrhea are extremely high in developing countries: 1.8 billion people globally use a source of drinking water with fecal contamination, and 2.4 billion people lack access to safely managed sanitation services (WHO and UNICEF, 2015), which can result in stunting and other developmental disadvantages for young or vulnerable household members (Spears, 2013). Sanitation and water access form the sixth of the Sustainable Development Goals, and while subsidies have been effective at increasing take-up of health and sanitation goods, the demand for such products often remains low (Kremer and Miguel, 2007; Cohen and Dupas, 2010; Dupas, 2014) suggesting that larger subsidies for these goods may be necessary, particularly for the poorest households. But relatively wealthy households who would purchase the mechanical service at prevailing market prices are happy to accept assistance, diverting subsidy dollars from poor households and attenuating treatment effects.

We adopt the perspective of a local municipality or NGO creating a platform to maximize utilization of mechanical desludging services, subject to a limited subsidy budget. Even when conditioning on observable or verifiable information about households in general — such as number of rooms in the dwelling, the size of its compound, or past municipal water or electricity bills — there is still uncertainty about a particular household's willingness-to-pay and the expected price it would face in the decentralized search market. An ideal toolbox for this kind of targeting problem is mechanism design, which studies the implementation of socially desirable outcomes when market participants have private information and individual agency. In standard optimal pricing problems, a profit-maximizing seller is only concerned with whether a consumer buys and at what price, but because of the externalities of inferior sanitation methods, the municipality here is also concerned with whether a household would purchase manual or mechanical desludging in the absence of the intervention. Consequently, the optimal platform trades off a social motive to quote lower prices and encourage mechanical take-up with a profit motive to quote higher prices and relax its budget constraint. It strikes a balance by using observable information not only to pick which households are subsidized but also how much assistance to provide. The optimal platform thus cross-subsidizes poor or socially valued households with profits made from richer ones, based on the platform's inference about the household's propensity to purchase mechanical desludging services and the costs of provision. This approach applies whenever a socially motivated actor like a government or NGO seeks to increase total market share of an improved product, including goods that generate externalities, like cookstoves or toilets, as well as ones that don't, such as solar lights. The main challenge is to avoid distributing subsidy dollars to rich households who already consume improved products, or failing to sufficiently subsidize poorer households so that the product comes within financial reach: our approach addresses both problems.

We then gather the data necessary to build the optimal platform in the context of desludging services in Ouagadougou. On the supply side, we invite mechanical service providers to participate in monthly neighborhood-by-neighborhood auctions in which the lowest bidders win and are paid the lowest rejected bid for every job they complete for the platform in that month, where jobs are distributed at random among the winners. On the demand side, we ask households to participate in an auction in which the highest bidders win but only have to pay the highest rejected bid. This generalization of the second-price auction gives the firms and households a weakly dominant strategy to bid honestly, providing us with unbiased estimates of firm costs and the households' willingness-to-pay for a desludging from the platform. Combining the willingness-to-pay reports with market data on the desludging purchases that households most recently made and the prices they paid allows us to predict which households are likely to purchase on their own, and which are the cheapest for the platform to convert to mechanical desludging through an attractive price offer. This is an important distinction, because the goal of the platform is not to maximize volume or total sales, but instead to maximize the total market share of mechanical desludging. That households can opt out of our market in favor of the prevailing, decentralized market differentiates our setting from many existing studies that focus on increasing demand for novel products that face little or no competition. By purchasing desludgings in bulk at low prices through competitive mechanisms, we can undercut the high prices offered to richer households in the existing search market. Wealthier households then either decline our offer in favor of buying a mechanical desludging at a more attractive price in the existing market, or purchase a desludging from us and thereby provide revenue that can be used to cross-subsidize poorer households. This kind of targeting has uses beyond sanitation services since it explicitly explores the demand curve below prevailing prices, providing policy-makers with information about the impact and sustainability of different subsidy levels.

We then use the estimates of cost and demand to opertionalize our theoretical results and design a pricing rule based on a set of observables that are known or easily verified by a local governmental authority or utility, like the Burkina Faso Office of Sanitation (ONEA), subject to a budget balance condition that losses per household not exceed a given subsidy threshold of \$3.00. The choice of observables is a crucial decision in the design of the platform, and they should be verifiable to the government authority, difficult to manipulate, cheap to measure, and correlated with wealth or the probability of selecting mechanical desludging. Otherwise, households have an incentive to lie or misrepresent themselves in order to receive more favorable treatment. For example, past water and electricity bills may be cheap for the municipality to obtain from local utilities, and the quality of a dwelling can easily be assessed by an enumerator in a short visit. In order to determine the price quoted to each type of household, we use: water and electricity expenditure; house type (precarious, concrete structure, or rooming house); whether the house is owned or rented, number of members in the household, number of women in the household, number of other households in the compound; desludging frequency; distance from the pit to the road; and whether the respondent has a high education level. Section 6 uses a counterfactual model to predict how alternative information structures would have performed, limiting the information set to mimic what would be available to a more constrained NGO or municipal authority. The approach is similar to Wolak (2016), who uses census, satellite and past usage data to design household water tariffs in California. Szabo (2015) estimates demand for water in South Africa and shows that adopting optimal non-linear tariffs can raise the same amount of revenue while improving allocative efficiency. The use of demand elicitation games to measure willingness-to-pay for health products has also been used by Berry et al. (2015), who estimate the demand for water purifiers in Ghana.

We then use a randomized controlled trial to measure the impact of the optimal platform. Access to the platform is provided to a second random sample of households, and their outcomes are compared with a third random sample of households that serves as a control group, who do not receive access to the platform. During the baseline interview for these participants, each treatment household receives a take-it-or-leave-it offer based on the observable information and the enumerator's subjective assessments. We find that neighborhoods with the targeting treatment have a 4.4 percentage point higher market share for the improved sanitation service than neighborhoods in the control group. There is no impact from the treatment on the wealthy households who have a high (99.3%) use of mechanized desludging services, even without the treatment. The treatment effect is generated entirely by the poorest households who were offered the lowest, below-market-average targeted prices: while market share of mechanical desludging among the poor households in the control group is 58.9%, market share among the poor households in the treatment group is 68.2%. These effects also appear among the poor, highly subsidized households at the household level: those with the largest subsidies were 8.2 percentage points less likely to purchase a manual desludging and 8.2 percentage points more likely to purchase a mechanical desludging. This improvement in the sanitation conditions also led to a decrease in diarrhea in children in neighborhoods with more poor households: treated neighborhoods with 10% more poor households (approximately 3 additional poor households) had a 2.2 percentage point drop in the probability of a report of diarrhea among children in a household relative to similar neighborhoods in the control group.

By offering a new approach to targeting that relies on data to decide how to distribute subsidies, this paper contributes to a number of different literatures. Targeting in existing programs has been found to be only moderately successful: for example, Coady et al. (2004) find that some targeting programs transfer only 25% more than random or universal allocation to poor households, with 27% of programs found to be regressive. Several methods of targeting aid and subsidies have been proposed and evaluated: proxy means tests based on the household's ownership of a basket of assets (Kidd and Wylde, 2011; Narayan and Yoshida, 2005; Banerjee et al., 2018); ordeal mechanisms in which the household must collect and submit coupons or undergo an application process (Alatas et al., 2012; Dupas et al., 2016; Alatas et al., 2016)[2]; and community-based targeting in which members of the local com-

[2]See Olken (2016) for a review.

munity or local government select which people should receive the program (Basurto et al., 2017). Jack (2013) places special emphasis on how auctions can reveal information that is also useful for targeting purposes. While these mechanisms may work well when the government has the resources to devote to a large anti-poverty program, in cases where the transfer is limited to a subsidy on a particular product, it may be possible to cross-subsidize between households by keeping the wealthier households engaged in purchasing through the platform. In this paper, we develop a framework for designing a pricing policy based on limited information about households to increase the take-up of a sanitation product with substantial externalities.

The randomized controlled trial rigorously tests the effectiveness of the proposed platform, but a number of questions remain about how other designs would have performed. In Section 6, we estimate a model that predicts the purchasing behavior of households and financial outcomes for the platform. We show that auctions, proxy-means testing, and price ceiling policies all deliver slightly lower average treatment effects, but much lower treatment effects among the poorest households. Auctions select on willingness-to-pay, making the platform essentially a transfer to relatively wealthy households. Proxy-means testing targets noisy measures of wealth that do not correlate well with actual behavior, nor use cross-subsidization or adjust assistance to ensure poor households are sufficiently subsidized to buy the expensive service. A price ceiling typically benefits richer households who faced price discrimination, without using subsidies to make the product affordable for poorer households. Indeed, to match the average treatment effect of the optimal platform, the auctions require an additional 1,704 CFA subsidy per household on average, proxy-means testing requires 1,643 CFA, and the price ceiling requires 2,478 CFA. To investigate what would happen if different kinds of information were used to design the platform, we adjust the variables used to mimic an NGO and a municipal authority, reducing the data available and re-solving for the optimal platform. In our case, the treatment effects remain relatively large, but unintended budgetary losses become larger as it becomes more difficult to target poor households and the intervention becomes unintentionally more generous. Finally, we investigate counterfactual values of the subsidy per household, letting it range from a negative subsidy (profit) of -750 CFA per household to 13,000 CFA. We find that even without positive subsidization, there can be a positive treatment effect because access to the market and competitive prices from the auctions increases welfare. In addition, we find that reaching the very poorest households can be expensive, requiring that subsidies reach 60% of the average market price to induce them to take up the healthier service.

2 Background

Lack of adequate sanitation is a primary cause of approximately 10% of global diseases, primarily through diarrheal diseases (Mara et al., 2010). While there has been substantial attention to increasing access to toilets for households (Guiteras et al., 2015; Kar and Pasteur, 2005) and open defecation (Gertler et al., 2015) in rural areas, there has been less attention to sanitation issues in urban environments where the impact of inadequate sanitation may be particularly high (Coffey et al., 2014). While the coverage of latrines in urban environments is high, latrines fill between every 6 months and 4 years, and inadequate management of the fecal sludge creates negative externalities and becomes a health hazard to the household and neighborhood. Attempts by NGOs to improve the sanitation issues caused by manual desludging have focused on heavily subsidizing as many mechanical desludgings as possible, but these programs typically run out of budget quickly.

How does the market for desludging services operate? Households can choose between mechanical emptying, in which a vacuum truck comes to the household, pumps the latrine sewage into the truck's tank, and empties the tank at a treatment center, and manual emptying, in which a family member or worker digs a trench in the road next to the household's compound and uses buckets to transfer the sewage from the latrine into the trench in the road. The externalities associated with manual desludging are substantial: the sewage dries over time in the street, but attracts bugs and parasites, affecting both the household itself and its neighbors. Mechanical desludging tends to be more expensive than manual desludging[3]. Rather than pay for either service, the poorest households often manually desludge their own latrine pits, compounding the potential for adverse health outcomes: at endline, 14.4% of households desludging manually desludged their own pit for free.

Why don't all households opt for the mechanical service? Households using manual desludging typically state that they would have preferred to use mechanical desludging, but choose manual because of the price. High prices are, however, often a symptom of other market failures. At baseline, 12% of households had searched for a mechanical desludger prior to getting their last manual desludging, and over 60% of those who searched for a mechanical desludger but used a manual desludger report searching for a week or more before going with a manual desludger. The median household reports looking for their last mechanical desludger for 12 days and having searched for a mechanical desludger for 24 days or more on at least one occasion in the past. The most common ways to find desludgers are calling the desludger that they used last time (44%), going to a parking lot (14%), and asking family or friends for a desludger phone number (8.5%). However, 30% of households report waiting because they had trouble finding a desludger or because the desludger with whom they had negotiated was not available. In economic terms, there is a missing market for mechanical desludging, creating search costs and the scope for price discrimination.

There is suggestive evidence of price discrimination. Prices tend to be higher for households that use an intermediary (1,700 CFA higher on average), call a number that they saw on a truck (945 CFA higher), or ask a desludger that they know lives nearby (500 CFA higher). Decentralized markets create the potential for the exercise of price power: a household negotiating with a desludger must weigh the likelihood of finding someone else to do the job at a lower price with the costs of further search and the burdens of a full latrine pit,

[3]The median price of both manual and mechanical desludging is 15,000 CFA (approximately $30), but the price of mechanical second order stochastically dominates the price of manual), see Figure 1.

and negotiations often break down. Financial constraints can also be a factor in delays: 42% of households report waiting because they had to collect funds to pay for the desludging. Households that cannot afford to pay or wait any longer ultimately turn to manual desludging. Thus, incomplete information about firm costs and household willingness-to-pay leads to inefficiency.

Centralizing the market addresses many of these problems, and we propose doing so through a call center that minimizes costs on the supply side through competitive procurement processes and maximizes impact on the demand side by exploiting available information to target prices based on observables and cross-subsidizing consumption by poor households with profits from wealthier ones.

3 Targeting through Mechanism Design

In general terms, "targeting" refers to the method by which beneficiaries are selected to receive aid from social programs (Alatas et al. (2016)), including auctions, proxy-means testing, social voting, and ordeal mechanisms. Participants privately hold important information about their ability- or willingness-to-pay, and incenting them to reveal this information can improve the performance of social programs. The problem of how to design these incentives leads naturally to a mechanism design analysis. This section poses and solves the mechanism design problem of a platform that competes alongside a prevailing market to maximize take-up of a socially beneficial health product, subject to incentive compatibility and individual rationality constraints, as well as a budget constraint requiring that its total losses not exceed a given subsidy level. We show that the platform acts as a "profit-minded social planner," optimally charging relatively high prices to households who would purchase otherwise and relatively low prices to households who will likely be unable to afford the mechanical service on their own.

There is a unit mass of households, all of whom must decide between purchasing manual or mechanical services. Each household has a privately known willingness-to-pay w, privately known outside price r, and publicly known observables x from a set[4] X. The willingness-to-pay w is the maximum price at which a household would be willing to switch from manual to mechanical desludging. The outside price r is the amount the household anticipates paying in the prevailing decentralized market for a mechanical desludging. The observable type x corresponds to characteristics observable to the market, like the household's neighborhood or the quality of its dwelling, or observable to a municipal authority such as ONEA, such as water or electricity bill expenditures.

The willingness-to-pay w and outside price r are distributed $F_w[w|x]$ and $F_r[r|x]$, with support on $[\underline{w}, \overline{w}]$ with densities $f_w[w|x]$ and $f_r[r|x]$, respectively. Conditional on x, r is independent of w since the market does not observe the household's private information[5]. A desludger can never reasonably charge more than the maximum willingness-to-pay of a household, and neither is it profitable for a firm with some price power to charge less than the minimum willingness-to-pay of a household. Thus, the support of r is also $[\underline{w}, \overline{w}]$. Assume the standard regularity conditions that $1 - F_w[w|x]$ and $1 - F_r[r|x]$ are log-concave[6].

Households have quasi-linear utility, so that consuming a mechanical desludging at a price of t yields a payoff $w - t$, while the payoff of consuming a manual desludging is normalized to 0. A household procures the mechanical service in the prevailing market only if its willingness-to-pay is sufficiently high, so that $w - r \geq 0$. The platform competes alongside a prevailing, decentralized market for mechanical services. Since our sample includes a small number of households relative to the overall size of the market, we assume that the

[4]In our applications, x will correspond to data including a mix of real-valued variables, integer-valued ones, dummy variables, and categorical variables or factors, so we place no additional structure on X beyond requiring it be a probability space, and the distributions and densities of w and r and the mechanism be $(\Omega_X, \mathcal{F}_X, \mu_X)$-measurable, so that expectations are well defined.

[5]It if were otherwise, knowing r would reveal additional information about w about x, implying that the market must be using additional variables to determine a price. We are assuming x includes all publicly available data on the household, from which the market uses some subset to set prices.

[6]This implies that a profit-maximizing monopolist's second-order condition is satisfied. See (Mussa and Rosen, 1978), (Myerson, 1981), and (Bagnoli and Bergstrom, 2005).

platform does not create general equilibrium effects that change the expected probability of trade or payment in the prevailing market[7].

What would happen in the full information benchmark where w and r are observable by the platform? With w and r known, the platform engages in perfect price discrimination against those households who would purchase anyway and are profitable to serve: those for whom $w \geq r$ and $r \geq c_x$. These profits $w - c_x$ from all such $x \in X$ relax the platform's budget constraint but fail to increase the share of mechanical consumption. To increase mechanical consumption, the platform redistributes these proceeds plus the subsidies to households who would otherwise fail to purchase mechanical, for whom $w < r$. The households for whom $w - c_x$ is the smallest are the least costly to induce to switch, so the platform finds the largest set of such households that satisfy the budget constraint.

Once incomplete information about w and r is introduced, however, this scheme of perfect price discrimination and cross-subsidization is not possible. Households that would purchase anyway will misrepresent themselves as households that would fail to purchase, and all households would strategically understate their willingness-to-pay. This is the essence of the targeting problem: scarce subsidy dollars can end up in the hands of relatively rich households instead of relatively poor ones, and this diversion of scarce subsidy dollars can make it impossible to provide low enough prices to relatively poor households to induce them to switch.

To solve the more demanding problem with private information, we invoke the Revelation Principle, which guarantees that any method the platform could use to arrange trade is equivalent to some direct mechanism, $\{p(w,r,x), t(w,r,x)\}_{w \in [\underline{w}, \overline{w}], r \in [\underline{w}, \overline{w}], x \in X}$ in which a household with observables x reports — not necessarily honestly — some type $(\hat{w}, \hat{r})$ and trade occurs with probability $p(\hat{w}, \hat{r}, x)$ at a price of $t(\hat{w}, \hat{r}, x)$. A direct mechanism is *incentive compatible* if[8] households find it to be in their best interests to participate

[7]The conclusion discusses some of the issues concerning selection onto the platform, which is an issue at scale.

[8]The more standard way of writing these constraints in the mechanism design literature is to subtract $t(w, r, x)$ from expected surplus without multiplying it by $p(w, r, x)$. Readers

honestly, or, for all observables x, types w and r, and reports $\hat{w}$ and $\hat{r}$,

$$\begin{aligned}\underbrace{p(w,r,x)}_{\text{Pr[Trade on the platform]}}\underbrace{(w-t(w,r,x))}_{\text{Platform payoff}}+\underbrace{(1-p(w,r,x))}_{\text{Pr[Trade off the platform]}}\underbrace{\max\{w-r,0\}}_{\text{Outside option}}\\ \geq p(\hat{w},\hat{r},x)(w-t(\hat{w},\hat{r},x))+(1-p(\hat{w},\hat{r},x))\max\{w-r,0\}\end{aligned}$$

or, converting to net quantities and noting that $w-\max\{w-r,0\}=\min\{w,r\}$,

$$p(w,r,x)(\min\{w,r\}-t(w,r,x))\geq p(\hat{w},\hat{r},x)(\min\{w,r\}-t(\hat{w},\hat{r},x)). \tag{1}$$

Similarly, a direct mechanism is *individually rational* if all households do at least as well participating as opting out, or, for all w, r, and x,

$$\underbrace{p(w,r,x)}_{\text{Pr[Trade on the platform]}}\underbrace{(w-t(w,r,x))}_{\text{Platform payoff}}+\underbrace{(1-p(w,r,x))}_{\text{Pr[Trade off the platform]}}\underbrace{\max\{w-r,0\}}_{\text{Outside option}}\geq\underbrace{\max\{w-r,0\}}_{\text{Outside option}},$$

or, again converting to net quantities,

$$p(w,r,x)(\min\{w,r\}-t(w,r,x))\geq 0. \tag{2}$$

In addition to the individual rationality and incentive compatibility constraints, the platform must also ensure that its total profits plus subsidies, s, are non-negative in expectation, or

$$\mathbb{E}_{(w,r,x)}\left[p(w,r,x)(t(w,r,x)-c_x)\right]+s\geq 0 \tag{3}$$

where c_x is the expected cost of serving a household with observables x. Call (3) the *expected budget balance* constraint.

As discussed in the introduction, there are significant negative externalities from the collection and disposal of human fecal sludge, especially on young children for whom exposure to human waste can lead to diarrhea, stunting,

more comfortable with this approach can substitute $\tilde{t}(w,r,x)=p(w,r,x)t(w,r,x)$, and the analysis and results will be identical.

and death. Let b_x be the net social benefit of a household of type x consuming the mechanical service rather than manual. The platform seeks to solve the *targeting problem*: pick the payments $t(w,r,x)$ and probabilities of trade $p(w,r,x)$ to solve:

$$\max_{\{p,t\}} \mathbb{E}_{(w,r,x)}\left[\underbrace{p(w,r,x)b_x}_{\text{Platform purchases}} + \underbrace{(1-p(w,r,x))\mathbb{I}\{w \geq r\}b_x}_{\text{Market purchases}}\right]$$

subject to incentive compatibility (1), individual rationality (2), and expected budget balance (3). Due to the challenges of eliciting households' preferences over their neighbors' consumption of the mechanical service[9], we focus on maximizing the share of mechanical services, setting $b_x = 1$.

Honest revelation of both w and r cannot both be incented: only the minimum of the two appears directly in the household's incentive constraints, so that the household will lie in the most advantageous way about the maximum of the two[10]. In order for a direct mechanism to be incentive compatible, it must then be a function only of the minimum of $\hat{w}$ and $\hat{r}$. Define $\eta = \min\{w,r\}$, and instead ask households to make a report of this value, $\hat{\eta}$; to distinguish this from the willingness-to-pay w, we refer to η as the household's *willingness-to-switch*. Transforming the problem in this way allows us to use standard tools to compute[11] the platform's profits in terms of the probabilities

[9]We piloted a variety of demand elicitation games that asked whether households would be willing to pay something if ensured their neighbors received mechanical services, but participants found this unnatural, given the political economy of their neighborhoods.

[10]If a household was going to purchase anyway, the relevant thing to lie about is the price it would have faced; if a household was not going to purchase mechanical services on its own, the relevant thing to lie about is its willingness-to-pay. So in either case, only w or r is payoff-relevant, not both.

[11]Appendix E provides a full analysis of the mechanism design problem, starting with the standard characterization of incentive compatibility in terms of the envelope theorem and monotonicity condition, computation of profits, derivation of a relaxed solution, and verification that under some set of reasonable regularity conditions, the monotonicity condition fails to bind at the optimum.

of trade in any incentive compatible direct mechanism:

$$\mathbb{E}_{(\eta,x)}\left[p(\eta,x)(t(\eta,x)-c_x)\right]=\mathbb{E}_{(\eta,x)}\left[p(\eta,x)\left\{\eta-\frac{1-F_\eta[\eta|x]}{f_\eta[\eta|x]}-c_x\right\}\right]. \quad (4)$$

The quantity

$$\psi_\eta[\eta|x]=\eta-\frac{1-F_\eta[\eta|x]}{f_\eta[\eta|x]}$$

is called the *virtual value*, and represents the expected marginal revenue[12] generated by providing a desludging to a household reporting η given x. It can be understood as the total surplus η, less an informational rent that accrues to the household due to the presence of private information, $(1-F_\eta[\eta|x])/f_\eta[\eta|x]$, that captures the cost to the platform of providing incentives for honest reporting.

The platform's objective function can similarly be simplified as a function of η:

$$\begin{aligned}&\mathbb{E}_{(w,r,x)}\left[p(\min\{w,r\},x)b_x+(1-p(\min\{w,r\},x))\mathbb{I}\{w\geq r\}b_x\right]\\&\quad=\mathbb{E}_{(\eta,x)}\left[p(\eta,x)\frac{h_w[\eta|x]}{h_w[\eta|x]+h_r[\eta|x]}b_x+\frac{h_r[\eta|x]}{h_w[\eta|x]+h_r[\eta|x]}b_x\right]\end{aligned} \quad (5)$$

where $h_z[\eta|x]$ is the hazard rate of the random variable z at η given x: $f_z[\eta|x]/(1-F_z[\eta|x])$. The *switching function*

$$\sigma(\eta,x)=\frac{h_w[\eta|x]}{h_w[\eta|x]+h_r[\eta|x]}$$

quantifies the platform's inference about a household's propensity to switch from manual to mechanical. It answers the question, "given a report of η, what is the probability that the household is switching from manual to mechanical, rather than switching from purchasing the mechanical service in the search

[12]To see this connection more explicitly, consider the classical monopolist's problem $\max_t D(t)(t-c)$, where demand is $D(t)=1-F(t)$. The first-order necessary condition satisfies $t^*-(1-F(t^*))/f(t^*)=c$, or the classical "marginal revenue equals marginal cost" condition. Thus, $\psi(t)=t-(1-F(t))/f(t)$ can be interpreted as marginal revenue. For more on this interpretation, see Bulow and Klemperer (1996).

market to buying from the platform?" This is the difference between classical non-linear pricing and targeting: the platform cares not only about the level of sales, but what the household would likely do in the absence of the intervention.

The optimal allocation rule $\{p(\eta, x)\}_{x \in X}$ then necessarily maximizes the Lagrangian

$$\mathcal{L}(p, \lambda) = \mathbb{E}_{(\eta,x)} \left[p(\eta, x) \left\{ \sigma(\eta, x) b_x + \lambda \left(\psi_\eta[\eta|x] - c_x \right) \right\} \right] \tag{6}$$

where λ is the multiplier on the expected budget balance constraint. The term in braces represents the marginal benefit of serving a household with observables x reporting η,

$$\underbrace{\sigma(\eta, x)}_{\text{Marginal propensity to switch at } (\eta,x)} b_x + \underbrace{\lambda}_{\text{Shadow value of profit}} \underbrace{(\psi_\eta[\eta|x] - c_x)}_{\text{Marginal profit from } (\eta,x)} . \tag{7}$$

When this term is positive, the platform prefers to provide a desludging to the (η, x) type and set $p(\min\{w, r\}, x, \lambda) = 1$, and otherwise set $p(\min\{w, r\}, x, \lambda) = 0$. The first term is the odds of a switch at η given x, capturing the social motive. The second term is the marginal profit generated by the sale to the (η, x) type weighted by the shadow value of the expected budget balance constraint, capturing the profit motive. If λ is small, the platform will generously distribute mechanical desludgings at low prices, while if λ is large, the budget constraint is relatively binding and it will behave more like a purely profit-maximizing platform. This illustrates how the platform is a "profit-minded social planner," who places some weight on profits and some on consumption of improved services, where the weight is endogenously determined by the balancing the budget with the relative likelihoods of the households to switch.

A full analysis of the problem characterizes[13] the optimal mechanism in this environment:

Theorem 1 *Suppose $\underline{w} - c_x \leq s$, so that the subsidy is not sufficiently large to provide every household in the market with a mechanical desludging. The*

[13]See Appendix E. The results are summarized in the statement of this Theorem to streamline the exposition.

optimal allocation can be implemented by making take-it-or-leave-it offers to each observable x at a price of

$$t_x^* = c_x + \frac{1 - F_\eta[t_x^*|x]}{f_\eta[t_x^*|x]} - \frac{\sigma(t_x^*, x)b_x}{\lambda^*} \tag{8}$$

where λ^* exists and is a solution to $\mathbb{E}_{(\eta,x)}\left[(1 - F_\eta[t_x^*(\lambda)|x])(t_x^*(\lambda) - c_x)\right] + s = 0.$

If the platform offers an attractive offer to a type (w, r, x) household, any household of type (w', r', x) with $\min\{w', r'\} > \min\{w, r\}$ can behave as if it was a type (w, r) household and receive the same attractive offer. Thus, the terms of trade for all households with observables x must be the same as the household with the lowest η with observables x that buys through the platform. This naturally leads to a cut-off rule where types η greater than some η_x^* purchase on the platform while those with $\eta < \eta_x^*$ do not, which is implementable by posted prices[14], $\eta_x^* = t_x^*$.

The operation of the optimal mechanism for a given $x \in X$ is illustrated in Figure 2. In the absence of the platform, only the households for whom $w \geq r$ would make a purchase on their own, corresponding to the set of types below the diagonal. If the platform offers a price t_x^*, this creates four groups:

i. *Non-buyers*: those who find neither the market nor the platform price attractive, and purchase a manual desludging ($r > w$ and $t_x^* > w$)

ii. *Non-participating buyers*: those who prefer the market price to the platform price, and purchase in the prevailing market ($w > r$ and $t_x^* > r$)

iii. *Participating buyers*: those who prefer the platform price to the market price, and purchase on the platform but would have purchased in the market ($w > r$ and $t_x^* < r$)

[14]If the virtual value $\psi_\eta[\eta|x]$ or switching function $\sigma(\eta, x)$ failed to satisfy the single-crossing property in η, it is possible that the standard monotonicity condition that $p(\eta, x)$ be non-decreasing in η would bind. This can introduce the usual pooling and randomization into the optimal mechanism, where a household of type x would be presented with a schedule of prices and probabilities of service and asked to select one. Similarly, if there were a finite number of households instead of a unit mass, the problem would more closely resemble an auction rather than non-linear pricing, and would result in more complicate pricing schemes.

iv. *Switchers*: those who prefer the platform price to the market price, and would not have purchased in the prevailing market ($r > w$ and $t_x^* < w$)

While participating buyers might be contributing profits that relax the platform's budget constraint, they do not increase the share of mechanical services purchased in the market: only the set of switchers corresponds to increased social welfare. The figure also illustrates the switching function $\sigma(\eta, x)$: the hazard rates $h_w[t_x^*|x]$ and $h_r[t_x^*|x]$ are the measures of households along the boundaries from switching from manual to mechanical and from the market to the platform, respectively. The switching function is the probability of being on the boundary from switching from manual to mechanical, given a report of $\hat{\eta} = t_x^*$. This taxonomy will form the basis of our empirical design strategy in Section 4.

The optimal price in (8) can be decomposed as

$$\underbrace{t_x^*}_{\text{Price}} = \underbrace{c_x}_{\text{Marginal cost}} + \underbrace{\frac{1 - F_\eta[t_x^*|x]}{f_\eta[t_x^*|x]}}_{\text{Informational Rent}} - \underbrace{\frac{\sigma(t_x^*, x) b_x}{\lambda^*}}_{\text{Social Discount}} .$$

A standard monopolist would set its price equal to marginal cost plus the informational rent, but the platform is instead maximizing take-up of the mechanical service, reflected in the social discount term. Each observable type x has some probability of switching on the margin and contributing a social benefit b_x, which is then deflated by λ^*, representing the opportunity cost of providing assistance to this type x over some other type x'. The households with the largest discount are those who deliver the largest social benefit b_x and are most likely to switch, quantified by $\sigma(t_x^*, x)$. At the optimum, some households receive assistance while others pay into the system, and an observable type x is *subsidized* if

$$t_x^* - c_x = \frac{1 - F_\eta[t_x^*|x]}{f_\eta[t_x^*|x]} - \frac{\sigma(t_x^*, x) b_x}{\lambda^*} \leq 0,$$

and *profitable* otherwise. Subsidization occurs in our application when the shadow price of the subsidy is greater than the hazard rate of the household's willingness to pay, so that

$$\lambda^* < \frac{f_\eta[t_x^*|x]}{1 - F_\eta[t_x^*|x]}\sigma(t_x^*, x) = h_w[t_x^*|x],$$

and otherwise x pays into the system. If the hazard rate at t_x^* is greater than the shadow price of the subsidy, the platform's social motive dominates and every such x receives a price below the cost of procurement. If the inequality is reversed, the profit motive dominates and every such x pays into the system and relaxes the budget constraint. This concept of subsidization will play a key role in evaluating the efficacy of different market design in the counterfactual analysis of Section 6.

This analysis of the optimal design highlights some drawbacks of commonly used mechanisms. An auction selects households entirely on the basis of willingness-to-switch, η, ignoring x. This will pass on cost savings and subsidies to relatively rich households at the expense of relatively poor households. Similarly, proxy means testing fixes a flat price t and seeks to exclude relatively wealthy households on the basis of x. This eliminates the possibility of exploiting profitable types and engaging in cross-subsidization and fails to adjust the level of aid provided to poor households to ensure take up of the healthier service. The optimal mechanism synthesizes features of both of these schemes to reach a superior alternative, targeted pricing.

4 Empirical Platform Design

Section 3 derived the structure of the optimal mechanism, but it relies heavily on how observables translate into willingness-to-switch values and propensities to purchase mechanical in the absence of the platform. In this section, we describe how market and experimental data were gathered and used to operationalize the optimal design. The process has two steps.

In the first step, we administer a baseline survey to a random sample from the target population, gathering market data on each household's most recent transactions in the search market and measuring each household's willingness-to-switch through a demand elicitation game. These data jointly form the basis of the the platform's beliefs about the distribution of prices that a household of observable type x faces and the probability that such a household would purchase mechanical on its own. One could then posit and estimate a structural model of decision-making to predict how a household might respond to a counterfactual change in the mechanical price it faces, but this would be sensitive to modeling assumptions about how the household searches for service operators and how negotiations proceed. We adopt a simpler approach: ask households their willingness-to-switch, η, in an incentive compatible game similar to a second-price auction, and exploit its correlation with past decisions, prices, and observables to estimate a household's probability of purchasing on the platform or the search market (see Figures 2 and 3, and page 18). This allows us to assign a probability to any given household of being a non-buyer, participating buyer, non-participating buyer, or switcher, conditional on price.

In the second step, we select prices that maximize the share of households that purchase mechanical services subject to an expected budget balance constraint. Since the sample on which the prices are designed is large and random, the prices are also approximately optimal with respect to the population overall. We then deploy the optimal platform on a new sample of households, with results reported in Section 5.

The Market Survey and Demand Elicitation Game were administered in December 2014, with $2,088$ participant households selected based on their proximity to 67 randomly selected grid points from 450 grid points evenly spaced across Ouagadougou[15]. Prior to randomization, grid points falling in the wealthiest neighborhoods, neighborhoods that were connected to the sewer system, and neighborhoods in which property rights are not well-defined were omitted. Enumerators were sent to map the households closest to the grid points prior to the survey, and households were randomly selected for participation in the survey from the mapped households near the 67 selected gridpoints. Because the demand elicitation game is a generalization of the second-price auction to allow for multiple winners, we call this the *Demand*

[15]Another 52 were reserved for the Treatment group and 40 for the Control group.

Elicitation group.

During the Market Survey, we gathered household characteristics x_i that would be available to a local municipal authority like ONEA as well as information on their most recent desludging. This information includes whether they purchased mechanical, $y_i = 1$, or manual, $y_i = 0$; the mechanical price if they purchased mechanical $r_{mech,i}$; and the manual price if they purchased manual, $r_{man,i}$. We model the determination of the manual and mechanical prices in the market and the household's decision as a Type V Tobit or an endogeneous regime switching regression[16] :

$$\tilde{y}_i = x_i\delta + \varepsilon_{0i} \tag{9}$$

$$r_{mech,i} = \begin{cases} z_i\beta_{mech} + \varepsilon_{mech,i}, & \tilde{y}_i \geq 0 \\ \varnothing, & \tilde{y}_i < 0 \end{cases} \tag{10}$$

$$r_{man,i} = \begin{cases} \varnothing, & \tilde{y}_i \geq 0 \\ z_i\beta_{man} + \varepsilon_{man,i}, & \tilde{y}_i < 0 \end{cases}, \tag{11}$$

where the latent index, $\tilde{y}_i$, determines selection into manual[17] or mechanical, and the shock $\varepsilon_i = (\varepsilon_{0i}, \varepsilon_{mech,i}, \varepsilon_{man,i})$ is trivariate normal, so that

$$\begin{bmatrix} \varepsilon_{0i} \\ \varepsilon_{mech,i} \\ \varepsilon_{man,i} \end{bmatrix} \sim \text{Normal}\left(\mu = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \Sigma = \begin{bmatrix} 1 & \rho_{0,mech} & \rho_{0,man} \\ \rho_{0,mech} & \sigma^2_{mech} & \rho_{mech,man} \\ \rho_{0,man} & \rho_{mech,man} & \sigma^2_{man} \end{bmatrix} \right).$$

Because a household must search for service providers, only the transaction price for the kind of desludging selected is observed, not the counterfactual price that would have been charged had the household selected the other kind of service[18].

[16]See, for example, (Amemiya, 1985) or (Maddala, 1983).

[17]Why estimate the manual price equation at all? It exploits more decision-relevant data, allowing the selection equation to better rationalize observed choices by incorporating information about the household's perceived outside option, manual.

[18]This is also why we did not use a standard multinomial logit model of demand: for the vast majority of households, the alternative price is not observed and there is no centralized market with stable prices. A structural approach would require instead estimating a search

The purpose of the model is to predict the distribution of prices that a household faces and whether it is likely to purchase mechanical or manual services. But the variables included must also be observable or verifiable: a household should not be able to manipulate, easily adjust, or lie about their value. A platform designed around variables which the household can misrepresent invites them to maximize their appearance of need. For example, the quality of the dwelling is a good candidate for an NGO or non-governmental actor, since it can easily be determined by an enumerator during a baseline interview. For a municipal authority, water and electricity bills are even better candidates, since these are objective, impose no additional data collection costs, and are highly correlated with household income. The variables we use are given in Table 2, and include

i. Information gathered by the enumerator during a household interview: housing type (precarious, concrete, or rooming house), whether other households lived in the compound, the distance from the latrine pit to the road in meters, the number of people living in the household, the number of women, and whether the respondent finished high school

ii. Information available to a municipal authority: whether the water bill was more than 5,000 CFA last month, the previous month's electricity bill in CFA, and whether the household owns the dwelling

iii. Information available to a continuously operating platform that can keep its own records: average number of months between desludgings, whether the last service episode required more than one trip because of the large size of the pit

In Section 6.2, we provide a counterfactual analysis of the information structure using subsets of these variables to determine which variables are the most useful and the consequences of using less or different kinds of information.

model that characterizes distributions of prices and the propensity to continue searching for a mechanical provider or take the option of manual desludging.

To ensure that the model is not identified solely from the functional form of the structural errors[19], we drop the electricity bill, the number of people in the household, the number of women in the household, and whether the respondent completed high school from estimation of the price equations (10) and (11). Our argument that the exclusion restriction is satisfied is based on price discrimination: at the time of contracting, a desludger in the search market might observe many characteristics about the household — particularly related to water consumption and sanitation — and adjust the price to extract rents from the household. The variables excluded from the second-stage are not observable to a desludger, and therefore can not affect price, but do shift the likelihood the household will purchase mechanical desludging: more highly educated household heads are more likely to understand the importance of health and sanitation, women typically value sanitation services at higher rates than men, electricity expenditure is unobserved by a one-time visitor, and larger households incur greater disutility from poor removal of sanitation.

We estimate $(\delta, \beta_{mech}, \beta_{man})$ in equations (9), (10), and (11) by maximum likelihood, and results are reported in Table 2. Measures of wealth like electricity bill, quality of the dwelling, and respondent education have a positive and statistically significant effect on the likelihood of mechanical desludging, while households that desludge more frequently or own their own dwelling are less likely to use mechanical. Households that rent often share their compound with other households (which is positive but not statistically significant), and likely share the cost with their landlord, which explains why households who own their dwelling and are presumably wealthier are less likely to purchase mechanical services. Similar patterns hold for the mechanical and manual price equations. A likelihood ratio test of a restricted model that drops the instruments against the unrestricted model has a test statistic of 126.56, rejecting the hypothesis that the coefficients of the instruments are jointly zero.

While the model predicts the distributions of prices households would face in the decentralized market and how they would select into manual or mechanical desludging in the absence of the platform, it is a reduced-form model

[19] Section 6 includes a case in which no instruments are used.

that is silent about how a household would respond to a price offer from the platform. In particular, such a model does not predict the price at which a household that has selected a manual desludging would switch to mechanical; indeed, prices do not even appear in the equation determining selection, (9). Instead, we supplement the market data with information from a willingness-to-switch elicitation experiment based on the second-price auction.[20] The rules of the *highest-reject bid auction* are:

i. Each household i is told it is facing N competitors, but only $K < N$ will be selected to win a desludging.

ii. Each household i is asked to make an offer, η_i, for a desludging.

iii. The highest K offers are accepted, and all winners are asked to pay the $K + 1$-st (highest losing) price when they purchase a desludging.

Since honest reporting is a weakly dominant strategy in the $K + 1$-st price auction, the offer η_i provides an unbiased estimate of the household's willingness-to-switch, the minimum of their willingness-to-pay and the price they expect to face in the prevailing market[21]. A histogram of the offers received and summary statistics are given in Figure 4 and Table 4.

Do these offers accurately reflect a household's true willingness-to-switch? Before revealing any outcomes in the elicitation game, we conduct a variety of thought experiments with the households to check they understand their incentives to report honestly and the potential for regret if they submit a dishonest bid and lose[22]. In particular, households were asked to confirm that

[20]The script is provided in Appendix A.

[21]Households may have an intrinsically higher willingness to pay, but make lower offers because of credit constraints that constrain their access to funds in the short run (see, for example, Yishay et al. (2017)). While the distinction between willingness- and ability-to-pay is important for understanding potential desludging demand absent these market constraints, we argue that for our purposes, the minimum of the two is what is relevant for maximizing short-run demand.

[22]Our motivation is not to test whether households would play the sincere dominant strategy on their own, but to convince them that it is in their interests to do so. These thought experiments simply help them work out this idea for themselves, before they are committed to their offer.

they would want to purchase a desludging at a price 5% lower than their bid if that was the highest rejected bid; 2% of the households said no. They were also asked to confirm that they would not regret losing the ability to purchase a desludging at a price 5% higher than their offer if the other households were to bid higher than them and they were the highest rejected bid; 18% of households stated that they would regret losing the ability to purchase. Households stating that they would regret their bid were then allowed to revise their bids, before learning the clearing price they faced. The enumerators reported that 99.5% of households understood by the end of the exercise, though 10.5% of households required multiple explanations[23].

There are many potential improvements to this overall approach to modeling and estimating household behavior. Picking observable variables to target is an important step, but they must be available or easily observable to a local governmental entity or NGO, since otherwise the resulting mechanism would not be incentive compatible: households would generally be able to guess how to lie to misrepresent themselves as poorer than they are, or engage in behaviors to hide wealth from enumerators. Ideally, variables used would be highly correlated with wealth and sanitation decision-making, be cheap to gather, and robust to measurement error. With respect to estimation and model selection, penalized regression could be applied to the log-likelihood of (9)–(11), similar to the LASSO, in order to avoid over-fitting or high-variance predictions. With the rapid advances in data-gathering methods and machine learning tools over recent years, we expect there are variety of other improvements that could be made.

With these estimates in hand, we turn to computing household-level demand. The vector of shocks ε_i determines the prices and decision the household would make in the absence of the intervention, conditional on x_i. If the platform quotes a lower price to a household that was already planning on

[23]In addition, the number of winners was randomly assigned, and a regression of offers on the number of winners yields a coefficient of -119 CFA (not statistically significant at conventional levels). Players who do not understand that honesty is a weakly dominant strategy of the game often shade their bids more when they face more opponents, so this provides additional evidence that participants understood the game.

purchasing a mechanical desludging in the decentralized market ($t_i < r_{mech,i}$ and $\tilde{y}_i \geq 0$), that household will be a participating buyer. If the household was not planning on buying a mechanical desludging ($\tilde{y}_i < 0$), then it only switches if $t(x_i) < \eta_i = w_i < r_{mech,i}$, which is determined by the joint distribution of (η_i, x_i). These relationships are summarized in Figure 3, which corresponds to the theoretical framework given in 2. The total demand for mechanical desludgings is then

$$D(t_i, x_i) = \mathbb{E}_{\varepsilon_i, \eta_i} \Bigg[\underbrace{\mathbb{I}\{\tilde{y}_i \geq 0 \cap t_i < r_{mech,i}\}}_{\text{Participating buyers}} + \underbrace{\mathbb{I}\{\tilde{y}_i \geq 0 \cap t_i > r_{mech,i}\}}_{\text{Non-participating buyers}} + \underbrace{\mathbb{I}\{\tilde{y}_i < 0 \cap t_i < r_{mech,i} \cap t_i \leq \eta_i \leq r_{mech,i}\}}_{\text{Switchers}} \Bigg| x_i \Bigg]. \quad (12)$$

In order to derive a tractable way of estimating and computing this quantity, we make the assumption that the joint density of (ε_i, η_i) takes the form $f[\varepsilon_{i0}, \varepsilon_{i,mech}, \eta_i | x_i] = f\varepsilon(\varepsilon_{i0}, \varepsilon_{i,mech}) f_\eta[\eta_i | x_i]$, so that the joint density is the product of a bivariate normal and a distribution that is independent of ε_i, conditional on x_i. This is a statistical and not economic assumption, and an alternative would be to use a semi- or non-parametric approach and dispense with the use of the Normal assumption. But the benefit of the assumption is that we can exploit the tri-variate normal and willingness-to-switch data to express demand as[24]

$$
\begin{aligned}
D(t_i, z_i) = & \underbrace{\int_{-\infty}^{t_i - z_i\beta_{mech}} 1 - \Phi\left(\frac{-x_i\delta - \frac{\rho_{0,mech}}{\sigma_{mech}}\varepsilon_{mech,i}}{\sqrt{1-\rho_{0,mech}^2}}\right) d\Phi_{mech}\left(\frac{\varepsilon_{mech,i}}{\sigma_{mech}}\right)}_{\text{Non-participating buyers}} \\
& + \underbrace{\int_{t_i - z_i\beta_{mech}}^{\infty} 1 - \Phi\left(\frac{-x_i\delta - \frac{\rho_{0,mech}}{\sigma_{mech}}\varepsilon_{mech,i}}{\sqrt{1-\rho_{0,mech}^2}}\right) d\Phi_{mech}\left(\frac{\varepsilon_{mech,i}}{\sigma_{mech}}\right)}_{\text{Participating buyers}} \\
& + \underbrace{\int_{t_i - z_i\beta_{mech}}^{\infty} \Phi\left(\frac{-x_i\delta - \frac{\rho_{0,mech}}{\sigma_{mech}}\varepsilon_{mech,i}}{\sqrt{1-\rho_{0,mech}^2}}\right) \max\{F_\eta[z_i\beta_{mech} + \varepsilon_{mech,i}|z_i] - F_\eta[t_i|z_i], 0\} d\Phi_{mech}\left(\frac{\varepsilon_{mech,i}}{\sigma_{mech}}\right)}_{\text{Switchers}}.
\end{aligned}
$$

These three terms integrate over the regions where households purchase a mechanical desludging derived from the theoretical model, yielding a tractable way of computing overall demand. One can think of the switchers' contribution to the sum as integrating the area under a demand curve between the prices t_i and $r_{mech,i}$, weighted by the household's characteristics and the decision not to purchase in the market. To estimate the conditional probability of

[24] If X and Y are jointly normally distributed random variables with $\sigma_y = 1$, then $Y|X$ is distributed normally, with mean $\mu_y + \frac{\rho\sigma_x}{\sigma_x}(x - \mu_x)$ and variance $(1-\rho^2)$, yielding the conditional distribution $F[y|x] = \Phi\left([y - \mu_y - \frac{\rho}{\sigma_x}(x-\mu_x)]/\sqrt{1-\rho^2}\right)$. We then have, for example for the switchers,

$$
\begin{aligned}
& \mathbb{E}_{\varepsilon_i,\eta_i}\left[\mathbb{I}\{\tilde{y}_i < 0 \cap t_i < r_{mech,i} \cap t_i \leq \eta_i \leq r_{mech,i}\}|\, x_i\right] \\
& \quad = \mathbb{E}_{\varepsilon_i,\eta_i}\left[\mathbb{I}\{x_i\delta + \varepsilon_{i0} < 0 \cap t_i < r_{mech,i} \cap t_i \leq \eta_i \leq z_i\beta_{mech} + \varepsilon_{i,mech}\}|\, x_i\right] \\
& \quad = \int_{t_i - z_i\beta_{mech}}^{\infty}\int_{-\infty}^{-x_i\delta}\int_{t_i}^{z_i\beta_{mech}+\varepsilon_{i,mech}\wedge t_i} f[\varepsilon_{i0}|\varepsilon_{i,mech}] f_\eta[\eta_i|x_i] d\eta_i d\varepsilon_{i,0} d\Phi\left(\frac{\varepsilon_{i,mech}}{\sigma_{mech}}\right) \\
& = \int_{t_i - z_i\beta_{mech}}^{\infty} \Phi\left(\frac{-x_i\delta - \frac{\rho_{0,mech}}{\sigma_{mech}}\varepsilon_{i,mech}}{\sqrt{1-\rho_{0,mech}^2}}\right) \max\{F_\eta[z_i\beta_{mech} + \varepsilon_{i,mech}|x_i] - F_\eta[t_i|x_i], 0\} d\Phi\left(\frac{\varepsilon_{i,mech}}{\sigma_{mech}}\right).
\end{aligned}
$$

The details of the other calculations are similar (and simpler, because they do not include η_i). To compute this, we use a Monte Carlo approach and take a large number of draws from the distribution of residuals of $\varepsilon_{mech,i}$, use the given closed form solutions for the integrand, and average over the results from all the draws.

a switch given the prices and observables, $F_\eta[r_{mech,i}|z_i] - F_\eta[t_i|z_i]$, we need only estimate $F_\eta[y|x_i] = Pr[\eta_i \leq y|x_i]$ for the set of relevant prices (those observed or those we intend to quote), then evaluate at $r_{mech,i}$ and t_i and take the difference. This only requires estimating the probability that η_i is above a given set of thresholds, for which we use a sequence of simple logit regressions, with results reported in Table 3. In practice, we compute the integrals by Monte Carlo integration, re-sampling the distribution of residuals of the mechanical pricing equation regression to approximate the distribution of $\varepsilon_{mech,i}$. Similarly, platform demand is given by

$$D^P(t_i, x_i) = \mathbb{E}_\varepsilon \left[\left. \underbrace{\mathbb{I}\{\tilde{y}_i \geq 0 \cap t_i < r_{mech,i}\}}_{\text{Participating buyers}} + \underbrace{\mathbb{I}\{\tilde{y}_i < 0 \cap t_i < r_{mech,i} \cap t_i \leq \eta_i \leq r_{mech,i}\}}_{\text{Switchers}} \right| x_i \right], \tag{13}$$

computed the same way as (12). This corresponds to the probability that household i with characteristics x_i decides to purchase from the platform, and will play a key role in the constrained optimization problem that determines the prices we quote. Estimated demand is illustrated in Figure 5: panel (a) plots a demand curve for each household in the Demand Elicitation group — illustrating total demand (12) — while panel (b) plots the average curve as well as the average curves by the price bin they are ultimately assigned to. Similarly, panels (c) and (d) plot household demand — illustrating platform demand (13) — and average demand for platform services, respectively.

Equations (12) and (13) highlight again the importance of combining the Tobit estimation with the demand elicitation data, because the prices at which buyers are willing to participate versus not can be estimated from the Tobit model while switchers would be out of sample: those who typically purchase manual would not have market data on their mechanical purchases, and so their prices must be calculated from the auction. In other words, this method allows the platform to explore the shape of the demand curve for prices that are not observed in the market, which is a serious challenge for policymakers when offering novel goods or assistance to poor populations.

With the demand side modeled and estimated, we now turn to the supply side. To procure services and determine the cost of a desludging c_x, we used *lowest rejected bid auctions* in each neighborhood:

i. Each firm j is told it is facing N competitors, but only $K < N$ will be selected to win the right to provide services in this neighborhood for this service period.

ii. Each desludger j is asked to submit an ask, a_j for each desludging they provide to the platform during this service period.

iii. The lowest K asks are accepted, and all winners receive the $K + 1$-st (lowest losing) price for each desludging they provide the platform during this service period.

Whenever a household called in to claim a desludging, we randomly selected one of the K winners to receive the job. We selected $K > 1$ and typically equal to two or three, so that if the first desludger was unavailable, there were other service providers who could take the job. As for households, it is a weakly dominant strategy for firms to bid their expected marginal costs, since all of the winners receive an equal share of the work and are paid the lowest rejected bid for each job completed. These auctions were conducted monthly, there were typically two or three winners selected, and the neighborhood-average clearing price[25] is illustrated in Figure 6. In particular, the average price was 17,500 CFA at the time the Targeted Pricing treatment began, which we take as our average cost, c_x.

We can now exploit Theorem 1: the optimal mechanism is posted prices, selected to maximize the mechanical market share subject to an expected budget constraint. More formally, the platform takes the sample $X = \{x_i\}_{i=1}^N$, the subsidy s, and the average cost of procuring a desludging c_x as given, and maximizes market demand,

$$\max_{t=(t_1,\ldots,t_n)} \frac{1}{N} \sum_{i=1}^{N} D(t_i, x_i) \tag{14}$$

[25] Demand Elicitation and Targeted Pricing households were in disjoint neighborhoods, so that we cannot simply use "neighborhood-level" prices.

subject to

$$0 \leq \frac{1}{N}\sum_{i=1}^{N} D_i^P(t_i, x_i)(t_i - c_x) + s \tag{15}$$

$$t_i \in T = \{8{,}000, \quad 10{,}000, \quad 12{,}500, \quad 15{,}000, \quad 17{,}500, \quad 20{,}000\}. \tag{16}$$

The set of prices T spans the observed transaction prices in the market data, and are the most commonly used denominations for payment[26]. The average subsidy per household (s) was 1,750 CFA or about \$3.00, and c_x was set to 17,500 CFA. Imposing the constraint (16) converts the maximization problem into a linear programming problem where each household i is assigned to a price t_i.

The optimal linear programming solution is presented in Figure 7, panel (a). It turns out that it is never optimal to offer 12,500: this is too high a price to induce a poorer household to switch to mechanical, and too low to relax the budget constraint. Very few (approximately 4%) of households are allocated to the 8,000 CFA bin, which we will ultimately shift to the 10,000 CFA bin[27]. Table 5 provides predicted treatment effects, using the model presented here: it predicts a 10% average treatment effect, and a 28.9 percentage point increase in the 10,000 CFA price bin.

Since the original sample, $X = \{x_i\}_{i=1}^N$, is random, the platform can replace the personalized prices for each household t_i^* with a function that maps characteristics x_i into prices, $t_i^* = t^*(x_i)$, and the same pricing rule should also maximize adoption of mechanical desludging across the population, by the weak law of large numbers:

$$\lim_{N\to\infty} \left\{ \max_{t=(t_1,\ldots,t_n)} \frac{1}{N}\sum_{i=1}^{N} D(t_i, x_i) \right\} \to_{pr} \max_{\{t^*(x)\}_{x\in X}} \mathbb{E}_x[D(t^*(x), x)] \tag{17}$$

[26]From an experimental perspective, having more prices also results in more treatment arms, creating a trade-off between a larger average treatment effect from a more complex price schedule and the statistical power to detect treatment effects for each price bin separately.

[27]In short, statistically separating these households from 10,000 CFA households is difficult and leads to more aggregate error than reassignment.

subject to

$$
\begin{aligned}
0 &\leq \mathbb{E}_x[D_i^P(t^*(x),x)(t^*(x)-c_x)]+s && (18)\\
t_i &\in T=\{8{,}000, \quad 10{,}000, \quad 12{,}500, \quad 15{,}000, \quad 17{,}500, \quad 20{,}000\}. && (19)
\end{aligned}
$$

While the solution to the linear program (14) — (16) is in terms of individual households, $\{(t_i,x_i)\}_{i=1}^N$, the solution to (17) — (19) is a mapping from observables x_i to prices, $t^*(x_i)$. Because we wish to apply the optimal rule to a new set of households, we must convert the first kind of solution into the second[28]. To do this, we use an ordered logit model, mapping household characteristics x_i to a latent index $\tilde{t}_i$, and then using the index to assign households to price bins. Because no households were allocated to the $12,500$ CFA bin and only 4% of households were allocated to the $8,000$ CFA bin, we shifted them to the $10,000$ CFA bin[29] and fit the index

$$\tilde{t}_i = x_i'\gamma + \varepsilon_{t,i},$$

[28]A promising alternative we considered was to parameterize a rule $t(x_i,\gamma)$ rather than choose personalized prices $\{t_i\}_{i=1}^N$, and optimize over γ directly rather than fit a mapping $t(x_i,\hat{\gamma})$. We think this approach has potential (and potentially better stochastic convergence properties), but is much more computationally intensive and sensitive to the specification of the mapping, since linear programming is a very robust solution method while non-linear programming is generally not.

[29]In addition, ordered logit and random forest models struggle to fit sparsely populated bins, particularly when they are on the interior of the set of prices offered, creating more error overall than if they are simply eliminated. In general, regression methods like OLS tend to predict very similar prices for all households clustered around the average price offered, necessitating the use of classification methods like ordered logit or random forest in order to create price dispersion and populate very low or high bins.

by maximum likelihood; results are reported in Table 7. Letting

$$\begin{aligned}
\pi_{10,000}(x_i) &= Pr[x_i'\hat{\gamma} + \varepsilon_{t,i} < 10,000] = \frac{e^{10,000 - x_i'\hat{\gamma}}}{1 + e^{10,000 - x_i'\hat{\gamma}}} \\
\pi_{15,000}(x_i) &= Pr[10,000 \leq x_i'\hat{\gamma} + \varepsilon_{t,i} < 15,000] = \frac{e^{15,000 - x_i'\hat{\gamma}}}{1 + e^{15,000 - x_i'\hat{\gamma}}} - \frac{e^{10,000 - x_i'\hat{\gamma}}}{1 + e^{10,000 - x_i'\hat{\gamma}}} \\
\pi_{17,500}(x_i) &= Pr[15,000 \leq x_i'\hat{\gamma} + \varepsilon_{t,i} < 17,500] = \frac{e^{17,500 - x_i'\hat{\gamma}}}{1 + e^{17,500 - x_i'\hat{\gamma}}} - \frac{e^{15,000 - x_i'\hat{\gamma}}}{1 + e^{15,000 - x_i'\hat{\gamma}}} \\
\pi_{20,000}(x_i) &= Pr[x_i'\hat{\gamma} + \varepsilon_{t,i} > 17,500] = 1 - \frac{e^{17,500 - x_i'\hat{\gamma}}}{1 + e^{17,500 - x_i'\hat{\gamma}}},
\end{aligned}$$

the assignment rule is

$$t^*(x_i) = \underset{t \in \{10000,15000,17500,20000\}}{\text{argmax}} \pi_t(x_i),$$

mapping x_i to the most likely bin under $\hat{\gamma}$. This resembles proxy means testing, but is constructed not by maximizing a classification target like the fraction of households in a training data set below the poverty line, but instead by approximating the optimal pricing schedule.

Figure 7 illustrates the linear programming and ordered logit pricing rules in the left panel, and the propensity for mis-classification in the right panel, and average deviations of the ordered logit from the linear programming price given in Table 6. The dark bars represent the optimal pricing rule and the light bars represent the proportion of these price quotes in the ordered logit approximation. The ordered logit tends to make too few 10,000 offers and too many 17,500 offers, but is correct approximately 79% of the time, and within 2,500 CFA of the correct bin 92% of the time. Conditional on the bin, Table 6 shows that the ordered logit rule tends to overcharge households assigned to the 10,000 bin by 440 CFA on average, while it tends to undercharge households assigned to the 20,000 bin by 2,000 CFA on average. Consequently, mis-classifications by the ordered logit rule should be expected to attenuate the treatment effect and lead to higher budget deficits than expected. We consider alternative, more algorithmic and automatic methods of assigning observables to prices in Section 6.

Our intervention is then to offer the pricing rule $t = t^*(x_{i'})$ to a new sample, $X' = \{x_{i'}\}_{i'=1}^{N'}$, under the same circumstances: a household survey is conducted, the results are recorded on a tablet computer, and in the background $x_{i'}$ is used to compute a price $t_{i'} = t^*(x_{i'})$. Take up of this targeted price group is then compared to the take up of a control group in a randomized controlled trial. We refer to this group as the *Targeting Pricing* treatment group, which is the focus of the next section.

How might the process we have just described be done at scale? Of particular concern is that a household who anticipates that its reports of η_i and $(r_{mech,i}, r_{man,i})$ will determine its future payoffs has an incentive to lie. There are straightforward ways to avoid this. For a given period of the program, a small, randomly selected subset of a large population can be drawn and offered the chance to participate in the market survey and game, just like the Demand Elicitation group. This period's market will be designed using information from this subset, but the subset will not be subject to the terms of the program this period. This maintains incentives for honest reporting, both within each period and across periods.

5 Experimental Results

We returned to Ouagadougou in August and September of 2015 in order to test the Targeted Prices treatment designed in Section 4. During the baseline survey, we collected basic information on the household. As the enumerators entered information into the tablets, the program calculated the price that the household should be offered according to the pricing algorithm. The pricing treatment was offered to the treatment households at the end of the baseline survey. The tablet provided treatment households with the price that the enumerator was to offer according to the targeted prices treatment. If a treatment household wished to accept this price, they were asked to pay a nominal deposit of 500 CFA (which was the same as the participation payment for the survey). They were then able to call in to the center at any time in the following 15 months in order to claim their desludging at the targeted price offered. Our endline survey took place in December of 2016, at which point we collected information from the households on any mechanical or manual desludgings they had purchased over the past 15 months.

5.1 Data

We ran three household surveys for this project: the demand elicitation survey, the baseline survey, and the endline survey. In addition, we collected data on the cost of desludgings provided through our call center, the bidding behavior of trucks in our call center, and the timing of calls by households. Each data source is described in turn below:

Selection for household surveys: We placed 450 evenly spaced grid points across Ouagadougou, and randomly selected 67 for the Auction survey, 52 for the Treatment group and 40 for the Control group.[30] Prior to randomization, grid points falling in the wealthiest neighborhoods, neighborhoods that were connected to the sewer system, and neighborhoods in which property rights are not well-defined were omitted. Enumerators were sent to map the households closest to the grid points but within 100 meters of a grid point. Households were randomly selected from among mapped households for inclusion into the project. Households without latrines were excluded from the survey. Each neighborhood cluster point included approximately 30 households.

Demand Elicitation Survey: The demand elicitation survey and the incentive-compatible demand elicitation game were administered in December 2014, with 2,088 participant households selected based on their proximity to the 67 randomly selected grid points selected for the demand elicitation treatment. The survey collected data on household choices related to sanitation, and at the end of the survey households were asked to participate in a demand elicitation auction in which they were asked to bid on a desludging in a K+1 price auction developed to elicit their incentive-compatible willingness to pay for a

[30] We selected more gridpoints for the auction survey in order to ensure that we would have adequate representation of different consumer types for generating the model. We increased the size of the treatment group by randomly selecting additional gridpoints after a bug in the algorithm on the enumerators' tablets was detected during the baseline.

mechanical desludging from the call center (the auction is further explained in section 4 and the script for the demand elicitation game is given in appendix A). The data collected for this survey was used primarily to inform the design of the pricing system.

Baseline Survey: The baseline survey took place in August and September of 2015, with 1,660 pricing treatment households and 1,284 control group households. During the survey, households were asked about their latrine pit, their sanitation practices, the process of search for a desludging operator, and their level of wealth.

Table 8 presents the balance of the treatment versus control groups, both at the cluster level and the household level. Balance is not perfect between the treatment and control groups, but the control group appears to be somewhat more wealthy than the treatment group (the average principal components index for the control group is significantly higher than both the treatment and auction group, and they are more likely to spend more than 5,000 CFA on water bills). We control for baseline variables that are not well balanced in the OLS specifications in our main regressions. We can also see from column (3) in the table that the balance between the control group and the demand eliciation group is not perfect, and again the control group is wealthier than the demand elicitation group. To the extent that the demand elicitation group was not fully representative of the treatment and control groups for the randomized controlled trials, we would expect that the pricing model would not perform as well, biasing our estimates toward 0.

Table 9 shows the balance between control and treatment for each of the price groups since we are interested in the difference in impacts across price groups. We find that in this case the balance is problematic on similar variables to those that were highlighted in the overall balance tests. In order to be sure that the key variables for balance have been selected, we also run a LASSO specification for each regression with 119 potential control variables on demographic characteristics and past use of mechanical and manual desludgings from the baseline survey.

Endline Survey: We returned to the households interviewed in both the Demand Elicitation survey and the Baseline survey in December of 2016. During the endline survey, households were asked about their latrine pit, their sanitation practices since the Baseline or Demand Elicitation survey, their search process for a desludging operator for any desludgings that they had done over the period, the diarrhea related health of their children, and their level of wealth. We find that households in the Control group used more desludgings total over the course of the study than households in the treatment group, which biases our estimates toward zero since the market share of mechanical desludgings is 84%. We discuss this further in Appendix B.

Call Center Data: We also collected data in our call center. This included data on the supply side: the bids made by desludging operators and the winners in each month, and data on the demand side: which households called for a desludging and the date on which they called. Deposit rates by price offered and use of the call center are shown in table 10. Use of the call center was somewhat lower than predicted, but among those who purchased a desludging in the first 6 months and deposited, use of the call center was quite close to the level expected from the model (and somewhat higher among the 20k price group).

At endline we asked households that deposited but did not call the call center why they had not called in, and the responses are shown in table 11. More than half stated that they had simply not needed a desludging during the study period. Many of the others forgot about it or found a better outside option. One attribute of the platform system is that by giving households a price offer, we helped them to better negotiate with their desludging operators. Much of the impact of the call center may have been through improving the negotiating power of households with desludging operators outside of the system. A simple regression comparing the average price paid for desludgings by treatment households who purchased a desludging outside of the system to the prices paid by control households shows that average prices were significantly lower (by about $2 on average) for households that were assigned to the treatment group. Households may have stayed with the desludging operator that they knew, but they were able to purchase at better prices and avoid switching to manual desludging.

5.2 Impact on Mechanical Utilization

For sanitation goods, social benefits accrue to the neighborhood when a household chooses an improved service over the (traditional) unimproved service; in this case mechanical over manual desludging. The objective is therefore not to increase the overall number of desludgings which is determined by the rate of fill of the tank, but to convince households to switch from manual desludging to mechanical. This is in contrast to many previously studied settings where the objective is to increase the overall level of purchases of some under-utilized good (for example, water purifiers, chlorine tablets or insecticide-treated mosquito nets). To capture the impact on mechanical utilization at both the neighborhood and household levels, we use three measures of impact: the market share of mechanical services in the neighborhood, the percent of household purchases which are mechanical, and whether a household purchased any manual or mechanical desludgings over the period of the study. We discuss each in turn below.

First, given that the externalities from manual desludging are a local phenomenon that impacts nearby households, our primary measure of the success of the program is based on the amount of switching between manual and mechanical desludgings at the neighborhood level: the *market share of mechanical services*. Market share[31] is defined as

$$Share_n = \frac{Mechanical_n}{Mechanical_n + Manual_n} \tag{20}$$

where $Mechanical_n$ and $Manual_n$ are the numbers of mechanical and manual desludgings done in neighborhood n, for each of 92 neighborhoods of 25-40 households during the intervention period. Each household that switches from manual to mechanical represents a reduction in fecal sludge in the environ-

[31] Market share is a common outcome variable in papers estimating market effects, particularly when estimating the coverage of a certain product (see, for example, Jensen and Miller (2017) or Nevo (2001)).

ment, so $Share_n$ is the best measure of the impact of the intervention at a neighborhood level, where the effects of the negative externalities are largest.

Second, we provide results for the household-level analog of market share, *percentage of mechanical purchases* during the intervention period:

$$PctMechanical_i = \frac{Mechanical_i}{Mechanical_i + Manual_i} \tag{21}$$

where $Mechanical_i$ and $Manual_i$ are the numbers of mechanical and manual desludgings done by household i, respectively, for each of the 1,199 households during the intervention period who purchased at least one desludging. While market share captures the consequences of access to the call center at the neighborhood level, $PctMechanical_i$ captures impact on mechanical utilization within the household, which is more closely related to the private benefits accruing to members of Targeted Pricing group households.

Third, we provide results for whether a household i purchased *any manual* desludgings:

$$AnyManual_i = \mathbb{I}\{Manual_i > 0\}, \tag{22}$$

for each of the 1,199 households during the intervention period who purchased at least one desludging. To the extent that even one manual desludging can have large, persistent, negative consequences at the household level, this captures the effectiveness of the program in eliminating the socially undesirable behavior. For completeness, we also provide results on

$$AnyMechanical_i = \mathbb{I}\{Mechanical_i > 0\}, \tag{23}$$

which is a more optimistic measure of household-level impact.

For each of these utilization measures $Share_n$, $PctMechanical_i$, $AnyManual_i$, or $AnyMechanical_i$, we provide two sets of regression results. First, we estimate the overall impact of the call center using a pooled average treatment effect β with the specification

$$y_i = \alpha + \beta TargetedPricesTreatment_i + \gamma' X_i + \varepsilon_i, \tag{24}$$

where $TargetedPricesTreatment_i$ takes the value 1 if i received the Targeted Prices treatment and 0 if it is in the Control group, X_i is a vector of control variables including the variables not balanced at baseline, the stratification variable, and the baseline values of variables related to the outcome,[32] and ε_i is a disturbance, clustered at the neighborhood level for Pct_i and $AnyManual_i$.[33] For household-level regressions, standard errors are clustered at the neighborhood level.[34]

Second, we estimate the effect of the call center on mechanical utilization by each price bin k taking values in the set $\mathcal{P} = \{10000, 15000, 17500, 20000\}$

[32]The control variables include: the distance from the latrine pit to the road, whether more than one trip was required to desludge the last time a desludging was done, whether the water bill cost more than 5,000 CFA (approximately \$9), whether the household has more than one pit, and whether there are unrelated households living in the compound and a principal components index of wealth where the principal components variable was constructed using indicators for whether the household owns a refrigerator, motorcycle, car, mobile phone, air conditioner, television with video recorder, whether the household owns or rents, the amount of phone credit the respondent uses in a week, the number of rooms in the house, and an index for roof quality. The stratification variable is whether the neighborhood had an above-median number of low walls. The controls for desludging behavior prior to the baseline for ANCOVA specification (McKenzie, 2012) include: whether the last desludging was manual or mechanical, whether the household had ever desludged in the past, and the percent of desludgings done prior to the baseline that were mechanical.

[33]In the market share regressions, the specification of interest is:

$$y_n = \alpha + \beta TargetedPricesTreatment_n + \gamma' X_n + \varepsilon_n, \tag{25}$$

the targeted prices treatment indicator $TargetedPricesTreatment_n$ takes the value of 1 if the neighborhood n is one of the neighborhoods randomly selected for treatment. Control variables are averaged at the neighborhood level.

[34]A programming error on the enumerators' tablets led to some households being offered prices higher at baseline than the model had predicted. The error occurred at 10.9% of households, and 27 of the 52 treatment neighborhoods. Nearly all of the households receiving incorrect prices received prices that were too high by 1 price bin. In cases in which the household received a price that was too high, we returned to the household to offer them the correct price, and if they had initially rejected the price offer they were given the opportunity to purchase. In the specifications in this paper, we use the price bin to which the household would have been assigned by the correct pricing system, even though this biases our results toward 0. The results controlling for which households were given a different price in error are available on request.

with the specification

$$y_i = \sum_{k \in \mathcal{P}} \alpha_k PriceGroup_{ki} + \sum_{k \in \mathcal{P}} \beta_k TargetedPricesTreatment_i \times PriceGroup_{ki} + \gamma' X_i + \varepsilon_i \tag{26}$$

where $PriceGroup_{ki}$ takes the value 1 if i is in the Targeted Prices treatment and is offered a price of k, or if i is in the Control group and *would have been offered* a price of k. Each coefficient β_k then captures the impact of being assigned to the Targeted Price group for a household with the characteristics that would place them in a particular price bin relative to remaining in the Control group and receiving no price offer. In the neighborhood level regressions, the dependent variable is the market share for a price group within a neighborhood cluster: the market share is calculated as the number of mechanical desludgings purchased by households of that price group in that neighborhood (k equals $10,000$, $15,000$, $17,500$, or $20,000$) divided by the total number of desludgings purchased by households of that price group in that neighborhood. We omit the constant in order to include indicator variables for each price group.

In addition to estimating the OLS model with controls for unbalanced variables, we also estimate the models using the post-double-selection LASSO (Belloni et al., 2014; Ahrens et al., 2018) in order to allow the model to flexibly control for any pre-existing differences at baseline between the control and treatment neighborhoods. The LASSO (Least Absolute Shrinkage and Selection Operator, (Tibshirani, 1996)) regression selects control variables from the full set of potential controls in order to minimize the potential for either over or underestimating the effect size. We include 119 potential control variables[35] from the baseline survey as potential controls, from which the LASSO algorithm selected five in the neighborhood level regression and 4 in the household level regression, and 38 in the household level regression controlling for

[35]In cases in which an observation of a variable is missing, either because the respondent declined to answer or the respondent did not know the answer, the missing observation was replaced with the mean value and an indicator variable was included which takes a value of 1 when the observation is missing and 0 otherwise. Each of the non-binary control variables has been standardized by subtracting the mean and dividing by its the standard deviation.

price-level interactions with the treatment variable.

5.2.1 Market Share

We first provide results for neighborhood market share of mechanical desludgings. Table 12 presents estimates for $y_i = Share_i$ for both the average specification (24) and by price level (26). We also provide the endline market shares in the Control group by the price group the households *would have received* if they were in the Targeted Prices treatment for comparison; in the absence of the Targeted Prices treatment, we would expect the same outcomes for the treatment neighborhoods.

The Targeted Prices treatment generates an increase in the neighborhood market share of the mechanical service of 4.4 percentage points in the OLS regression (significant at the 10% level) and 3.9 percentage points in the LASSO regression (also signficant at the 10% level). This is a 5.2% effect at the mean mechanical desludging market share of 84%. Breaking the treatment effect out by price bins as in specification (26), reveals that almost all of the impact is concentrated in the 10,000 CFA bin, where the OLS regression coefficient is 9.6 percentage points (significant at the 10% level) and the LASSO 7.9 percentage points (significant at the 10% level), while the other bins exhibit small and statistically insignificant effects.

This pattern shows that the Targeted Prices intervention is working as intended by providing aid to the poor households who are unlikely to purchase mechanical on their own. Control group households that would have been quoted relatively high prices are already purchasing mechanical desludgings at high rates, so there is little scope for the intervention to increase utilization of the healthy service: 99% of those households in the control group who would have been quoted a price of 20,000 CFA purchased mechanical services anyway, 91% for those quoted 17,500 CFA, and 85% for those quoted 15,000 CFA. The majority of the potential switchers are in the 10,000 CFA bin, where control group mechanical utilization is only 59%. The average treatment effect in the pooled regression of 4.4 percentage points is driven by the impact in the 10,000 CFA bin, which has a treatment effect of 9.6 percentage points.

This pattern appears in the majority of our results, illustrating how the Targeted Prices intervention successfully implemented the key elements of the theoretically-motivated design. This also highlights that the treatment effect must be coming from the change in market share for the low price groups, not a change in behavior by the high-price groups which already include primarily *participating buyers*.

5.2.2 Percent Mechanical

The market share results demonstrate that across the neighborhoods, households substitute from manual to mechanical desludging when they have access to the targeted price treatment on average, reducing the level of negative externalities. To measure impacts on individual household behavior, we focus on the analogous measure, percent mechanical, $PctMechanical_i$. The same qualitative patterns appear: the average treatment effect is driven by take-up in the 10,000 CFA bin, where potential switchers are concentrated.

Table 13 presents the estimates for $y_i = PctMechanical_i$ for specifications (24) and (26). The sample is limited to households for which percent mechanical is defined: those households which purchased at least one desludging during the sample period. We find that in the pooled regression, the percent of mechanical done increases by 3.2 percentage points (not statistically significant) and 3.0 percentage points in the post-double selection LASSO (not statistically significant). The reduction in the average treatment effect is a consequence of the regression averaging the effects across all groups; the treatment effect is only expected to change the percent of desludgings done that were mechanical on the 10,000 CFA price group, which only accounted for 27% of the sample. Specification (26) separates the average effect by price bin, given in columns (3) and (4) of Table 13. The impact on the 10,000 CFA households is 7.4 percentage points (not statistically significant) in the OLS regression, and 8.5 percentage points (significant at the 10% level) in the LASSO. The other price groups have lower estimates of the impact of the treatment, and none are statistically significant. This is a large effect. In the control group, mechanical accounts for 59% of the desludgings in the 10,000 price group but 85% of the desludgings in the 15,000 price group. The 8.5 percentage point impact on market share in the 10,000 price group in the Targeted Prices treatment bridges 33% of the 26 percentage point gap between the market share of mechanical desludging in the 10,000 price group and that in the 15,000 price group.

Households in the control group purchased more desludgings on average, which also created a bias toward 0 in our estimates since the market share of mechanical desludgings was 84%. Desludging frequency is closely related to the height of the water table and rain levels, so neighborhood level sampling may have meant that we randomly selected control neighborhoods which were more likely to get more desludgings. This is discussed more in appendix B.

5.2.3 Any Mechanical and Any Manual

To the extent that even one manual desludging can generate negative health impacts for household members and neighbors, we are also interested in the measure of whether Targeted Prices households purchased any manual desludgings during the project period. We also present results for whether they purchased any mechanical desludgings, because the treatment allowed for a subsidy only on the first mechanical desludging that the household purchased (ie we should have an impact on the first desludging, but possibly not on additional desludgings they did during the time period). In both the regressions on *Any mechanical* and *Any manual*, we restrict the sample to households which purchased at least one desludging: this is done both to maintain a consistent sample size with the regressions on percent mechanical and to avoid downward bias from the households which did not need a desludging over the time period.

Table 14 presents results for $AnyMechanical_i$ and $AnyManual_i$, for specifications (24) and (26). The overall probability that a household purchases a mechanical desludging increases for the targeted prices treatment group by 3.3 to 3.5 percentage points (not statistically significant), and manual decreases at similar rates (2.3-2.7 percentage points, not significant). Similar to the results for the percentage of desludgings that were mechanical, effects are largest in the lowest price bin: the probability that a household purchases a mechanical desludging if they purchase any desludgings increases by 7.0 percentage points (not significant) in the OLS regression and 8.2 percentage points in the LASSO regression (significant at the 10% level). The probability that a household in the 10,000 CFA price bin purchases a manual desludging decreases by 7.6 percentage points (not statistically significant), 8.2 percentage points in the LASSO specification (significant at the 10% level).

5.3 Health Impacts

The ultimate goal of the program was to reduce the use of manual desludging in order to improve local health and sanitation conditions. We test the impact of the targeted subsidies on child diarrhea rates. We focus on children for two reasons. First, their health is most sensitive to environmental conditions, and they are most likely to be affected by bouts of diarrhea resulting in developmental disadvantages. Second, if the intervention improves the health of children, adults will also benefit, even if these effects are more difficult to observe because of the more disparate environmental factors to which adults have exposure.

We use the specification from (24) to estimate the pooled effect. Observations are at the household level and standard errors are clustered at the neighborhood-cluster level (92 clusters in total). The vector of controls X_i includes the same variables as in our main specifications: variables unbalanced across neighborhoods at baseline, the stratification variable, and an indicator for whether the household was offered an incorrect price in the baseline. We also control for whether the household had a child suffering from diarrhea at baseline for the ANCOVA specification to increase efficiency, following McKenzie (2012). Results are shown in table 15 columns (1) and (2). We are under-powered to find an effect in the pooled regression, but the point estimate on children's diarrhea overall is a 1.4-1.5 percentage point decrease (not statistically significant). At the mean of 13.1% of households reporting that at least one of their children had diarrhea in the last week at baseline, this is a 10.6 percent effect (not statistically significant).

We are also interested in the different effects across the price groups. We run the following specification from (26) to estimate the differential effects on the lowest price group, controlling for baseline diarrhea rates. We use the same control variables and clustering as for the previous regression. Results are shown in table 15, columns (3) and (4). The 10,000 CFA Target Pricing households are 6.0 percentage points less likely to report that a child in their household had diarrhea than a low price household in the control group (significant at the 10% level in the OLS regression; in the LASSO regression the coefficient is -4.6 percentage points, but not statistically significant). Diarrhea rates are higher among children in the households that receive the most subsidized prices: among highly subsidized households, 18.8 percent of households reported their children having had diarrhea in the past week in the control group. At the mean, the OLS specification suggests a 32 percent effect on these households relative to the control group.

Note that there is an important timing difference which attenuates the measured impact: the survey question asks about diarrhea in the past week, while desludging by the household or households in the neighborhood could have been done at any time over the 15 months of the program. The primary impact of the manual desludging on health will be in the first weeks after the desludging is done, before the sludge dries. The effect of desludging choices made by the household and its neighbors just prior to the endline should therefore have substantially more impact than the choices made in the beginning of the time period. We therefore expect any effects on children's diarrhea to be a lower bound.

Children contract illnesses based on the sanitation conditions of the neighborhood, not just the sanitation decisions made by the household. This motivates an analysis of the spillover effects of sanitation decisions within the neighborhood based on the percent of households in the neighborhood that fall into each of the price groups. We use the following specification to estimate the differential effect of the treatment in neighborhoods with larger numbers of each of the price groups, with price k taking values in the set

$\mathcal{P} = \{10000, 15000, 17500, 20000\}$:

$$AnyChildrenDiarrhea_i = \sum_{k \in \mathcal{P}} \alpha_k PriceGroup_{ki} + \sum_{k \in \mathcal{P}} \Theta_k PctPriceGroup_{ki} + \sum_{k \in \mathcal{P}} \beta_k TargetedPricesTreatment_i * PctPriceGroup_{ki} + \gamma' X_i + \varepsilon_i, \quad (27)$$

we use the same control variables and clustering as in the main regression.

The results for the spillovers model are presented in table 15, columns (5) and (6). The diarrhea effects of the treatment are concentrated in neighborhoods with a higher percentage of households placed in the 10,000 CFA Targeted Pricing group. We can see that a 10% increase in the number of households receiving the 10,000 CFA price (about 3 households with our neighborhood sizes of 25-40), leads to a 2.1-2.2 percentage point reduction in reports of a child with diarrhea among households in the neighborhood (statistically significant in OLS and the LASSO at the 5% level). Recall that at the baseline mean of 13.1 percent of households reporting an episode of child diarrhea, this is a substantial effect. Note that the interpretation of this effect is different from that of a typical randomized controlled trial in which prices have been randomized across households: this is the effect of having the targeted prices intervention in a neighborhood in which three more households are poor relative to not having the Targeted Prices intervention in an otherwise similar neighborhood.

We compare the effects found in this paper to those reported in Fewtrell et al. (2005), a large epidemiology meta-study of the impacts of water and sanitation interventions on diarrhea rates in children. They compare relative risks of falling ill with a specified disease for the treatment group versus the control group. They could find only 4 sanitation studies, and report an average relative risk ratio[36] following sanitation treatments of 0.68. As expected from the point estimates, the relative risk ratio for our pooled sample is 0.89 which is close to 1, which suggests little impact in the pooled sample as a whole. However,

[36]The relative risk ratio is the ratio of the rates in the control and treatment groups: $\frac{Outcome_{Treated}}{Outcome_{Control}}$

when the sample is constrained to the households which would receive a price of $10,000$, the relative risk ratio for this group is 0.64. This is a large effect: the diarrhea rates are self reports of households over diarrhea in their children under 12 in the past week, while manual desludgings in their neighborhood could have taken place at any time during the treatment period. Desludging a toilet to restore it to working order and avoiding improper disposal through mechanical desludging therefore has similar impacts to providing households with toilets at the low end of the income distribution.

5.4 Who Receives the Subsidies?

The goal of targeting on observables was to ensure subsidy dollars reached the households who are the cheapest to convert from manual to mechanical, but who are these households? Indeed, targeting the poorest households might not be optimal: if the poorest households were "too poor to help" given the subsidy budget, the platform would instead have allocated them to *higher* prices in order to capitalize on any participation by them, and instead directed subsidies to a wealthier segment of the population more likely to take up. This section examines how *non*-targeted variables vary across pricing bins to shed more light on how subsidy dollars were distributed.

We can observe the extent to which the model is targeting relatively poor households who are more likely to get manual desludgings in Table 16. Households that receive a price of 10,000 CFA (approximately $20, and subsidized by approximately $10) spend an average of 2,200 CFA per week on phone credit while households that receive a price of 17,500 spend nearly twice that, an average of 4,512 CFA for those receiving 17,500 and 5,631 per week for those receiving 20,000 CFA. On average, approximately one quarter of the households receiving a price of 10,000 CFA have a refrigerator, while households receiving 20,000 CFA as their price have on average 1.5 refrigerators. Motorcycles are the most common type of transport in Ouagadougou, and we see that again the households receiving the largest subsidies have fewer motorcycles on average (1.8) than the households receiving no subsidies (2.2

on average for those receiving a price of 17,500 and 3.1 for those receiving a price of 20,000 CFA). We see very similar trends for other asset markers of wealth: cars, televisions, mobile phones, and air conditioners. We see similar differences in terms of key summary statistics about the household's use of desludging services. Households in the highest subsidy group get desludgings the most infrequently (just under four years between desludgings, while households in the 20,000 CFA price group get desludgings just less than once per year).

Examining households' expectations of their future purchase behavior at baseline, we also see lower expected usage of mechanical desludging among those in the most highly subsidized price group: 80% of those in the 10,000 CFA price group state that they expect their next desludging will be mechanical, while 89% in the highest price group state that they expect their next desludging will be mechanical. The differences are even larger when we compare the last desludging of each group: 69% of households in the lowest price group got a mechanical desludging for their last desludging while 94% of those in the highest price group purchased a mechanical desludging for their previous desludging. If we compare manual desludgings in the past, we see that the gap widens even further: 76% of households in the lowest price group have ever purchased a manual desludging, while only 40% of households in the highest price group have ever purchased a manual desludging.

Taken together, this evidence from non-targeted variables implies that the intervention did target poor households for assistance, and was effective in determining the right amount of aid to provide in order to induce switching.

6 Counterfactual Experiments

While the intervention in Section 5 demonstrates that the Targeted Pricing treatment had a variety of statistically significant impacts on behavioral and health outcomes, at least three questions remain: How would other, alternative market designs have fared? If different design variables had been used, what would the consequences have been? How do treatment effects respond to the subsidy level, s? This section uses the Control, Demand Elicitation, and Targeted Pricing group data to predict how behavioral and financial outcomes would have changed under alternative designs, as well as to demonstrate the channels through which the platform operates.

We have three types of counterfactual experiments in mind: varying the market design (Section 6.1), the platform's information structure (Section 6.2), and the level of the subsidy (Section 6.3), which require counterfactual models of pricing and behavior. In the intervention, household i was quoted a price t_i; household i then decided whether or not to deposit, $Deposit_i \in \{0,1\}$, and whether or not to arrange a desludging through the platform, $Arrange_i \in \{0,1\}$; which resulted in the percentage of mechanical desludgings it purchased given that the household did or did not deposit, $m_i^1 \in [0,1]$ or $m_i^0 \in [0,1]$, respectively. To measure the overall impact of the platform and its financial prospects under alternative designs, we want to predict what would have happened at new prices $\{t_i'\}_{i=1}^I$ and a potentially different subsidy level s' with cost of service c_i, and compute for each household i in the Targeted Pricing group,

- Expected mechanical share: $\widehat{Share}_i' = \widehat{Deposit}_i' \times \hat{m}_i^{1'} + (1 - \widehat{Deposit}_i') \times \hat{m}_i^{0'}$, where predicted values of $\hat{m}_i^{1'}$ or $\hat{m}_i^{0'}$ above 1 or below 0 are rounded[37] to 1 or 0, respectively.
- Profit: $\widehat{Profit}_i' = \widehat{Arrange}_i' \times (t_i' - c_i)$.
- Budget balance: $\widehat{BB}_i' = \widehat{Arrange}_i' \times (t_i' - c_i + s')$.
- Subsidization rate: $\widehat{SR}_i' = \widehat{Arrange}_i' \times \dfrac{(t_i' - c_i)}{c_i}$.

These metrics of performance each capture an important aspect of platform operation. Market share is our general measure of success at converting households from manual consumption to mechanical. Different designs might, however, achieve higher mechanical utilization by incurring greater expected losses,

[37]This linear probability model rarely predicts values below 0, but does predicts values above 1, and this logical impossibility tends to benefit the Targeted Pricing treatment when compared to other designs.

motivating us to also evaluate financial measures of performance like profits and expected budget balance. Finally, we wish to quantify which households are paying into the system and which are benefiting from it in expectation. The subsidization rate is a "reverse Lerner index" measuring the percentage of the procurement cost that is paid by the household: negative values correspond to profit losses to the platform on the sale, while positive values correspond to profits. The subsidization rate is particularly effective at identifying how different designs achieve their effects, since it provides a dimensionless measure of the assistance received by each household.

We now turn to estimating models of these key quantities, starting with a model of mechanical and manual desludging decisions — $Deposit_i$, m_i^1, and m_i^0 — analogous to the treatment effects regressions (24) and (26) in Section 5.2. Recall that Targeted Pricing households were asked whether or not they wished to forgo a 500 CFA gift as a deposit at the time of the baseline survey. For a household i in the Targeted Pricing group, we observe its decision whether to pay a deposit to the platform or not, $Deposit_i \in \{0, 1\}$, and the percentage of mechanical desludgings it purchases, $PctMechanical_i \in [0, 1]$, as a function of price t_i and covariates z_i. We model this as an endogeneous regression switching model (Amemiya, 1985),

$$
\begin{aligned}
Deposit_i &= \mathbb{I}\{z_i\alpha_d + \beta_d t_i + \varepsilon_i\}, && (28)\\
PctMechanical_i &= \begin{cases} x_i\delta_1 + \pi t_i + \varepsilon_{i1}, & Deposit_i = 1 \\ x_i\delta_0 + \varepsilon_{i0}, & Deposit_i = 0 \end{cases} && (29)
\end{aligned}
$$

where x_i is a subset of z_i and ε_i has a standard normal distribution, but the shocks $(\varepsilon_{i0}, \varepsilon_{i1})$ are correlated with ε_i but not necessarily normally distributed. Price is excluded from the $PctMechanical_i$ equation when $Deposit_i = 0$ because if the household fails to leave the deposit, it no longer has access to the platform, and the price it was quoted should no longer play a role in its service choice.

We control for variables related to wealth and past desludging behavior, including a wealth index based on a principal components analysis of the household's assets; whether the last desludging was mechanical or manual; whether the current residents ever desludged at that household; the share of desludgings in the previous five years that were mechanical; and the respondent's age. Also included is a predicted value of the household's willingness-to-switch, $\hat{\eta}$, created by estimating the Demand Elicitation group's reported willingness-to-switch values using the LASSO and then predicting values for the Targeted Pricing group; the value of the latent variable that determines the household's assignment to a pricing bin, called Weight; the enumerator's subjective assessment of whether the household's responses were believable or accurate, called Reliable Responses; and price.

We estimate α_d and β_d in (28) by maximum likelihood, with results reported in the first and second columns of Table 17, which are the coefficients and marginal effects at the mean, respectively. The marginal effect at the mean of price is $-.03$, statistically significant at the 10% level, so that households facing higher prices are indeed less likely to purchase. Respondent Age and $\hat{\eta}$ are statistically significant and positive, and a Likelihood Ratio test rejects the hypothesis that the model is jointly insignificant.

The deposit choice, however, potentially creates selection: the Targeted Pricing households who forwent the 500 CFA gift in order to secure a mechanical desludging at the price quoted might differ systematically from those households who failed to do so[38]. In order to provide unbiased estimates of the $PctMechanical_i$ equation (29), we adopt the semi-parametric approach described in (Powell, 1994) or (Newey, 2009) and evaluated empirically in (Newey et al., 1990). Using the estimated coefficients from the probit model, define

$$\hat{\xi}_i = -(z_i'\hat{\alpha}_d + \hat{\beta}_d t_i), \tag{30}$$

which can roughly be interpreted as the negative of the expected net utility of making a deposit, conditional on household characteristics z_i and price t_i. Plots of the empirical cumulative distribution function of $\hat{\xi}_i$ conditional on

[38]More formally, $\mathbb{E}[\varepsilon_{i1}|x_i, t_i, Deposit_i = 1] \neq 0$ and $\mathbb{E}[\varepsilon_{i0}|x_i, Deposit_i = 0] \neq 0$ because households who deposit are systematically more likely to purchase mechanical, so that OLS regression of $PctMechanical_i$ on x_i and t_i will result in biased estimates.

$Deposit_i = 0$ and $Deposit_i = 1$ are provided in Figure 8, and a Kolmogorov-Smirnoff test rejects the hypothesis of equality of the distributions from which the two samples were drawn at any conventional level of significance ($D = 0.263$, p-value $= 5.665 \times 10^{-10}$). We then regress $PctMechanical_i$ on x_i, t_i, and powers[39] of $\hat{\xi}_i$ for the group that deposited and the group that did not,

$$\begin{aligned} m_i^0 &= x_i\delta_0 + \sum_{\ell=1}^{L^0} \rho_\ell^0 (\hat{\xi}_i)^\ell, \quad Deposit_i = 0 && (31) \\ m_i^1 &= x_i\delta_1 + \sum_{\ell=1}^{L^1} \rho_\ell^1 (\hat{\xi}_i)^\ell + \pi t_i, \quad Deposit_i = 1. && (32) \end{aligned}$$

Since $\hat{\xi}_i$ is a linear combination of variables in (z_i, t_i), some first-stage variables must be excluded from the second-stage in order to achieve identification. For a variable to be valid for exclusion in the second stage, it should shift the household's propensity to leave a deposit conditional on the price quote, but

[39] The intuition for this procedure can be seen by comparing it with the Heckman two-step approach, which begins with a discrete choice

$$d_i = \begin{cases} 1, & \varepsilon_i \geq \xi_i \\ 0, & \varepsilon_i < \xi_i, \end{cases}$$

so that $Pr[d_i = 1] = 1 - \Phi(\xi_i)$, and the second-stage equation

$$y_{ij} = x_i\delta + \varepsilon_{ij},$$

but $\mathbb{E}[\varepsilon_{ij}|d_i] \neq 0$ due to selection, where $cov(\varepsilon_i, \varepsilon_{ij}) \neq 0$. The structural assumption of joint normality implies that this conditional expectation can be computed analytically, so that

$$\mathbb{E}[y_i|d_i] = x_i\delta + \underbrace{\rho_{\varepsilon_i,\varepsilon_{ij}}\sigma_{\varepsilon_{ij}}\lambda(\xi_i)}_{\text{Heckman Correction}},$$

where $\lambda(z)$ is the Mills ratio $\phi(z)/\Phi(z)$ or inverse Mills ratio $-\phi(z)/(1-\Phi(z))$, depending on whether the data are observed given $d_i = 1$ or $d_i = 0$. The semi-parametric approach "replaces" $\rho_{\varepsilon_i,\varepsilon_{ij}}\sigma_{\varepsilon_{ij}}\lambda(z)$ with a flexible polynomial,

$$\rho_{\varepsilon_i,\varepsilon_{ij}}\sigma_\varepsilon\lambda(\xi_i) \approx \sum_{\ell=1}^{L} \rho_\ell(\xi_i)^\ell \approx \mathbb{E}[\varepsilon_{ij}|d_i],$$

allowing a similar two-step approach to estimation that relaxes the structural assumption of normality and provides consistent estimates of δ.

not their propensity to get a mechanical desludging conditional on the deposit choice and price quote. We use Reliable Responses — a dummy variable that takes the value 1 if the enumerator judged the household's responses to be dishonest or unreliable, thus indicating distrust or skepticism of the project and a diminished propensity to deposit — as the excluded variable.

We estimate (31) and (32) by OLS, and results are presented in Columns 3 and 4 of Table 17, where standard errors are bootstrapped[40] to account for the estimated regressors. We used LASSO to select the number of powers[41] of ξ_i. This results in the selection of two powers for m^0, although the first three powers are all significant, and only one power for m^1, although selecting zero was a possibility. The index $\hat{\xi}_i$ can be interpreted as the *negative* of the expected net benefit of depositing, so that the negative coefficients on $\hat{\xi}_i$ imply that households with higher benefit are more likely to purchase mechanical. The excluded variable, Reliable Responses, was not significant at conventional levels (p-value = .190), but has a large marginal effect at the mean of $-.181$. Indicators of past mechanical purchases increase the likelihood of purchasing mechanical, even if the household did not leave the deposit. Price in the m^1 regression has a negative coefficient, but it is small in magnitude and not statistically significant. We interpret these results as suggesting that households were price sensitive at the deposit stage, but conditional on depositing, were very likely to purchase mechanical even if they likely wouldn't have purchased otherwise.

Is modeling the deposit step really necessary? Couldn't $PctMechanical_i$ simply be regressed directly on platform price t_i and household covariates, instead of a two-step procedure including the deposit step? Such an approach overestimates the effect of the platform price on mechanical share, because it imposes a stronger negative relationship between price and consumption than

[40]Some bootstrap samples resulted in positive price coefficients. These samples were discarded because they could not be consistent with the underlying data generating process.

[41]In practice, there is no definitive way to select the number of powers. (Newey et al., 1990) use generalized cross validation and other authors have used the Akaike or Bayesian Information Criteria. We used the LASSO to select which powers $\{\hat{\xi}_i^\ell\}_{\ell=1}^{25}$ to include to minimize cross-validated mean-squared error — allowing for the possibility of selecting none — which is similar to Newey et al's approach.

exists. In reality, households can respond to a high platform price by opting out of purchasing mechanical services from the platform but still purchasing in the search market. A simple probit regression of $PctMechanical_i$ on t_i and z_i is presented in Table 21 and the predicted counterfactual shares for different market designs in Table 22. The regression coefficients are similar to the deposit regression in Table 17, but the predicted market shares are not credible, with many values below the predicted control (compare with Table 25). Modeling the deposit step therefore explicitly allows households to be more price elastic with respect to the platform than their demand for mechanical services, delivering more accurate counterfactuals.

We now turn to the financial side of the platform's performance, and focus on predicting which households actually arrange a desludging through the platform, conditional on their observables and the price, $Arrange_i$. This only requires estimating the mapping from prices and covariates to arrangement decisions, so we use a simple probit model

$$Arrange_i = \mathbb{I}\{\alpha_a' z_i + \beta_a t_i + \varepsilon_i\} \tag{33}$$

where ε_i is distributed standard normal. In our setting, we can control directly for the latent index (Weight) that determines the group to which each household i is assigned: there are no unobserved household or product characteristics which might bias β_a, which is the usual problem in estimating demand models[42]. Results are reported in columns 5 and 6 of Table 17. Again, Respondent Age, $\hat{\eta}$, and a dummy indicating whether the last desludging was mechanical are all statistically significant and positive, and price is negative and statistically significant, with a marginal effect at the mean of -0.022 (statistically significant at the 10% level).

Measuring impact requires a counterfactual evaluation of what would have happened in the absence of the intervention. To provide an accurate prediction of what would have happened if the Targeted Pricing group was denied access to the platform, we estimate $PctMechanical_i$ by LASSO using the Control

[42] See, for example, (Berry et al., 1995) or Petrin and Train (2010).

group data, and then predict what mechanical consumption for the Targeted Pricing group would have been. To leverage the data as much as possible, we include the same variables as in z_i and then use the LASSO to select from among 115 other control variables to minimize prediction error.

Do these models fit the data? Table 18 provides realized and predicted values for $Deposit_i$, m_i^0, m_i^1, and $Arrange_i$, with bootstrapped 90% confidence intervals[43]. For population averages, the largest deviations on average and in the 10,000 CFA bin are only 0.6 of a percentage point and 2 percentage points, respectively. The empirical and estimated mechanical market shares — our main outcome of interest — are given in Table 19, where all of the deviations are under a percentage point. The predicted average treatment effect is 4.0 percentage points and the predicted average treatment effect on the 10,000 CFA group is 10.6 percentage points, in line with estimates from the randomized controlled trial[44]. Note that the results of Table 18 are not estimated to specifically reproduce or target the results of the intervention, and the control column is even estimated using a separate LASSO regression independent of Targeted Pricing data. Finally, Table 20 provides a financial statement for the platform. The platform is designed to balance the *expected* budget, but ex post losses or gains are possible. Here, the realized loss was 102 CFA on average, or about 0.18 USD per household[45], compared with a predicted loss of 116 CFA. Overall, the counterfactual model (28) – (32) credibly fits the quantitative and qualitative features of the household decisions, market outcomes, and platform finances.

Another concern with this procedure is that the model might be extrapolating outside the domain of what the data can credibly explain. In particular,

[43]Some bootstrap samples resulted in positive price coefficients in the m_i^1 regression. These samples were discarded because they could not be consistent with the underlying data generating process, and produced incoherent results (counterfactuals with systematically lower prices achieved lower mechanical market shares, for example). This leads to asymmetric confidence intervals around the estimated value, which are based on the variation in the data and not asymptotic formulas.

[44]See page 42 and 43 respectively.

[45]On a per-household basis, this is a small amount, but could add up at scale. We think that using machine learning tools like classification through random forest[46] at the design stage could improve further improve this aspect of the platform's performance.

because the prices are deterministically administered on the basis of observables, the type of household who was quoted a price of 10,000 CFA by the platform is never observed responding to a price of, say, 20,000 CFA. Consider Figure 9, which provides empirical cumulative distribution functions of the Demand Elicitation group's willingness-to-switch values. The distribution of willingness-to-switch values for the 20,000 CFA bin first-order stochastically dominates that of the 17,500 CFA bin, the 17,500 CFA bin dominates 15,000 CFA bin, and the 15,000 CFA bin dominates the 10,000 CFA bin, but each group has essentially the same support, so there is substantial overlap between the groups' willingness-to-switch values. In our models, we use the Demand Elicitation data to predict willingness-to-switch values for the Targeted Pricing group households, $\hat{\eta}$, incorporating finer-grained information about their propensity to deposit or arrange into the estimation. This jointly expands the range of data being used to predict counterfactual behavior beyond the discrete deposit and arrange choices alone.

In our results, we present the counterfactual mechanical share average for the whole Targeted Pricing group, as well as by each pricing bin. The comparison across pricing bins implicitly assumes the bins correctly identify relative wealth of households. We argue, however, that the results in Table 16 from Section 5 show that the Targeted Pricing bins accurately target households based on non-targeted measures of wealth, and therefore that maintaining comparisons across these bins allows for a reasonable and parsimonious comparison of the treatment effects across different market designs.

We make one important deviation from the design of the original platform in computing counterfactual prices. For the original pricing rule, we used an ordered logit model to map observables to prices after trying a variety of different approaches, including adding and dropping different pricing bins to ensure a good overall fit. In general, we would like to adopt a methodology that automatically maps design variables to prices, is robust, and easily replicated. At this point, we use the random forest algorithm for classification[47] to map

[47]See Appendix G for an explanation of the random forest algorithm, or (Hastie et al., 2017). In short, the random forest algorithm uses a large number of decision trees to fit the

design variables to price where necessary.

6.1 Alternative Market Designs

While the randomized controlled trial provides causal estimates of the treatment effect of access to the platform for households in the Targeted Pricing group, it does not address the question of how alternative designs would have performed. This section addresses this question and illustrates the channels through which Targeted Pricing operates. We consider three alternative designs:

i. *Auctions*: the clearing prices in the procurement auctions are passed directly to households who call in from those neighborhoods, who then either accept or reject.

ii. *Proxy-means testing*: We predict the income per household member in the demand elicitation group using a LASSO regression with 113 potential control variables[48], and use the results to predict the income per household member of the targeted price treatment households. We then create an indicator variable for whether the household per capita income falls above or below Burkina Faso's urban central region per capita poverty line, and 150 percent of the poverty line. We quote households classified as poor with the auction clearing price in their neighborhood minus the subsidy.

iii. *Price Ceiling*: We offer the average price available in the decentralized market, 16,833 CFA, to all households, as if a municipal authority were able to fix the price and sufficiently penalize firms that deviated from it.

The proxy-means testing counterfactual is constructed assuming the same subsidization level as our experiment, \$3.00. We also provide results for *Subsidized*

observed data, then averages over the trees to deliver a more robust decision rule for how to assign households to prices.

[48]Variables directly related to income were removed, which is why there are fewer variables included here than in the main regressions.

Procurement Auctions and a *Subsidized Price Ceiling*, where the clearing price or average price are reduced by the same subsidy level, to provide a more fair comparison to Targeted Pricing[49].

Table 23 provides counterfactual mechanical shares, and Figure 10 provides a visualization of the main results. On average, Targeted Pricing yields an 80.8 mechanical share, while the Auctions yield a 78.8 share, 100% Proxy-Means yields 79.3, and the Price Ceiling yields 78. The main difference is in how these designs achieve these outcomes. The treatment effect on the 10,000 CFA group is 10.6 percentage points for Targeted Pricing, 4.2 for Proxy-Means, 2.2 for the Auction, and 1.5 for the Price Ceiling. So while the average treatment effects are close, Targeted Pricing achieves its average effect by increasing the mechanical share among the poorest households who are the least likely to get a mechanical desludging on their own, while the other designs achieve their effects by causing switching in the higher-priced groups.

Why does Targeted Pricing outperform the other counterfactual designs, on average and in the 10,000 CFA price bin? Figure 11 illustrates the expected subsidization rates for the households in the Targeted Pricing group and Table 24 provides the averages. The Targeted Pricing treatment achieves average subsidization rates of around -11.8% for the 10,000 CFA group, -1.3% for the 15,000 CFA group, 1.2% for the 17,500 CFA group, and 3.1% for the 20,000 CFA group. Note that we are able to achieve positive subsidization rates but still get positive take-up because we are able to buy through competitive processes, and then undercut some of the high prices the relatively wealthy households expect to receive in the market. On average, the 17,500 CFA and 20,000 CFA groups were paying into the system by covering a positive percentage of the cost of procurement, while the poorest households received almost a 12% discount on the platform's cost of procurement. In contrast, no other treatment achieves a subsidization rate larger than -2.0% on the 10,000 CFA group, and the only other design that achieves positive subsidization rates

[49]Later results will provide a fully "apples-to-apples" comparison by computing the subsidy level for the alternative designs that would have achieved the same treatment effects as the Targeted Pricing intervention.

is the Price Ceiling, where a small number of households are purchasing at a relatively high price. This illustrates how Targeted Pricing combines data-driven targeting and cross-subsidization to funnel subsidies and profits to the poorest households[50].

One potential objection could be that Targeted Pricing achieves these results by running much larger budget deficits than the other designs, but this is not the case. Budget balance calculations are reported in Table 25. The Auctions and Proxy-Means have average loses of zero[51], the Price Ceiling raises 78 CFA per household, and Targeted Pricing loses 100 CFA per household ($.18) on average. While $.18 might not be an insignificant amount at scale, it is also not so large that the differences in performance between Targeted Pricing and the other designs can be explained solely by violating the budget constraint.

An alternative way of comparing designs is to answer the question, "How large a subsidy is required for Proxy-Means, Auctions, and the Price Ceiling to achieve the same performance as Targeted Pricing, on average and for the 10,000 CFA price group?" We call this quantity the *design variation*, π^*, since it solves the implicit equation $ATE_{TP}(s) = ATE_j(s + \pi^*)$, adjusting the subsidy by π^* to equate average treatment effects between Targeted Pricing and an alternative design j, similar to the equivalent variation in consumer theory. Table 26 provides the design variations for Auctions, Proxy-Means Testing, and Price Ceiling, on average and for the 10,000 CFA group. To

[50]For the reader that disagrees with the use of the pricing bins as a way of "cutting the data" for this analysis, there are two things to note. First, the intervention is designed to maximize overall take-up and not just the impact on the 10,000 CFA bin, and within that bin, the baseline utilization of mechanical desludging is about 40%. This suggests that even if other partitions based on measures of poverty or utilization were used, the results would be similar. Second, no other treatment achieves subsidization rates on the order of 10% for *any households* — see Figure 11 — so that even if other subgroups were of particular interest, the other designs considered would not achieve similar impacts on them.

[51]These calculations for the alternative designs are based on *ex post* costs, so the subsidy plus the price paid minus the cost equal zero. For the Targeted Pricing treatment, the costs were unknown at the time of contracting, so the budget does not balance exactly. Using the realized costs here advantages the counterfactual designs. We do this in order to provide the most realistic and fair comparisons, rather than speculate on what a designer might have predicted would happen at the time of design. In practice, predicting the exact amount that would be spent on Proxy Means or Subsidized Auctions would be as difficult as predicting the cost of the Targeted Pricing program, and would also run a deficit or surplus in practice.

achieve the same average treatment effect, an Auction requires an additional subsidy of 1,704 CFA, Proxy-Means Testing requires 1,643 CFA at the 100% level and 350 CFA at the 150% level, and the Price Ceiling requires 2,478 CFA. In order to match the average treatment effect on the 10,000 CFA group, the numbers are much larger: 6,132 CFA for Auctions, 5,388 CFA for Proxy-Means at the 100% level and 4,518 at the 150% level, and 6,825 for the Price Ceiling, all values in excess of $10.

Why are the design variations so large? In short, because the alternative designs fail to target likely switchers, and decreasing returns quickly blunt the effect of additional subsidies on those households the methods do target. Consider Figures 12 and 13 (and corresponding Tables 29 and 30), which illustrate subsidization rates for the different designs. In order to provide enough aid to poor households to match the Targeted Pricing treatment effect on the 10,000 CFA group, the alternative designs have to provide a significant additional amount to create the same number of switchers to match subsidization rates of -11%. But these alternative designs employ weaker methods of screening on willingness-to-switch, so larger subsidies don't just go to the poorest households, but to all households who participate. The consequence is that a significant amount of the additional aid is going to relatively wealthy households, and additional subsidies do little to increase their propensity to purchase mechanical. The decision to exclude households based on a strict test in Proxy-Means Testing becomes a liability, since many potential switchers receive no aid at all while others who were already receiving aid are less and less likely to switch on the margin. This illustrates how conventional designs can fail to target switchers, resulting in attenuated treatment effects and more costly designs.

Why does Targeted Pricing consistently outperform Proxy-Means Testing? Aren't the two approaches similar, at least in principle? Proxy-Means Testing takes a training data set and attempts to predict poverty using wealth, then subsidize those who are diagnosed as poor. This process focuses on trying to predict a latent household variable, poverty, rather than whether the household will actually purchase the product or how much it would need to be subsidized to do so. In principle, such a model could correctly predict who is poor and still fail to increase take-up if the subsidized price is not low enough to create switchers. Similarly, if household decisions are relatively noisy conditional on wealth, such a rule will perform poorly in predicting who needs assistance, and limited subsidies will end up in the hands of people who would have purchased anyway. Our approach instead focuses on understanding what households do by eliciting their willingness-to-switch and past decisions, and then mapping observables into predicted actions conditional on assistance. It turns out that focusing on this correspondence — rather than the correspondence of observables to poverty — is more predictive of behavior and a better approach to the design of social policy.

6.2 Information Structures

A key choice in the design of the platform is selection of variables included in the pricing equations, (9). These variables must be observable or verifiable, so that households cannot manipulate them to their advantage. This naturally leads to two questions: If less or different information had been used, how would the platform have performed? What variables are the most useful in practice? This section addresses these issues by exploring counterfactual outcomes using alternative sets of design variables that correspond to information held by state or non-state actors, running the platform optimization exercise in Section 4 from scratch, and then predicting counterfactual outcomes.

The version of the model which we took to the field used a variety of variables based on enumerators' subjective assessments, municipal records, and platform records[52]. Now, we wish to consider more conservative versions

[52] Recall from page 23: information gathered by the enumerator during a household interview including housing type (precarious, concrete, or rooming house), whether other households lived in the compound, the pit's distance to the road in meters, the number of people living in the household, the number of women living in the household, and whether the respondent finished high school; information available to a municipal authority, including last month's water and electricity bills, whether the household owns the dwelling; and information available to a continuously operating platform that can keep its own records, including average months between desludgings, whether the last service episode required more than one trip because of the large size of the pit. Instruments included electricity ex-

of this model, giving special attention to the case of an NGO that does not have access to municipal records and a municipal authority that lacks a budget for household surveying.

Figure 14 succinctly illustrates the relationships between the original and alternative information structures. The first alternative information structure slightly reduces the information used in the version we took to the field, providing a more conservative version of the information available to a continuously operating platform working with local municipal authorities. Alternative Version 1 removes Own House from the set of controls, and Number of Women in the Household and Respondent Finished High School from the set of instruments. Alternative Version 2 is limited to the information that would be available to a municipal authority with a limited budget: the controls include Water Bill More than 5,000 CFA and Own House, and the instruments include only the Electricity Bill. In many settings, property rights may be weak or difficult to determine, so we provide Alternative Version 2B that also removes Own House. Alternative Version 3 approximates the information that an NGO or non-state actor could easily collect through visits and interviews. It includes dummy variables for housing quality and type, Other Households in Compound, and Pit Meters From Road as controls, and Household Size as an instrument. In some cases, information might be restricted only to what an enumerator can observe during a short visit. To address this, Version 3B removes Household Size as an instrument, relying entirely on the functional form of the tri-variate normal to estimate the platform demand model.

We re-estimate prices based on these alternative sets of design variables, and then use the model from Section 6 to simulate which households would have purchased a mechanical desludging under the alternative estimated pricing model. One key task in designing the system is to match observables to prices. Figure 15 illustrates the optimal Linear Programming, Ordered Logit, and Random Forest rules side-by-side for the Demand Elicitation group. Our modification of the Ordered Logit rule outperforms the Random Forest rule

penditure, the number of people in the household, the number of women in the household, and whether or not the respondent completed high school.

when using the full information set (especially in the 20,000 CFA price bin), but then is typically out-performed for the other information structures in terms of fit by the random forest. While designers might benefit by tailoring more idiosyncratic rules based on the context and the most up-to-date methods, here we focus on the random forest classification rule since it is straightforward to implement and performs reasonably well in matching observables to prices.

Counterfactual shares are presented in Table 31 and visualized in Figure 10 (b). Note that, compared with the ordered logit version (80.8 on average and 68.6 in the 10,000 CFA bin), the predicted treatment effects are slightly lower using the random forest rule (80.6 and 66.3). On average, the Original and Conservative versions achieve the same market share of 80.6, and the Conservative version exhibits a decline of only .04 percentage points in the 10,000 CFA bin.

Surprisingly, the municipal information sets in Versions 2 and 2B exhibit higher average market shares at 81.5 and 82.1, respectively, compared to the original version, which achieves an 80.8 mechanical share, but lower values in the 10,000 CFA bin, at 66 and 65.6, respectively, compared to 66.3. The explanation for this can be found in Table 32, which provides the budget balance values for each information structure: Version 2 and 2B lose 177 and 306 CFA per household on average compared with 24 CFA in Version 0, while the loss in the 10,000 CFA bin is 894 CFA for Version 0, but only 808 and 775 CFA for Versions 2 and 2B, respectively. This illustrates that as information is restricted, effective targeting becomes more difficult, leading to interventions that can be more generous than intended and target the "wrong" households. While these numbers might be "small" in the context of this experiment, the differences might matter at scale. In principle, machine learning methods like penalized regression can be adapted to the Tobit model given the available data, and bootstrapping methods can be used to simulate the distribution of expected losses.

Version 3 exhibits a similar pattern to Versions 2 and 2B, with a comparable average market share of 80.7 versus 80.6, but a diminished 10,000 CFA bin share of 64.6 versus 66.3. For Version 3B — the one information structure without any instruments in the Tobit V estimation, so that the model is identified off the structural assumption of normality — the results are markedly worse. The average and 10,000 CFA bin market shares are indistinguishable from Predicted Control at 77.6 and 63.3, respectively. This suggests that the role of the instruments in identifying the Tobit V model is important not just for producing estimates that are robust to violations of the assumption of normality, but also for producing pricing rules that consistently mapping observables to prices[53].

Can the value of the variables in designing the pricing rules be quantified? One by-product of the construction of a random forest is a measure of variable importance[54], given in Table 33. In the original version, the most informative variables are Electricity bill, Desludging frequency, and Housing type dummies, which are all either observable on a visit from an enumerator or included in municipal records. As variables are removed from consideration, the remaining variables capture more unexplained variation and become more relevant, similar to linear regression. In Versions 2 and 2B, Electricity bill is still by far the most informative variable, but Own/rent and Water Bill > 5k both become informative. In Versions 3 and 3B, the most important variables are the Housing Dummies, which are again strong predictors of wealth and relatively cheap to survey. Overall, this analysis provides one explanation

[53]Indeed, the only change from Version 3 to 3B is the removal of the instrument, which as Table 33 illustrates, is not even particularly informative in the Version 3 pricing rule; see the next paragraph. In Figure 15, Version 3B is also qualitatively different than the other rules, essentially selecting only 10000 and 20000 CFA as prices.

[54]Appendix G gives a more precise definition of random forests and variable importance, but the next few sentences provide a rough explanation for those uninterested in the details: A decision tree is a collection of binary decision rules — e.g., "is the electricity bill greater or less than 16,000 CFA?", "is the distance from the pit to the road greater or less than 8 meters?" — that assigns a price prediction, and a random forest is a collection of decision trees constructed on bootstrapped samples using different, random selections of explanatory variables that assigns the modal price prediction made by all of its trees. When adding a new binary rule to the tree, the algorithm scans the tree and finds the addition that minimizes the heterogeneity in final predictions as measured by the Gini coefficient. The variable importance measure of a variable takes the average reduction in the Gini coefficient across all of the decision trees in the forest in which that variable appears. With a large number of trees, this is a measure of the reduction in misclassification attributable to the variable.

for why the mapping from design variables to prices (and resulting treatment effects) is fairly robust across information structures: the most informative information happened to be the easiest to gather, and some useful information goes a long way in identifying needy households.

6.3 Sustainability

If a design requires substantial and perpetual financial assistance to operate, it may not be sustainable in the long run. This section varies the subsidy and provides counterfactual estimates of mechanical market share, studying the trade-off between the level of subsidization required and the treatment effect that can be achieved.

Note that because maximizing mechanical market share (14) subject to the expected budget constraint (15) has, as its dual problem, profit maximization subject to a constraint that a certain mechanical market share is achieved, all of the analysis can be conducted by varying the subsidization level, s. When s is positive, it corresponds to a scenario in which an outside agent must pay into the platform to fund it, while when s is negative, it corresponds to a scenario in which the platform is making positive profits from operating. We vary this parameter from -750 CFA (-\$1.36) to $13,000$ CFA (\$23.64), solving for the prices that would have been quoted, and then using the counterfactual model to predict what the treatment effects would have been[55]. This allows us to construct estimates of the trade-off between subsidization and treatment effect.

In addition to using a random forest algorithm to map observables to prices, this exercise departs from the original design methodology in two important ways. First, we use an average procurement cost of $13,750$ CFA, which is the average of the lowest costs we achieved across neighborhoods at the time the

[55] The pricing model predicts that at any s less than -750 CFA, the linear programming problem (14) — (16) has no feasible points. The prices associated with -750 CFA therefore correspond to the profit-maximizing prices, given the platform's beliefs based on the experiment with the Demand Elicitation group. Note that this value varies with the information structure, so might best be interpreted as the profit-maximizing price *this platform believes it can extract.*

project concluded, rather than $17,500$ CFA, which is the average clearing price at the time the Targeted Pricing intervention began, and, second, we expand the set of available prices to $\{0, 5000, 10000, 12500, 15000, 17500, 20000, 22500\}$, allowing for greater freedom in pricing behavior by the platform. We adopt these changes because we are interested in what an optimal platform could achieve in the future given what we now know, not what the optimal platform could have achieved given what we knew when the intervention began.

Panel (a) of Figure 16 shows how the average Mechanical share and the share for the 10,000 CFA price bin households vary with the subsidy. Even at a negative subsidy value of -750 CFA, there is a 3.3 percentage point increase on average, and 5.2 percentage points in the 10,000 CFA price group. At a subsidy value of 0 CFA, there is a 4.0 percentage point increase in mechanical share and 8.0 points in the 10,000 CFA price group. As the subsidy increases to approximately 4500 CFA, the impact on the $10,000$ CFA households flattens out, increasing again at a subsidy of $10,000$ CFA. The reason for this is illustrated in Panel (b) of Figure 16: the budget constraint prevents the platform from offering the 0 price offer until it has a substantial amount of funds available[56]. Once it does, the platform begins offering free desludgings, and the treatment effect on the poorest group begins increasing again; indeed, kinks in panel (a) correspond to subsidization levels in panel (b) at which the set of prices offered changes.

This analysis illustrates two points. First, access to a centralized market can improve welfare even without subsidization. Indeed, a large portion of the treatment effect can be attributed to putting service providers into price competition on the platform, driving down procurement costs to the benefit of households who would otherwise have faced higher prices. Which households benefit then depends on the ability of the intermediary to target and cross-subsidize effectively. Many non-state actors or financially constrained governments cannot perpetually subsidize goods or services, but can build platforms that minimize procurement costs, engage in cross-subsidization, and

[56] As this model is based on a stochastic discrete choice model, the market share never reaches 1, but gets arbitrarily close as the subsidy level increases past $12,500$ CFA.

potentially even return a modest profit. Our data and experiments cannot address whether there are general equilibrium effects or how large they are, but a robust platform with a large market share might also become competitive with the decentralized market, leading to further welfare gains. Second, this analysis shows that reaching the poorest households may require very large subsidies. Between 3,000 CFA and 10,000 CFA there is a flat segment because the subsidy is not yet generous enough for the platform to offer free mechanical desludgings to the poorest households. While some $10,000$ CFA price bin households exhibit immediate gains from access, many are left behind until the subsidy reaches approximately 60% of the average price of the mechanical service at 10,000 CFA, and the mechanical share begins increasing again.

7 Conclusion

Increasing the take-up of health and sanitation goods may require large subsidies, and targeting subsidies to only those marginally failing to adopt is difficult. Households who would purchase in the absence of the subsidy are those most interested in receiving the subsidy, but this raises the cost of converting the poorest households to improved products and services. We solve the mechanism design problem required to design and implement a platform in which the households most at risk for purchasing the negative-externality producing service receive the highest subsidies, and those who would purchase the improved service anyway have access to the platform, but their purchases are used to augment the subsidy budget for others. In contrast to much of the current literature on increasing take-up of health and sanitation goods, this is accomplished in the presence of a prevailing decentralized market; consumers can opt out of purchasing through our platform–a common situation in developing countries (for example, water purification tablets and basic mosquito nets are often available in local markets). Subsidies can be more effectively employed to raise take-up of key products and services with externalities if a data-driven approach is adopted so that the budget is more effectively dedicated to those who would not have purchased the good in the absence of the subsidy.

We see model and implementation criticism as an important feature of a project such as this. While the demand model was deliberately selected to be simple and a workhorse economic model, we might have exploited other tools to deliberately focus on prediction rather than point estimation, drawing on the machine learning literature. A particularly difficult parameter to estimate is the correlation between mechanical and manual price shocks, which is typically unidentified in the Type V Tobit model. Our solution was to estimate this parameter off the subset of households that recalled both mechanical and manual prices for the last job, but there is obviously selection into this group, since "shoppers" will likely get lower prices. In addition, the linear program did not account for correlation in shocks between households in clusters. We see this as explaining a large proportion of deviations in stage two outcomes from the stage one estimates. Finally, we test the platform at relatively small scale–further research remains to be done on the general equilibrium aspects of full scale up of subsidy programs such as this.This is a short list of shortcomings, but we hope to further investigate and refine the methodology by exploiting the availability of the datasets from the two stages, as was done in the counter-factual subsidy exercise.

The existing research on sanitation shows that it is extremely difficult to have significant impacts on health through improving sanitation choices made by households. The evidence on encouragement campaigns and shaming is mixed, but CLTS campaigns have been found to have little impact except for the most intensive programs (see, for example, Gertler et al. (2015)). For comparison, in a randomized controlled trial of water and sanitation and nutrition programs of over 8000 households, Null et al. (2018) find no impact of the interventions on diarrhea in children, and small impacts in year two on height only for the treatment arm which included sanitation, hand washing, and nutrition. McIntosh and Zeitlin (2018) find that a USAID program aimed at nutrition and water and sanitation and costing $142 per household had little impact on health indicators and served only to somewhat increase savings levels in the households. Many health and sanitation campaigns have been focused on rural areas where open defecation is common, yet because of the high population density in urban areas, the externalities from improper disposal of fecal waste may be much larger (Kresch et al., 2019).

While this paper focuses on a platform operated by a local government, the general methodology employed could be useful for a variety of other actors. In particular, NGOs often face questions of impact and sustainability. The approach used here answers both questions, by first gathering exactly the kind of data required to predict how much impact a market intervention may have, and then testing the optimal design. By further refining this kind of methodology, pilot studies and small grants might be made more effective in channeling limited public and international aid dollars into well-designed programs with impact.

References

Abdulkadiroglu, A. and Sonmez, T. (2003). School choice: A mechanism design approach. *American Economic Review*, 93:729 – 747.

Ahrens, A., Hansen, C., and Schaffer, M. (2018). Pdslasso: Stata module for post-selection and post-regularization ols or iv estimation and inference. *Statistical Software Components S458459*.

Alatas, V., Banerjee, A., Hanna, R., Olken, B., and Tobias, J. (2012). Targeting the poor: Evidence from a field experiment in Indonesia. *American Economic Review*, 104(2):1206–1204.

Alatas, V., Banerjee, A., Hanna, R., Olken, B., Wai-Poi, M., and Purnamasari, R. (2016). Self targeting: Evidence from a field experiment in Indonesia. *Journal of Political Economy*, 124(2):371–427.

Amemiya, T. (1985). *Advanced Econometrics*. Harvard University Press.

Bagnoli, M. and Bergstrom, T. (2005). Log-concave probability and its applications. *Economic Theory*, 26:445–469.

Banerjee, A., Hanna, R., Olken, B., and Sumarto, S. (2018). The (lack of) distortionary effects of proxy-means tests: Results from a nationwide experiment in indonesia.

Basurto, P., Dupas, P., and Robinson, J. (2017). Decentralization and efficiency of subsidy targeting: Evidence from chiefs in rural Malawi.

Belloni, A., Chernozhukov, V., and Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies*, 81(2):608–650.

Berry, J., Fischer, G., and Guiteras, R. (2015). Eliciting and utilizing willingness to pay: Evidence from field trials in northern Ghana. *Working Paper*.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 63(4):841 – 890.

Bulow, J. and Klemperer, P. (1996). Auctions versus negotiations. *American Economic Review*, 86(1):180–194.

Coady, D., Grosh, M., and Hoddinott, J. (2004). Targeting outcomes redux. *World Bank Research Observer*, 19(1):61–85.

Coffey, D., Haque, S., Hathi, P., Pant, L., and Spears, D. (2014). Place and child health: The interaction of population density and sanitation behavior in developing countries. *Demography*, pages 1–24.

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

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

Dupas, P., Hoffman, V., Kremer, M., and Zwane, A. P. (2016). Targeting health subsidies through a nonprice mechanism: A randomized controlled trial in kenya. *Science*, 353(6302):889–895.

Fewtrell, L., Kaufmann, R., Kay, D., Enanoria, W., Haller, L., and Colford, J. M. (2005). Water, sanitation, and hygiene interventions to reduce diarrhoea in less developed countries: a systematic review and meta-analysis. *Lancet infectious diseases*, 5(1):42–52.

Gertler, P., Shah, M., Alzua, M. L., Cameron, L., Martinez, S., and Patil, S. (2015). How does health promotion work? evidence from the dirty business of eliminating open defecation. *NBER Working paper*, (20997).

Guiteras, R., Levinsohn, J., and Mobarak, A. M. (2015). Encouraging sanitation investment in the developing world: A cluster-randomized controlled trial. *Science*, 348(6237):903–906.

Hastie, T., Tibshirani, R., and Friedman, J. (2017). *The Elements of Statistical Learning*. Springer Series in Statistics. Springer Verlag.

Hussam, R., Rigol, N., and Roth, B. (2017). Targeting high ability entrepreneurs using community information: Mechanism design in the field.

Jack, K. (2013). Private information and the allocation of land use subsidies in Malawi. *American Economic Journal: Applied*, 5(3):113–135.

Jayachandran, S., de Laat, J., Lambon, E., Stanton, C., Audy, R., and Thomas, N. (2017). Cash for carbon: A randomized trial of payments for ecosystem services to reduce deforestation. *Science*, 357:267–273.

Jensen, R. and Miller, N. (2017). Information, demand and the growth of firms: Evidence from a natural experiment in India. *Working Paper*.

Kar, K. and Pasteur, K. (2005). Subsidy or self respect? Community led total sanitation: an update on recent developments. *IDS Working Paper*, (257).

Kidd, S. and Wylde, E. (2011). Targeting the poorest: An assessment of the proxy means test methodology. *Australian Aid*.

Kremer, M. and Miguel, E. (2007). The illusion of sustainability. *Quarterly Journal of Economics*, 122(3):1007–1065.

Kresch, E., Lipscomb, M., and Schechter, L. (2019). Social network effects on waste management and sanitation in urban and rural neighborhoods.

Maddala, G. (1983). *Limited Dependent and Qualitative Variables in Econometrics*. Econometric Society Monographs. Cambridge University Press.

Mara, D., Lane, J., Scott, B., and Trouba, D. (2010). Sanitation and health. *PLoS Medicine*, 7(11).

McIntosh, C. and Zeitlin, A. (2018). Benchmarking a child nutrition program against cash: Experimental evidence from rwanda.

McKenzie, D. (2012). Beyond baseline and follow up: the case for more T in experiments. *Journal of Development Economics*, 99(2):210–221.

McKenzie, D. and Sansone, D. (2017). Man vs. machine in predicting successful entrepreneurs: evidence from a business plan competition in nigeria. *World Bank Policy Research Working Paper*, (8271).

Milgrom, P. and Segal, I. (2019). Clock auctions and radio spectrum reallocation. *Journal of Political Economy*, Forthcoming.

Milgrom, P. and Segal, U. (2002). Envelope theorems for arbitrary choice sets. *Econometrica*, 70:583–601.

Mussa, M. and Rosen, S. (1978). Monopoly and product quality. *Journal of Economic theory*, 18(2):301–317.

Myerson, R. (1981). Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73.

Narayan, A. and Yoshida, N. (2005). Proxy means test for targeting welfare benefits in Sri Lanka. *World Bank PREM Working Paper Series*, (SASPR-7).

Nevo, A. (2001). Measuring market power in the ready to eat cereal industry. *Econometrica*, 69(2):307–342.

Newey, W. (2009). Two-step estimation of sample selection models. *The Econometrics Journal*, 12(1):217–229.

Newey, W., Powell, J., and Walker, J. (1990). Semiparametric estimation of selection models: Some empirical results. *American Economic Review*, 80(2):324.

Null, C., Stewart, C. P., Pickering, A. J., Dentz, H. N., Arnold, B. F., Arnold, C. D., Benjamin-Chung, J., Clasen, T., Dewey, K. G., Fernald, L. C., et al. (2018). Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural kenya: a cluster-randomised controlled trial. *The Lancet Global Health*, 6(3):e316–e329.

Olken, B. (2016). Hassles versus prices. *Science*, 6302:864–865.

Petrin, A. and Train, K. (2010). A control function approach to endogeneity in consumer choice models. *Journal of Marketing Research*, 47(1):3 – 13.

Powell, J. (1994). Estimation of semiparametric models. *Handbook of Econometrics*, 4:2443 – 2521.

Rochet, J. C. (1987). A necessary and sufficient condition for rationalizability in a quasi-linear context. *Journal of Mathematical Economics*, 16:191–200.

Roth, A. and Peranson, E. (1999). The redesign of the matching market for american physicians: Some engineering aspects of economic design. *American Economic Review*, 89:748–780.

Roth, A., Sonmez, T., and Unver, U. (2004). Kidney exchange. *Quarterly Journal of Economics*, 119:457–488.

Roth, A. E. (2002). The economist as engineer: Game theory, experimentation, and computation as tools for design economics. *Econometrica*, 70(4):1341–1378.

Spears, D. (2013). How much international variation in child height can sanitaiton explain? *Working Paper*.

Szabo, A. (2015). The value of free water: Analyzing south africa's free basic water policy. *Econometrica*, 83:1913–1961.

Tibshirani, R. (1996). Regression shrinkage via the lasso. *Journal of the Royal Statistical Society*, 58(1):267–288.

WHO and UNICEF (2015). 25 years progress on sanitation and drinking water: 2015 update and MDG assessment.

Wolak, F. A. (2016). Designing nonlinear price schedules for urban water to balance revenue and conservation goals. *Working Paper*.

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

A Demand Elicitation Script

At the end of the market survey, the enumerator reads the following script to the participant in their native language (Moore or Diola depending on the preference of the participant), and records the value that they state:

We had a study of desludging businesses in Ouagadougou, and we purchased some of their services.

We are selling the services of the desludgers that we purchased in your neighborhood and in a few other neighborhoods in Ouagadougou.

We are asking households for their price for the services and we will sell the services to the households that suggest the highest prices.

We would like to sell you a desludging service, but the price is not yet set.

The offer that you make for the desludging service will determine if you win and if you win the price that you pay will always be lower than what you have offered.

Here is the way we will determine who get the desludging services and how much they will pay:

I will ask you how much you are willing to pay for the desludging service.

We will leave a sticker here with the number that you can call to arrange the desludging.

When you call, the operator will compare your price to those of 8 other households who also need desludgings. There will be [randomized K number of winners] desludgings available.

The [randomized K number of winners] households that offer the highest prices will win, and each of the winners will pay the amount offered by the household that offered the highest amount but still lost.

The winners will pay for the desludging at the time that they get a desludging.

For example, suppose [8 minus randomized K] each offer 25,000 CFA and [randomized K minus 1] households offer 15,000 CFA.

If you were to offer more than 15,000 CFA, you would win and pay 15,000 CFA.

If you offered less than 15,000 CFA, then you would lose and you would not have access to the desludging.

Not read aloud: (If the respondent asks about ties, then the enumerator should explain that ties are resolved by randomization).

If you win, the price that you pay will always be less than the price that you offer.

You should never make an offer larger than what you would really want to pay, otherwise you could lose money.

You should never make an offer lower than what you would want to pay, because you would risk losing the opportunity to have a good price.

Is this clear to you, or would you like me to explain part of it again?

What offer would you like to make?

To be sure, if you win and the next household offers [households price minus 5%], would you want to purchase the desludging at that price?

If you lose, and you were to find out later that the price was [households price plus 5%], would you regret not having offered more?

If yes, what new offer would you like to make?

B Differences in number of desludgings used since baseline by households

A large percentage of desludgings (88% at baseline) are mechanical, which means that neighborhoods with more desludgings will mechanically have more mechanical desludgings if the total number of desludgings are not included as a control variable. At endline, we find that the number of desludgings procured in households in the control neighborhoods was statistically significantly higher than the number of desludgings procured in households in the treatment neighborhoods (households in the control neighborhoods which purchased any desludgings purchased on average 2.10 desludgings while households in the treatment neighborhoods purchased on average 1.95 desludgings), the 0.15 desludging difference is significant at the 5% level. Disparities in number of desludgings across neighborhoods are directly controlled for in market share estimates which divide by the total number of desludgings at the neighborhood level, therefore this is our preferred specification.

At endline we asked households a number of questions about the state of their latrine pit and whether they delayed desludgings over the treatment period. Households in the treatment group reported no difference in the number of days it took to get a desludging relative to the control group. Mean number of days to desludging is 8, and treatment households take 0.39 days less to get a desludging (insignificant, with a p-value of 0.97). From among those who did delay their desludging by more than 7 days, we find that households in the treatment group were 8.8% less likely to delay their desludging due to lack of funds (significant at the 5% level), but 1.1% more likely to delay their desludging due to accessibility issues (significant at the 10% level) and 3.9% more likely to delay due to difficulties in coordinating with the desludger (significant at the 10% level).

When asked what pushes households to get more desludgings, households and desludgers typically respond that the frequency with which households need desludgings depends on factors about the latrine pit such as: the size and type of latrine pit that the household has; factors about the households such as: the frequency with which they use water and the number of people using the latrine pit, factors about the geography of the region (which can vary substantially across a city including: elevation, the height of the water table, and soil type. Many of these factors are not known to the household and are not readily available (few households are able to tell us the size of their latrine pit–only 467 of the 2944 households surveyed gave an answer to the question, and many of the sizes reported are far outside standard sizes so are likely to be incorrect). We find that 31% of the variation in the number of desludgings that the household gets during the treatment period can be explained by a combination of the household baseline variables and geographic variables about the area. When we use these variables as controls as an alternative to controlling directly for the number of desludgings that the household purchased during the treatment period, the point estimate on the treatment effect increases but the standard errors also increase (which is to be expected with a less precise control).

C Tables

Table 1: Baseline Prices of Mechanical and Manual Services

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Mechanical:	10000	15000	15000	16935	17500	55000
Manual:	0	6000	15000	12134	15000	30000

Table 2: Demand model: Tobit

	(1) Selection (δ)	(2) Mechanical Price (β_{mech})	(3) Manual Price (β_{man})
Constant	1.787***	16.601***	3.542
	(0.551)	(0.917)	(6.834)
Average months between desludgings	-0.006***	0.005	0.017*
	(0.001)	(0.005)	(0.01)
Water Bill more than 5,000	0.054	0.453	-1.341
	(0.113)	(0.412)	(0.991)
House type: Precarious	-1.936***	0.315	6.194
	(0.541)	(0.986)	(6.413)
House type: Concrete	-1.521***	0.8	5.711
	(0.527)	(0.735)	(6.241)
House type: Rooming House	-1.332**	0.92	6.595
	(0.592)	(1.152)	(6.804)
Other households in compound	0.051	-0.02	0.406
	(0.033)	(0.114)	(0.305)
Own house	-0.351**	-0.91*	-0.077
	(0.153)	(0.499)	(1.476)
Pit meters from road	-0.005	0.038	-0.016
	(0.012)	(0.041)	(0.1)
More than 1 trip last desludging	0.452	5.91***	-4.643
	(0.359)	(0.88)	(3.853)
Electricity bill	0.038***		
	(0.006)		
Household size	0.006		
	(0.013)		
Number of women in household	0.045		
	(0.034)		
Respondent finished high school	0.424***		
	(0.138)		
atanh(ρ_{mech})	-0.798***		
	(0.247)		
atanh(ρ_{man})	-0.498***		
	(0.122)		
log(σ_{mech})	8.022***		
	(0.091)		
log(σ_{man})	4.274***		
	(0.039)		
N	773	530	243
LR test statistic: $-2\ln(\lambda)$	309.438***		
LR test statistic, Instruments: $-2\ln(\lambda)$	126.563***		

Selection equation estimated from households in the Demand Elicitation group who purchased a desludging prior to our survey. Mechanical (Manual) Price equation estimated from households in the Demand Elicitation group who purchased a mechanical (manual) desludging for their most recent desluding. Omitted housing type is concrete, multi-level. Estimated by Maximum Likelihood, model given in equations (9) to (10). $\rho_{0,man}$ ($\rho_{0,mech}$) is the correlation between the manual (mechanical) price shock and the selection shock; higher prices lead to a lower likelihood of purchasing mechanical. σ_{man} (σ_{mech}) is the standard deviation of the manual (mechanical) price shock. One correlation, $\rho_{mech,man}$ is not identified by the Tobit 5 model; we compute $corr(\varepsilon_{mech,i}, \varepsilon_{man,i})$ for a small number of households who did have prices for both services, and takes the value $-.01645$: the shocks are close to independent controlling for observables.

Table 3: Demand model: Logit Regressions

	10,000	15,000	17,500	20,000
Constant	-8.609***	-0.037	1.712***	2.264
	(0.443)	(0.254)	(0.634)	(1.415)
Average months between desludgings	0.003	0.001	0	0.001
	(0.007)	(0.002)	(0.004)	(0.027)
Water Bill more than 5,000	-0.134	0.07	-0.193	-0.807
	(0.233)	(0.163)	(0.369)	(1.176)
House type: Precarious	7.15***	0.275	-0.094	-0.345
	(0.422)	(0.221)	(0.516)	(1.031)
House type: Concrete	7.043***	0.093	0.281	0.339
	(0.413)	(0.198)	(0.453)	(0.784)
House type: Rooming House	6.974***	0.023	0.534	0.115
	(0.86)	(0.263)	(0.638)	(1.025)
Other households in compound	-0.086	0.011	-0.002	0.117
	(0.079)	(0.049)	(0.12)	(0.402)
Own house	0.39	0.275	0.389	0.382
	(0.392)	(0.209)	(0.395)	(0.842)
Pit meters from road	-0.01	-0.032*	-0.07***	-0.076
	(0.046)	(0.019)	(0.024)	(0.075)
More than 1 trip last desludging	-6.702***	-0.438	-0.69	-0.913
	(2.087)	(0.303)	(0.541)	(0.878)
Electricity bill	-0.022	-0.013*	-0.01	0
	(0.025)	(0.007)	(0.012)	(0.02)
Household size	0.039	0.017	0	0.052
	(0.029)	(0.025)	(0.054)	(0.09)
Number of women in household	-0.12*	-0.021	0.131	0
	(0.071)	(0.06)	(0.169)	(0.322)
Respondent finished high school	-0.266	-0.12	-0.088	-0.151
	(0.439)	(0.192)	(0.375)	(0.614)
N	773	773	773	773
% above, actual	0.846	0.481	0.135	0.105
% above, predicted	0.846	0.481	0.135	0.105

Estimated by non-linear least squares, standard errors bootstrapped with 5,000 repetitions. Each column corresponds to a logit regression predicting whether the household's offer, conditional on observables, is above the value given in the first row of the table.

Table 4: Offers, Summary Statistics

	Min	1Q	Median	Mean	3Q	Max
Offer, η:	2000	10000	12500	12995.472	15000	40000

Table 5: Model-Based Predicted Treatment Effects

	Control	Treatment Effect
Average	0.725	0.102
10,000	0.429	0.290
15,000	0.725	0.076
17,500	0.932	0.002
20,000	0.938	0.000

Control level and treatment effect computed based on (9) model and optimal Linear Programming prices.

Table 6: Deviations of Ordered Logit from Linear Programming Pricing Rule

	Min	1Q	Median	Mean	3Q	Max
Average	-5000	0	0	446	0	5000
10,000	0	0	0	1371	5000	5000
15,000:	-5000	0	0	189	0	5000
17,500	-2500	0	0	304	0	2500
20,000	-2500	-2500	-2500	-2014	-2500	0

Ordered logit rule mapping observables to prices, $t^*(x_i)$, computed by maximum likelihood. Standard errors omitted.

Table 7: Ordered Logit Pricing Rule

	10,000	15,000	17,500	20,000
Constant	-8.609***	-0.037	1.712***	2.264
	(0.443)	(0.254)	(0.634)	(1.415)
Average months between desludgings	0.003	0.001	0	0.001
	(0.007)	(0.002)	(0.004)	(0.027)
Water Bill more than 5,000	-0.134	0.07	-0.193	-0.807
	(0.233)	(0.163)	(0.369)	(1.176)
House type: Precarious	7.15***	0.275	-0.094	-0.345
	(0.422)	(0.221)	(0.516)	(1.031)
House type: Concrete	7.043***	0.093	0.281	0.339
	(0.413)	(0.198)	(0.453)	(0.784)
House type: Rooming House	6.974***	0.023	0.534	0.115
	(0.86)	(0.263)	(0.638)	(1.025)
Other households in compound	-0.086	0.011	-0.002	0.117
	(0.079)	(0.049)	(0.12)	(0.402)
Own house	0.39	0.275	0.389	0.382
	(0.392)	(0.209)	(0.395)	(0.842)
Pit meters from road	-0.01	-0.032*	-0.07***	-0.076
	(0.046)	(0.019)	(0.024)	(0.075)
More than 1 trip last desludging	-6.702***	-0.438	-0.69	-0.913
	(2.087)	(0.303)	(0.541)	(0.878)
Electricity bill	-0.022	-0.013*	-0.01	0
	(0.025)	(0.007)	(0.012)	(0.02)
Household size	0.039	0.017	0	0.052
	(0.029)	(0.025)	(0.054)	(0.09)
Number of women in household	-0.12*	-0.021	0.131	0
	(0.071)	(0.06)	(0.169)	(0.322)
Respondent finished high school	-0.266	-0.12	-0.088	-0.151
	(0.439)	(0.192)	(0.375)	(0.614)
N	773	773	773	773
% above, actual	0.846	0.481	0.135	0.105
% above, predicted	0.846	0.481	0.135	0.105

Estimated by non-linear least squares, standard errors bootstrapped with 5,000 repetitions. Each column corresponds to a logit regression predicting whether the household's offer, conditional on observables, is above the value given in the first row of the table.

Table 8: Balance Tests: Cluster and Household Level

	Cluster Level Averages		Household Level		Demand Elicitation	
	Control	Control-Treat	Control	Control-Treat	Control-DE	Treatment-DE
Household Size	6.798	0.107	7.862	0.379	0.070	0.308
	(1.14)	(0.23)	(4.66)	(0.32)	(0.30)	(0.272)
Number of Women in Household	2.439	0.045	2.742	0.034	-0.016	0.050
	(0.37)	(0.08)	(1.81)	(0.12)	(0.11)	(0.115)
Respondent Finished High School	0.316	0.037	0.278	0.009	0.048	-0.039
	(0.15)	(0.03)	(0.45)	(0.04)	(0.35)	(0.029)
Precarious Housing	0.119	0.001	0.107	0.010	-0.010	0.019
	(0.11)	(0.02)	(0.31)	(0.02)	(0.026)	(0.020)
Concrete Building	0.769	0.014	0.795	0.013	0.028	-0.016
	(0.12)	(0.02)	(0.40)	(0.03)	(0.029)	(0.028)
Rental Dormatories	0.046	-0.014	0.051	-0.031	-0.004	-0.027
	(0.05)	(0.01)	(0.22)	(0.02)	(0.01)	(0.015)
Own house	0.769	0.000	0.819	0.035	0.008	0.027
	(0.09)	(0.02)	(0.39)	(0.02)	(0.02)	(0.024)
Water bill more than 5,000 CFA	0.490	0.064*	0.563	0.106***	0.132***	-0.026
	(0.13)	(0.03)	(0.50)	(0.03)	(0.03)	(0.032)
Electricity Bill	14.266	1.744	13.808	0.389	0.787	-0.400
	(5.58)	(1.03)	(14.19)	(0.98)	(0.96)	(0.800)
Pit meters from Road	5.581	0.802*	5.281	0.857**	-0.371	1.228***
	(1.60)	(0.34)	(3.71)	(0.30)	(0.34)	(0.342)
More than 1 trip last desludging	0.024	0.016**	0.025	0.010	-0.001	0.011
	(0.03)	(0.01)	(0.16)	(0.01)	(0.01)	(0.007)
Average Months between desludgings	27.301	-2.799	21.960	-1.298	-1.678	0.380
	(7.98)	(1.59)	(26.84)	(1.86)	(1.83)	(1.74)
Other households in compound	1.112	-0.093	1.392	-0.349*	0.006	-0.355**
	(0.57)	(0.12)	(2.27)	(0.18)	(0.15)	(0.173)
Respondent Arranges Desludgings	0.576	0.018	0.610	0.039	-0.032	0.071
	(0.14)	(0.03)	(0.49)	(0.03)	(0.32)	(0.029)
Respondent is the Household Head	0.549	0.010	0.555	0.004	-0.005	0.010
	(0.12)	(0.02)	(0.50)	(0.04)	(0.04)	(0.030)
Years respondent lived in Compound	18.316	0.315	21.269	1.570	-0.181	1.751
	(5.48)	(1.06)	(13.92)	(1.33)	(1.24)	(1.116)
Number of households sharing pit	1.061	-0.061	1.338	-0.297	0.126	-0.422
	(0.55)	(0.12)	(2.21)	(0.17)	(0.13)	(0.149)
Ever used Manual Desludging	0.660	-0.034	0.544	-0.037	0.103***	-0.140***
	(0.11)	(0.03)	(0.50)	(0.03)	(0.03)	(0.030)
Ever used Mechanical Desludging	0.555	0.028	0.786	0.048	0.048	0.000
	(0.17)	(0.04)	(0.41)	(0.03)	(0.03)	(0.029)
Never desludged at this residence	0.313	-0.026	0.114	-0.045	-0.030	-0.015
	(0.16)	(0.03)	(0.32)	(0.02)	(0.02)	(0.023)
Percent of desludgings mech before BL	0.884	0.019	0.881	0.015	-0.252***	0.006
	(0.07)	(0.02)	(1.62)	(0.02)	(0.02)	(0.029)
Last Desludging was Mechanical	0.728	0.034	0.726	0.008	0.007	0.013
	(0.16)	(0.04)	(0.446)	(0.03)	(0.03)	(0.030)
Number of income earners	1.502	0.065	1.617	0.071	0.000	0.070
	(0.25)	(0.05)	(1.28)	(0.08)	(0.08)	(0.078)
Respondent Earns income	0.628	0.031	0.628	0.026	0.027	0.001
	(0.11)	(0.02)	(0.48)	(0.03)	(0.03)	(0.025)
Wealth Index (1st principal Component)	0.274	0.255*	0.408	0.224**	0.372***	-0.148
	(0.59)	(0.11)	(0.48)	(0.10)	(0.09)	(0.096)
N	40	52	551	648	840	840

The first column provides the variable average and standard deviation in the control group averaged at the neighborhood level. The second column provides the difference between the treatment group and the control group averages at the neighborhood level, with standard errors in parentheses. The third column provides the control group averages at the household level. The fourth column provides the difference between the control group and the demand elicitation group with standard errors in parentheses, and the fifth column provides the difference between the demand elicitation group and the treatment group with standard errors in parentheses. Standard errors are clustered at the neighborhood cluster level. Sample is restricted to the households that purchased any desludgings during the period to be consistent with the main regressions; 551 households for the control group, 648 households for the treatment group, and 840 households for the demand elicitation group.

Table 9: Baseline Balance Checks within price group

	(1) 10k	(2) 15k	(3) 17.5k	(4) 20k
Household Size	-0.690	-0.188	-0.972*	1.527
	(0.616)	(0.422)	(0.501)	(1.174)
Number of Women in Household	-0.079	0.020	-0.328	0.914
	(0.213)	(0.168)	(0.212)	(0.692)
Respondent Finished High School	0.166	-0.019	0.0333	-0.149
	(0.036)	(0.039)	(0.062)	(0.128)
Precarious Housing	-0.101	0.017**	0.00	0.00
	(0.080)	(0.008)	(0.00)	(0.00)
Concrete Building	0.094	-0.050**	-0.021	-0.016
	(0.081)	(0.023)	(0.060)	(0.119)
Rental Dormatories	0.008	0.037*	0.475	-0.003
	(0.008)	(0.020)	(0.491)	(0.453)
Own house	-0.001	-0.031	-0.098*	0.065
	(0.023)	(0.029)	(0.055)	(0.096)
Water bill more than 5,000 CFA	-0.039	-0.118**	-0.164***	0.037
	(0.063)	(0.031)	(0.059)	(0.056)
Electricity Bill	-0.129	-0.338	-0.718	1.784
	(0.744)	(0.532)	(1.767)	(4.695)
Pit meters from Road	-0.880***	-0.732*	-0.698**	-2.552
	(0.315)	(0.373)	(0.607)	(1.207)
More than 1 trip last desludging	0.000	-0.008	-0.021	-0.016
	(0.00)	(0.006)	(0.028)	(0.091)
Average Months between desludgings	3.543	0.505	1.248	-1.310
	(5.311)	(1.721)	(2.373)	(4.958)
Other households in compound	-0.098	0.363	0.904**	-0.457
	(0.248)	(0.222)	(0.359)	(0.330)
Respondent is the Arranger for Desludgings	0.101	-0.081*	-0.782	0.006
	(0.071)	(0.036)	(0.605)	(0.103)
Respondent is the Household Head	-0.008	-0.008	-0.049	0.154
	(0.072)	(0.044)	(0.071)	(0.140)
Years respondent has lived in Compound	-0.427	-1.38	-3.055	-1.947
	(2.102)	(1.556)	(1.895)	(3.215)
Ever used manual desludging	-0.058	0.099	-0.010	-0.137
	(0.057)	(0.040)	(0.059)	(0.127)
Ever used Mechanical Desludging	-0.074	-0.069*	0.027	-0.051
	(0.069)	(0.040)	(0.045)	(0.067)
Never desludged at this residence	0.014	0.082***	-0.0134	0.051
	(0.483)	(0.030)	(0.040)	(0.067)
Last Used Mechanical	-0.053	-0.081*	0.044	0.016
	(0.068)	(0.041)	(0.054)	(0.080)
Number Income Earners	0.091	-0.098	-0.218	0.214
	(0.168)	(0.094)	(0.155)	(0.354)
Respondent Earns Income	0.079	-0.080**	-0.018	0.064
	(0.062)	(0.036)	(0.054)	(0.090)
Wealth Index	-0.347***	-0.149	-0.308	0.046
	(0.143)	(0.122)	(0.211)	(0.451)
Percent Mechanical prior to BL	-0.007	-0.078**	-0.042	-0.030
	(0.060)	(0.040)	(0.496)	(0.112)
N	244	616	275	64

The columns provide the difference between the treatment group and the control group in the 10k-20k price bin, with standard errors in parentheses. Standard errors are clustered at the neighborhood cluster level. Sample is restricted to the 1,199 households that purchased any desludgings during the period to be consistent with the main regressions.

Table 10: Call Center Take Up

Targeted Price Level	10000	15000	17500	20000	Total
Pct Offered Price	28	49	18	4	100
Deposited	55	52	35	38	49
Percent take-up through CC					
1st 6 months	100	58	78	75	70
Percent take-up through CC					
(from deposited and desludged)	62	47	38	47	50
Modeled Take up	94	59	33	0	58

Shown are percentages of each group. "Percent offered price" is the percent of the treatment group that were offered each of the price levels in accordance with the price targeting model. "Deposited" is the percent of those offered each price who accepted the price offer and paid a deposit. "Percent take-up through Call Center 1st 6 months" is the percentage of people who called the call center from among those that paid the deposit and ended up purchasing a desludging–separated between those who purchased a desludging in the first 6 months of the program and those that purchased a desludging at some point between baseline and endline. "Modeled take-up" is the expected level of take-up generated from the pricing model.

Table 11: Reasons Households did not Call the Call Center

	Targeted Price
Didn't need a desludging	368
Forgot about it	60
Better Outside option	59
Too Confusing/didn't understand	46
New to the compound	24
Not in charge of desludging	32
Other/refusal	20
Total	606

Households were able to select multiple responses. Sample restricted to treatment households that paid a deposit at baseline but did not use the call center between baseline and endline.

Table 12: Market Share Effects of Treatment

Dependent Variable:	Market share Control Group, at EL	Market share OLS	LASSO	Mkt share, cluster-price level OLS	LASSO
	(1)	(2)	(3)	(4)	(5)
Overall	0.840	0.044*	0.039*		
		(0.023)	(0.023)		
Treat × 10k group				0.096*	0.079*
				(0.049)	(0.047)
Treat × 15k group				0.017	0.015
				(0.046)	(0.044)
Treat × 17.5k group				-0.002	0.002
				(0.049)	(0.048)
Treat × 20k group				-0.036	-0.033
				(0.070)	(0.070)
10k group	0.589			0.297***	0.000
				(0.094)	(.)
15k group	0.852			0.426***	0.119
				(0.113)	(0.096)
17.5k group	0.906			0.444***	0.079
				(0.125)	(0.145)
20k group	0.993			0.490***	0.127
				(0.145)	(0.216)
N		92	92	300	300
R^2		0.343		0.942	
$mean$		0.852	0.852	0.832	0.832

Column (1) gives the average neighborhood market share for each price group of mechanical desludging at endline for the control group where market share is defined as $\frac{\#mechanicaldesludgings}{\#mechanical+\#manualdesludgings}$. Column (2) provides the OLS estimate of the pooled effect with observations at the neighborhood cluster level. Column (3) provides the LASSO estimate of the pooled effect with observations at the cluster level, with 119 potential control variables. Column (4) gives the OLS estimate for the market share effect for each price group in a neighborhood cluster (level of observation is neighborhood-price, but not all neighborhoods include households from each price group). Column (5) gives the LASSO estimate for the market share effect for each price group in a neighborhood cluster, with 119 potential control variables. Controls are included in the OLS regressions for the neighborhood means (in column 2) or neighborhood-price group means (in column 4) of the variables not balanced at baseline at the household level (water bill more than 5,000 CFA, latrine pit distance to road, two tanks used last desludging, other households in compound, wealth index). Percent of desludgings for which the household purchased mechanical, last desludging mechanical (dummy) and never desludged (dummy) are included as controls for past desludging behavior of the household. A control has also been included for the average of the stratification variable–above median number of households in neighborhood have high compound walls. The 10k price group has been omitted in the LASSO specification because of collinearity with the pdslasso selected control variables.

Table 13: Household Level Market Share: Effects on the Percent of Mechanical at the Household level

	(1) Pct mech	(2) Pct mech	(3) Pct mech	(4) Pct mech
Targeted Price Group	0.032	0.030		
	(0.022)	(0.021)		
Treat × 10k group			0.074	0.085*
			(0.052)	(0.047)
Treat × 15k group			0.017	0.026
			(0.028)	(0.028)
Treat × 17.5k group			0.020	0.026
			(0.035)	(0.033)
Treat × 20k group			0.028	0.050
			(0.058)	(0.055)
10k group			0.435***	0.524***
			(0.057)	(0.159)
15k group			0.513***	0.485***
			(0.058)	(0.145)
17k group			0.580***	0.518***
			(0.061)	(0.150)
20k group			0.583***	0.493***
			(0.078)	(0.162)
N	1199	1199	1199	1199
R^2	0.210		0.855	
Mean	0.804	0.804	0.804	0.804

Observations are at the household level. The constant is dropped in specifications (3) and (4). Included in the OLS specifications, but not shown for brevity are controls for the variables not balanced at baseline: water bill more than 5,000 CFA, latrine pit distance to road, two tanks used last desludging, other households in compound, and wealth index. Percent of desludgings for which the household purchased mechanical prior to the baseline, last desludging mechanical (dummy) and never desludged (dummy) are included as controls for past desludging behavior of the household. The LASSO specifications have 119 potential control variables. A control for the stratification variable: less than half of compound walls in the neighborhood are high, is also included. Standard errors, clustered by neighborhood, are in parentheses.

Table 14: Decomposition of Market Share: Effects on purchases of Mechanical and Manual

	(1) Any Mechanical	(2) Any Mechanical	(3) Any Mechanical	(4) Any Mechanical	(5) Any Manual	(6) Any Manual	(7) Any Manual	(8) Any Manua
Targeted price group	0.035	0.033			-0.027	-0.023		
	(0.022)	(0.022)			(0.021)	(0.021)		
Treat × 10k group			0.070	0.082*			-0.076	-0.082*
			(0.053)	(0.049)			(0.055)	(0.050)
Treat × 15k group			0.027	0.035			-0.007	-0.014
			(0.029)	(0.029)			(0.027)	(0.027)
Treat × 17.5k group			0.007	0.008			-0.040	-0.048
			(0.033)	(0.031)			(0.035)	(0.034)
Treat × 20k group			0.062	0.081			0.051	0.026
			(0.054)	(0.052)			(0.063)	(0.064)
10k group			0.450***	0.498***			0.589***	0.407**
			(0.059)	(0.161)			(0.059)	(0.189)
15k group			0.529***	0.471***			0.490***	0.425**
			(0.059)	(0.150)			(0.058)	(0.179)
17k group			0.609***	0.521***			0.439***	0.394**
			(0.063)	(0.153)			(0.065)	(0.169)
20k group			0.586***	0.478***			0.395***	0.390**
			(0.080)	(0.164)			(0.077)	(0.170)
N	1199	1199	1199	1199	1199	1199	1199	1199
R^2	0.198		0.860		0.179		0.349	
Mean	0.804	0.804	0.804	0.804	0.804	0.196	0.804	0.196

Observations are at the household level. Included in the OLS specifications, but not shown for brevity are controls for the variables not balanced at baselir water bill more than 5,000 CFA, latrine pit distance to road, two tanks used last desludging, other households in compound, and wealth index. Percent desludgings for which the household purchased mechanical prior to the baseline, last desludging mechanical (dummy) and never desludged (dummy) are includ as controls for past desludging behavior of the household. A control for the stratification variable: less than half of compound walls in the neighborhood are hig is also included but not shown. The LASSO specification has 119 potential control variables. Standard errors, clustered by neighborhood, are in parentheses

Table 15: Impact of Treatment on Children's Diarrhea

	Dependent Variable: Child Diarrhea					
	(1)	(2)	(3)	(4)	(5)	(6)
Targeted group	-0.015	-0.014				
	(0.017)	(0.017)				
Treatment*10k group			-0.060*	-0.046		
			(0.030)	(0.032)		
Treat × 15k group			0.006	0.012		
			(0.021)	(0.020)		
Treat × 17.5k group			-0.005	0.002		
			(0.036)	(0.035)		
Treat × 20k group			0.012	0.009		
			(0.059)	(0.057)		
10k group			0.129***	-0.071		
			(0.036)	(0.142)		
15k group			0.090**	-0.087		
			(0.037)	(0.137)		
17k group			0.121***	-0.030		
			(0.045)	(0.138)		
20k group			0.114**	-0.024		
			(0.057)	(0.130)		
Treat × Pct Nbhd in 10k group					-0.217**	-0.207**
					(0.096)	(0.095)
Treat × Pct Nbhd in 15k group					0.078	0.091
					(0.076)	(0.074)
Treat × Pct Nbhd in 17k group					-0.037	-0.059
					(0.148)	(0.137)
Treat × Pct Nbhd in 20k group					0.265	0.297
					(0.314)	(0.297)
Pct Nbhd in 10k group					0.636	0.861
					(0.625)	(0.649)
Pct Nbhd in 15k group					0.491	0.716
					(0.634)	(0.666)
Pct Nbhd in 17.5k group					0.595	0.834
					(0.637)	(0.665)
Pct Nbhd in 20k group					0.260	0.434
					(0.656)	(0.671)
N	2017	2017	2017	2017	2017	2017
R^2	0.025		0.155		0.030	
Mean	0.131	0.131	0.131	0.131	0.131	0.131

Observations are at the household level, standard errors are clustered at the neighborhood cluster level. The dependent variable is whether the participant reports that a child in the household has had diarrhea in the past 7 days at endline. Odd specifications are estimated with OLS, even specifications with LASSO. Specifications (1) and (2) are the pooled effect across all price groups. Specifications (3) and (4) provide the estimates of the effect for each price group. Specifications (5) and (6) show the spillover effects of the treatment within the neighborhood: the regressions control for the percent of the neighborhood that would be assigned to each price group, and estimate the effect of the treatment for each price group. The sample includes only households with children. The diarrhea question was posed as follows: "In the past seven days, of the children in your household, how many had diarrhea, even once?" Children are defined in the survey as being 14 and younger, if the number is 1 or more, the variable is classified as a 1, otherwise 0. Included in the OLS specification, but not shown for brevity are controls for the variables not balanced at baseline: water bill more than 5,000 CFA, latrine pit distance to road, two tanks used last desludging, other households in compound, and wealth index. Percent of desludgings for which the household purchased mechanical, last desludging mechanical (dummy) and never desludged (dummy) are included as controls for past desludging behavior of the household, and

Table 16: Mean Baseline Characteristics by Price Group

	10000	15000	17500	20000	Pooled
Phone Credit use over past week	1107	1754	4078	4882	2157
	(1929)	(2930)	(5660)	(10195)	(4152)
Number of Refrigerators	0.168	0.507	0.927	1.371	0.530
	(0.430)	(0.665)	(0.771)	(0.726)	(0.706)
Number of Cars	0.061	0.298	0.671	1.529	0.357
	(0.248)	(0.577)	(0.838)	(1.073)	(0.683)
Number of Air Conditioners	0.016	0.081	0.477	1.486	0.199
	(0.157)	(0.359)	(1.004)	(1.909)	(0.722)
Ever Desludged Mech	0.357	0.571	0.621	0.686	0.524
	(0.479)	(0.495)	(0.486)	(0.468)	(0.499)
Expected Price Mechanical (CFA)	12792	14103	15847	16716	14243
	(4717)	(4743)	(5550)	(7120)	(5173)
Last used Manual	0.510	0.219	0.153	0.030	0.263
	(0.500)	(0.414)	(0.361)	(0.171)	(0.440)

This table provides means for each variable at baseline by the price group to which they were assigned by the pricing model. Standard deviations are in parentheses.

Table 17: Deposit, Mechanical, and Arrange Regressions

	$Deposit_i$				$Arrange_i$	
	Coefficient	M. Effect	m_i^0	m_i^1	Coefficient	M. Effect
Constant	-0.375		0.184	0.409	-1.698**	
	(0.698)		(0.211)	(0.768)	(0.742)	
Wealth Index	0.035	0.014	-0.001	0.023	-0.016	-0.004
	(0.035)	(0.014)	(0.012)	(0.026)	(0.039)	(0.011)
Last Desludging Mechanical	0.085	0.033	0.248***	0.363**	0.357*	0.092*
	(0.186)	(0.073)	(0.08)	(0.161)	(0.207)	(0.05)
Never Desludged	0.289	0.11	-0.043	0.355	0.07	0.02
	(0.186)	(0.068)	(0.091)	(0.316)	(0.213)	(0.06)
% Mechanical at Baseline	0.176	0.069	0.046	0.077	-0.169	-0.046
	(0.167)	(0.065)	(0.063)	(0.181)	(0.186)	(0.05)
Respondent Age	0.018***	0.007***	-0.002	0.004	0.011***	0.003***
	(0.004)	(0.001)	(0.003)	(0.016)	(0.004)	(0.001)
$\hat{\eta}$	0.129***	0.051***	-0.01	0.016	0.158***	0.043***
	(0.05)	(0.02)	(0.019)	(0.093)	(0.052)	(0.014)
Weight	-0.066	-0.026	0.043*	0.004	-0.045	-0.012
	(0.06)	(0.024)	(0.025)	(0.057)	(0.065)	(0.018)
Reliable Responses	-0.459	-0.181			-0.14	-0.036
	(0.351)	(0.135)			(0.424)	(0.101)
Price	-0.076*	-0.03*		-0.022	-0.08*	-0.022*
	(0.042)	(0.016)		(0.096)	(0.044)	(0.012)
$\hat{\xi}$			-0.001	-0.226		
			(0.156)	(0.835)		
$\hat{\xi}^2$			0.109			
			(0.11)			
N	648	648	278	370	648	648
R^2			0.455	0.71		
LR Stat.	63.119***				33.66***	

Columns 1 and 2 provide probit estimates and marginal effects at the mean for (28), column 3 provides OLS estimates for 31, column 4 provides OLS estimates for 32, and columns 5 and 6 provide probit estimates and marginal effects at the mean for 33. Variables m_i^0 and m_i^1 correspond to $PctMechanical_i$, conditional on depositing or not, respectively. Variable $\hat{\eta}$ is predicted willingness-to-switch, and ξ is the predicted latent variable from the deposit regression. Reliable Responses is the excluded instrument in specifications 3 and 4. Standard errors in columns 1, 2, 5, and 6 computed analytically, and in columns 3 and 4, by the bootstrap.

Table 18: Model Fit, Household Decisions

	Deposit	$\widehat{\text{Deposit}}$	m^0	$\hat{m}^0$	m^1	$\hat{m}^1$	Arrange	$\widehat{\text{Arran}}$
Average	57.1	57	81.3	81.3	80.9	80.3	20.2	20.2
	(54.5,59.8)	(54.4,59.8)	(79.5,83.3)	(79.4,83.2)	(78.4,83.5)	(77.7,82.72)	(18.5,22)	(18.5,2
10,000	70.7	72	71.7	71	61.5	63.5	30.1	31
	(67.3,74.4)	(68.4,75.8)	(67.8,75.82)	(67.7,74.8)	(54.8,68)	(57.8,69.1)	(27,33.3)	(28,34
15,000	60.7	58.3	82.1	82.9	78.6	76.5	21	20
	(57.2,64.1)	(55.5,61)	(79.9,84)	(80.7,85)	(75.2,82.4)	(73.4,79.6)	(18.5,23.4)	(18.1,2
17,500	38.5	43.3	89.5	88.2	89.8	90	11.5	12.8
	(34.3,43.12)	(40.2,47.4)	(85.7,93.8)	(85.9,90.4)	(86.9,92.5)	(87,91.9)	(9.3,14)	(11.1,1
20,000	50	45.4	98	96.1	95.6	96	11.8	12.2
	(40.9,60)	(38.7,49.9)	(96.3,100)	(91.8,98.8)	(92.5,100)	(93.6,98.4)	(5.3,18.2)	(9,15.1

Compares actual and predicted household desludging decisions based on counterfactual model. Variables m_i^0 an
correspond to $PctMechanical_i$, conditional on depositing or not, respectively. Bootstrapped 90% confidence inte
reported below point estimate.

Table 19: Model Fit, Mechanical Shares

	Realized	$\widehat{\text{Treatment}}$	$\widehat{\text{Control}}$
Average	81.1	80.8	76.7
	(79.5,82.8)	(79.1,82.3)	(76.1,78)
10,000	68.7	68.6	57.9
	(65.38,72.4)	(65.8,71.8)	(57,61.1)
15,000	80.7	80.4	77.9
	(78.68,82.7)	(78.3,82.3)	(77,79.1)
17,500	89.7	89.3	86.5
	(87.3,91.9)	(87.3,90.7)	(85.3,88.1)
20,000	96.8	95.9	95.9
	(94.8,98.8)	(93.74,97.4)	(94.9,98.2)

Realized mechanical market shares in the intervention, predicted treatment based on counterfactual model, and predicted control based on LASSO regression on the control group predicted for the Targeted Pricing group. Bootstrapped 90% confidence intervals reported below point estimate.

Table 20: Model Fit, Finances

	Profit	$\widehat{\text{Profit}}$	Budget	$\widehat{\text{Budget}}$	SR	$\widehat{\text{SR}}$
Average	-433	-433	-102	-116	-3	-3
	(-489,-374)	(-488,-375)	(-150,-53)	(-146,-84)	(-3,-2)	(-3,-2)
10,000	-1830	-1917	-1334	-1043	-11	-12
	(-2027,-1635)	(-2114,-1732)	(-1477,-1189)	(-1179,-919)	(-13,-10)	(-13,-11)
15,000	-231	-214	116	70	-1	-1
	(-267,-189)	(-240,-185)	(91,144)	(59,85)	(-2,-1)	(-2,-1)
17,500	156	193	334	190	1	1
	(125,188)	(167,228)	(273,401)	(160,232)	(1,1)	(1,1)
20,000	495	495	689	343	3	3
	(222,764)	(367,612)	(308,1064)	(232,447)	(1,5)	(2,4)

Platform profit equals $Profit = \frac{1}{N}\sum_{i=1}^{N} Arrange_i \times (t_i - c_i)$, budget balance equals $BB = \frac{1}{N}\sum_{i=1}^{N} Arrange_i \times (t_i - c_i + s)$, and subsidization rate equals $SR = \frac{1}{N}\sum_{i=1}^{N} Arrange_i \times \frac{(t_i - c_i)}{c_i}$, where c_i is the cost of procurement for household i. Columns 1-4 are in CFA, columns 5-6 are percentages. Bootstrapped 90% confidence intervals reported below point estimate.

Table 21: Probit Regression Without Selection

	$PctMechanical_i$	M. Effect
Constant	-1.35	
	(0.946)	
Wealth Index	0.054	0.013
	(0.046)	(0.011)
Last Desludging Mechanical	0.99***	0.276***
	(0.207)	(0.064)
Never Desludged	0.398**	0.083**
	(0.189)	(0.034)
% Mechanical at Baseline	0.2	0.048
	(0.214)	(0.052)
Respondent Age	0	0
	(0.004)	(0.001)
$\hat{\eta}$	-0.012	-0.003
	(0.062)	(0.015)
Weight	0.195***	0.047***
	(0.072)	(0.017)
Reliable Responses	0.176	0.043
	(0.137)	(0.033)
Arranger	-0.17	-0.041
	(0.129)	(0.03)
Price	-0.079*	-0.019*
	(0.048)	(0.011)
N	648	648

Provides estimates of probit regression of $PctMechanical_i$ directly on covariates, without semi-parametric selection. Estimated by maximum likelihood. Variable $\hat{\eta}$ is predicted willingness-to-switch.

Table 22: Model Fit, No selection

	Realized	$\widehat{\text{Control}}$	$\widehat{\text{Treatment}}$	Auction	Auction (S)	PMT(100)	PMT(150)	Ceiling	Ceiling(S)
Average	81.1	76.8	80.4	76.8	79.7	52.7	65.8	75.4	78.4
10,000	68.7	58	67.3	52.1	56.3	46.4	53.7	50.4	54.6
15,000	80.7	77.9	80.6	78.5	81.7	59.1	72.2	77	80.3
17,500	89.7	86.6	88.4	90.3	92.1	42.2	59.4	89.3	91.2
20,000	96.8	96.1	95.2	97.5	98.1	60.8	77.9	97.1	97.8

Market shares for alternative platform designs, on average and by price group. Variable definitions given on page 6.1.

Table 23: Counterfactual Shares: Alternative Market Designs

	Realized	$\widehat{\text{Control}}$	$\widehat{\text{Treatment}}$	Auction	Auction(S)	Proxy-Means(100)	Proxy-Means
Average	81.1	76.7	80.8	78.8	80.7	79.3	80.4
	(79.5,82.8)	(76.1,78)	(79.1,82.3)	(75.1,80.8)	(78.6,82.2)	(77.6,80.8)	(77.9,81.
10,000	68.7	57.9	68.6	60.2	62	61.6	62.2
	(65.38,72.4)	(57,61.1)	(65.8,71.8)	(50,65.7)	(53.4,66.5)	(54.5,65.9)	(53.94,66.
15,000	80.7	77.9	80.4	79	81.3	80.1	80.9
	(78.68,82.7)	(77,79.1)	(78.3,82.3)	(75.6,81.2)	(79.5,83.36)	(78.8,81.96)	(78.9,82.
17,500	89.7	86.5	89.3	90.7	92.2	89.2	91.8
	(87.3,91.9)	(85.3,88.1)	(87.3,90.7)	(88.5,92.8)	(89.4,94.4)	(87.8,91)	(89.2,93.
20,000	96.8	95.9	95.9	98.1	98.4	97.2	98
	(94.8,98.8)	(94.9,98.2)	(93.74,97.4)	(95.9,98.9)	(96.4,99.1)	(95.8,98.5)	(95.5,98.

Gives market shares for alternative platform designs. Realized is the experimental value, $\widehat{\text{Control}}$ is the lasso-pre group using control group data, $\widehat{\text{Treatment}}$ is the model predicted value, Auction is the predicted outcome when ho auction prices, Proxy-Means(100) and Proxy-Means(150) refer to predicted outcomes using proxy-means testing at Cieling is a policy that constrains prices to be the average price in the search market, and (S) denotes an additional confidence intervals reported below point estimate.

Table 24: Counterfactual Subsidization Rates: Alternative Market Designs

	Realized	$\widehat{\text{Treatment}}$	Auction	Auction (S)	Proxy-Means(100)	Proxy-Means(150)	Ceiling	Ceiling(S)
Average	-2.66	-2.66	0	-2.12	-1.43	-1.79	0.76	-1.02
	(-3.01,-2.31)	(-3,-2.31)	(0,0)	(-2.33,-1.93)	(-1.57,-1.3)	(-1.95,-1.58)	(0.64,0.93)	(-1.15,-0.87)
10,000	-11.35	-11.82	0	-2.04	-1.7	-1.97	0.63	-1.12
	(-12.58,-10.17)	(-13.03,-10.67)	(0,0)	(-2.55,-1.63)	(-2.13,-1.34)	(-2.38,-1.54)	(0.42,0.98)	(-1.44,-0.83)
15,000	-1.41	-1.31	0	-2.2	-1.57	-1.95	0.79	-1.06
	(-1.62,-1.16)	(-1.47,-1.14)	(0,0)	(-2.38,-2)	(-1.71,-1.42)	(-2.12,-1.72)	(0.65,0.99)	(-1.2,-0.9)
17,500	0.98	1.22	0	-1.93	-0.88	-1.23	0.75	-0.83
	(0.78,1.19)	(1.06,1.45)	(0,0)	(-2.42,-1.51)	(-1.07,-0.66)	(-1.51,-0.9)	(0.62,0.91)	(-1.03,-0.66)
20,000	3.15	3.12	0	-2.42	-1.43	-1.97	1.02	-0.98
	(1.47,4.86)	(2.31,3.85)	(0,0)	(-3.32,-1.49)	(-1.95,-0.81)	(-2.69,-1.19)	(0.68,1.33)	(-1.4,-0.61)

Realized is the experimental value, $\widehat{\text{Treatment}}$ is the model predicted value, Auction is the predicted outcome when households are given the procurement auction prices, Proxy-Means(100) and Proxy-Means(150) refer to predicted outcomes using proxy-means testing at 100% and 150% of the poverty line, Cieling is a policy that constrains prices to be the average price in the search market, and (S) denotes an additional subsidy of $3.00. Bootstrapped 90% confidence intervals reported below point estimate.

Table 25: Counterfactual Budget Balance: Alternative Market Designs

	Realized	$\widehat{\text{Treatment}}$	Auction	Auction (S)	Proxy-Means(100)	Proxy-Means(150)	Ceil
Average	-102	-100	0	0	0	-230	78
	(-150,-53)	(-147,-52)	(0,0)	(0,0)	(0,0)	(-251,-208)	(64,
10,000	-1334	-1404	0	0	0	-275	82
	(-1477,-1189)	(-1551,-1267)	(0,0)	(0,0)	(0,0)	(-343,-216)	(54,1
15,000	116	116	0	0	0	-252	86
	(91,144)	(98,138)	(0,0)	(0,0)	(0,0)	(-275,-228)	(70,1
17,500	334	404	0	0	0	-140	53
	(273,401)	(352,473)	(0,0)	(0,0)	(0,0)	(-171,-106)	(42,
20,000	689	696	0	0	0	-226	98
	(308,1064)	(516,861)	(0,0)	(0,0)	(0,0)	(-309,-129)	(59,1

Realized is the experimental value, $\widehat{\text{Treatment}}$ is the model predicted value, Auction is the predicted outcome wh
given the procurement auction prices, Proxy-Means(100) and Proxy-Means(150) refer to predicted outcomes using pr
at 100% and 150% of the poverty line, Cieling is a policy that constrains prices to be the average price in the searc
denotes an additional subsidy of $3.00. Bootstrapped 90% confidence intervals reported below point estimate.

Table 26: Design Variations

	π^*_{ATE}	$\pi^*_{\text{ATE, 10k group}}$
Auction	1704	6132
	(1530,2254)	(6017,6257)
PMT(100)	1665	5407
	(907,4176)	(4747,7338)
PMT(150)	350	4518
	(44,1292)	(4331,4818)
Market	2478	6825
	(2321,2973)	(6807,6875)

Design variations for alternative market designs in CFA. Gives the additional subsidy requires for each alternative market design to match the Targeted Pricing treatment effect on average (column 1) and in the 10,000 CFA price bin (column 2).

Table 27: Design Variations: Mechanical Share

	Realized	$\widehat{\text{Control}}$	$\widehat{\text{Treatment}}$	Auction	Auction(S)	Proxy-Means(100)	Proxy-Means(150)	Ceiling	Ceiling(S)
Average	81.1	76.7	80.8	80.8	80.8	80.8	80.8	80.8	80.8
	(79.5,82.8)	(76.1,78)	(79.1,82.3)	(79.1,82.3)	(79.1,82.4)	(79.1,82.3)	(79.1,82.4)	(79.1,82.3)	(79.1,82.3)
10,000	68.7	57.9	68.6	62.1	62.1	63.5	62.6	62.2	62.2
	(65.38,72.4)	(57,61.1)	(65.8,71.8)	(54.2,66.7)	(54.2,66.7)	(57.04,67.5)	(55.5,67.1)	(54.34,66.7)	(54.34,66.7)
15,000	80.7	77.9	80.4	81.4	81.4	82	81.3	81.4	81.4
	(78.68,82.7)	(77,79.1)	(78.34,82.3)	(79.7,84)	(79.7,84)	(80.1,84.4)	(79.6,84.1)	(79.7,84)	(79.7,84)
17,500	89.7	86.5	89.3	92.3	92.3	89.7	92	92.2	92.2
	(87.3,91.9)	(85.3,88.1)	(87.3,90.7)	(89.4,94.5)	(89.44,94.6)	(88.3,91.4)	(89.4,93.8)	(89.4,94.5)	(89.4,94.5)
20,000	96.8	95.9	95.9	98.5	98.5	97.3	98	98.4	98.4
	(94.8,98.8)	(94.9,98.2)	(93.74,97.4)	(96.3,99.1)	(96.4,99.1)	(96,98.6)	(95.6,98.8)	(96.3,99.1)	(96.3,99.1)

Gives design variations for alternative platform designs. Realized is the experimental value, $\widehat{\text{Control}}$ is the lasso-predicted value for the targeted pricing group using control group data, $\widehat{\text{Treatment}}$ is the model predicted value, Auction is the predicted outcome when households are given the procurement auction prices, Proxy-Means(100) and Proxy-Means(150) refer to predicted outcomes using proxy-means testing at 100% and 150% of the poverty line, Cieling is a policy that constrains prices to be the average price in the search market, and (S) denotes an additional subsidy of $3.00. Bootstrapped 90% confidence intervals reported below point estimate.

Table 28: Design Variations, 10,000 CFA Group: Mechanical Share

	Realized	$\widehat{\text{Control}}$	$\widehat{\text{Treatment}}$	Auction	Auction(S)	Proxy-Means(100)	Proxy-Mean
Average	81.1	76.7	80.8	86.4	86.4	84.1	85.4
	(79.5,82.8)	(76.1,78)	(79.1,82.3)	(82.7,89.2)	(82.7,89.2)	(81.8,86.16)	(82.1,87
10,000	68.7	57.9	68.6	68.6	68.6	68.6	68.6
	(65.38,72.4)	(57,61.1)	(65.8,71.8)	(65.8,71.8)	(65.8,71.8)	(65.74,71.8)	(65.8,71
15,000	80.7	77.9	80.4	88.4	88.4	86.1	87.2
	(78.68,82.7)	(77,79.1)	(78.3,82.3)	(83.5,92.3)	(83.5,92.3)	(82.7,88.6)	(82.6,90
17,500	89.7	86.5	89.3	94.9	94.9	90.5	93.4
	(87.3,91.9)	(85.3,88.1)	(87.3,90.7)	(91,96.8)	(91,96.8)	(89,92.2)	(90.3,9
20,000	96.8	95.9	95.9	99	99	97.5	98.4
	(94.8,98.8)	(94.9,98.2)	(93.74,97.4)	(97.24,99.6)	(97.24,99.6)	(96.3,98.9)	(96.2,99

Gives design variations for alternative platform designs. Realized is the experimental value, $\widehat{\text{Control}}$ is the lasso-pr group using control group data, $\widehat{\text{Treatment}}$ is the model predicted value, Auction is the predicted outcome when h auction prices, Proxy-Means(100) and Proxy-Means(150) refer to predicted outcomes using proxy-means testing at Cieling is a policy that constrains prices to be the average price in the search market, and (S) denotes an additional confidence intervals reported below point estimate.

Table 29: Design Variations: Subsidization Rate

	Realized	$\widehat{\text{Treatment}}$	Auction	Auction(S)	PMT(100)	PMT(150)	Market	Market(S)
Average	-2.66	-2.66	-2.2	-2.2	-3.42	-2.25	-2.18	-2.18
	(-3.01,-2.31)	(-3.1,-2.33)	(-3.07,-1.93)	(-3.08,-2.08)	(-7.14,-2.36)	(-3.52,-1.85)	(-3.01,-1.91)	(-3.02,-1.92)
10,000	-11.35	-11.82	-2.12	-2.12	-4.08	-2.48	-2.26	-2.26
	(-12.58,-10.17)	(-13.34,-10.75)	(-3.33,-1.72)	(-3.34,-1.83)	(-9.45,-2.68)	(-4.15,-1.96)	(-3.42,-1.85)	(-3.44,-1.86)
15,000	-1.41	-1.31	-2.28	-2.28	-3.75	-2.46	-2.27	-2.27
	(-1.62,-1.16)	(-1.53,-1.15)	(-3.18,-1.99)	(-3.18,-2.16)	(-7.71,-2.6)	(-3.83,-2.02)	(-3.11,-1.98)	(-3.11,-1.98)
17,500	0.98	1.22	-2.01	-2.01	-2.11	-1.55	-1.88	-1.88
	(0.78,1.19)	(1.01,1.44)	(-3.15,-1.57)	(-3.15,-1.67)	(-4.39,-1.33)	(-2.62,-1.14)	(-2.87,-1.48)	(-2.87,-1.49)
20,000	3.15	3.12	-2.52	-2.52	-3.38	-2.48	-2.28	-2.28
	(1.47,4.86)	(2.2,3.83)	(-4.31,-1.55)	(-4.36,-1.63)	(-6.88,-1.74)	(-4.5,-1.54)	(-3.88,-1.44)	(-3.9,-1.45)

Realized is the experimental value, $\widehat{\text{Treatment}}$ is the model predicted value, Auction is the predicted outcome when households are given the procurement auction prices, Proxy-Means(100) and Proxy-Means(150) refer to predicted outcomes using proxy-means testing at 100% and 150% of the poverty line, Cieling is a policy that constrains prices to be the average price in the search market, and (S) denotes an additional subsidy of $3.00. Bootstrapped 90% confidence intervals reported below point estimate.

Table 30: Design Variations, 10,000 CFA Group: Subsidization Rate

	Realized	$\widehat{\text{Treatment}}$	Auction	Auction(S)	PMT(100)	PMT(150)
Average	-2.66	-2.66	-12.14	-12.14	-10.26	-10.35
	(-3.01,-2.31)	(-3.09,-2.34)	(-15.46,-9.44)	(-15.46,-9.44)	(-15.08,-7.71)	(-13.2,-8.02)
10,000	-11.35	-11.82	-11.78	-11.78	-12.31	-11.47
	(-12.58,-10.17)	(-13.3,-10.76)	(-13.25,-10.71)	(-13.25,-10.71)	(-17.94,-10.13)	(-13.31,-10.17)
15,000	-1.41	-1.31	-12.51	-12.51	-11.16	-11.2
	(-1.62,-1.16)	(-1.52,-1.15)	(-15.56,-9.95)	(-15.56,-9.95)	(-16.24,-8.36)	(-14.15,-8.73)
17,500	0.98	1.22	-11.29	-11.29	-6.44	-7.27
	(0.78,1.19)	(1.01,1.44)	(-16.59,-7.34)	(-16.59,-7.34)	(-10.28,-4.06)	(-10.68,-4.63)
20,000	3.15	3.12	-13.52	-13.52	-9.99	-11.04
	(1.47,4.86)	(2.19,3.84)	(-20.91,-7.17)	(-20.92,-7.17)	(-16.33,-5.12)	(-17.38,-5.8)

Realized is the experimental value, $\widehat{\text{Treatment}}$ is the model predicted value, Auction is the predicted outcome w procurement auction prices, Proxy-Means(100) and Proxy-Means(150) refer to predicted outcomes using proxy-me of the poverty line, Cieling is a policy that constrains prices to be the average price in the search market, and (S) d $3.00. Bootstrapped 90% confidence intervals reported below point estimate.

Table 31: Counterfactual Shares: Alternative Information Designs

	Realized	$\widehat{\text{Control}}$	Original(0)	Conservative(1)	Municipal(2)	Municipal(2B)	NGO(3)	NGO(3B)
Average	81.1	76.7	80.6	80.6	81.5	82.1	80.7	77.6
	(79.5,82.8)	(76.1,78)	(78.6,82)	(78.6,82.1)	(79.6,82.8)	(80.1,83.5)	(78.6,82)	(72.44,80.3)
10,000	68.7	57.9	66.3	65.9	66	65.6	64.6	63.3
	(65.38,72.4)	(57,61.1)	(62.2,69.4)	(61.64,69)	(61.6,69.1)	(60.9,68.8)	(59.04,67.9)	(58.34,67.3)
15,000	80.7	77.9	80.4	80.4	81.9	82.4	80.5	76.8
	(78.68,82.7)	(77,79.1)	(78.2,82.2)	(78.2,82.3)	(80,83.6)	(80.4,84.3)	(78.4,82.3)	(70.44,80)
17,500	89.7	86.5	90	90.4	90.8	92.4	91.4	88.1
	(87.3,91.9)	(85.3,88.1)	(88,91.5)	(88.3,92.1)	(88.6,92.6)	(89.6,94.46)	(89,93.6)	(84.64,89.7)
20,000	96.8	95.9	97.7	97.7	97.7	98.3	98.1	96.4
	(94.8,98.8)	(94.9,98.2)	(95.3,98.6)	(95.3,98.6)	(95.3,98.6)	(96.2,99)	(95.9,98.9)	(94.5,97.8)

Gives market shares for alternative information structures, defined in Figure 14. Bootstrapped 90% confidence intervals reported below point estimate.

Table 32: Budget Balance: Alternative Information Designs

	Realized	Original(0)	Conservative(1)	Municipal(2)	Municipal(2B)	NGO(3)	NG
Average	-102	-24	-22	-177	-306	-18	
	(-150,-53)	(-62,18)	(-55,14)	(-224,-120)	(-383,-225)	(-43,14)	(2
10,000	-1334	-894	-808	-849	-776	-553	
	(-1477,-1189)	(-1000,-780)	(-910,-699)	(-965,-724)	(-884,-652)	(-629,-452)	(-5
15,000	116	109	100	-158	-242	94	
	(91,144)	(84,139)	(73,128)	(-211,-87)	(-325,-150)	(74,117)	(1
17,500	334	331	283	225	-132	137	
	(273,401)	(288,390)	(245,333)	(191,271)	(-227,-58)	(109,170)	(2
20,000	689	530	530	520	155	300	
	(308,1064)	(372,677)	(372,677)	(364,670)	(94,213)	(191,409)	(8

Gives market shares for alternative information structures, defined in figure 14. Bootstrapped 90% confidence
reported below point estimate.

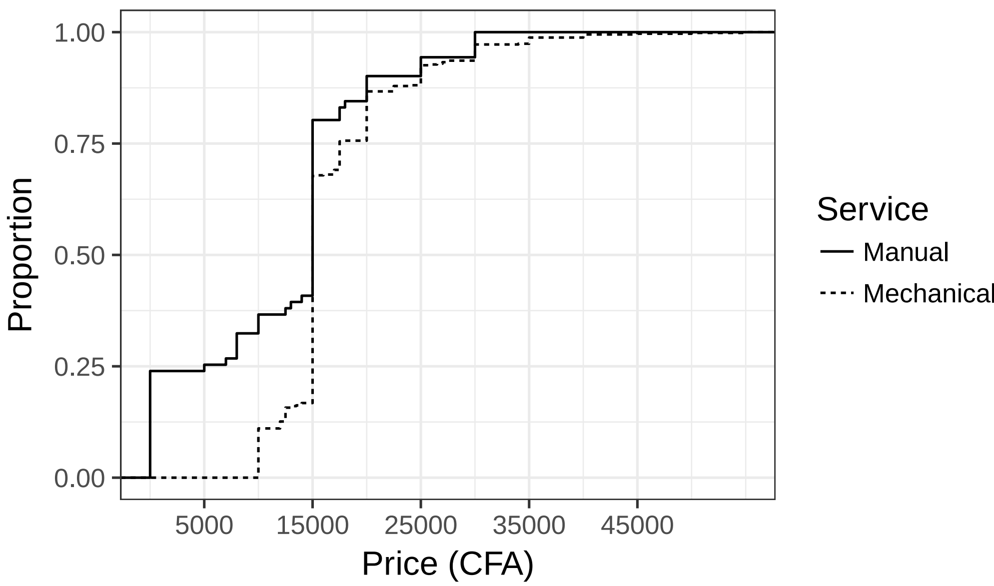
Table 33: Variable Importance of Design Variables (Gini coefficient)

	Original(0)	Conservative(1)	Municipal(2)	Conservative Municipal(2B)	NGO(3)	Conservative NGO(3B)
Average Months Between Desludgings	68	65.728				
Water Bill More Than 5,000 CFA	9.018	10.646	14.748	16.617		
House Type: Precarious	29.491	38.374			97.735	52.681
House Type: Concrete	23.124	29.911			81.786	13.392
House Type: Rooming House	4.947	5.382			12.559	2.759
Other Households in Compound	16.264	14.336			6.632	18.71
Own House	16.399	14.96	39.702			
Pit Meters From Road	19.208	16.077			5.229	57.635
More than 1 Trip Last Desludging	11.521	10.462				
Electricity Bill	151.488	174.222	236.082	189.703		
Household Size	25.631	22.908			11.336	
Number of Women in Household	21.412					
Respondent Finished High School	19.257					

Variable Importance averages the amount by which adding the variable to one of the decision trees in the random forest reduced misclassification as measured by the Gini coefficient at terminal nodes of the decision tree (see Section 6.2 or Appendix G).

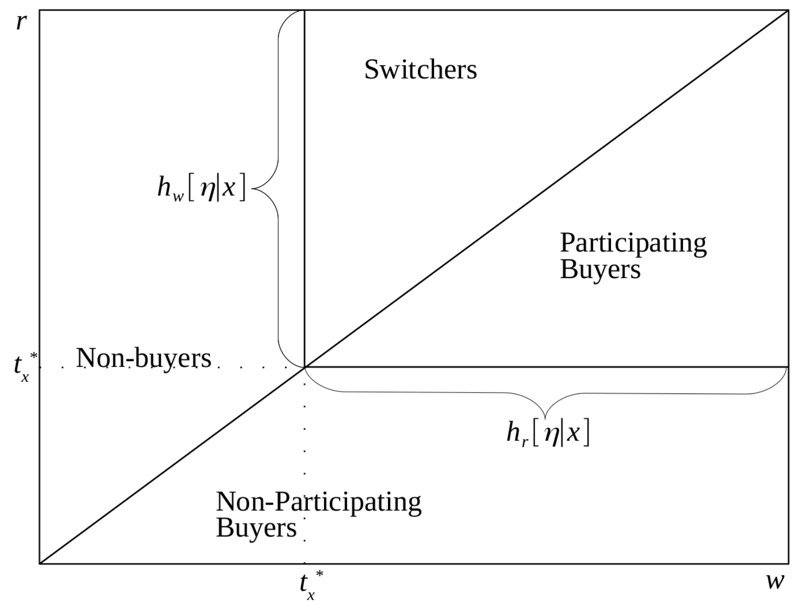
D Figures

Figure 1: Baseline Prices of Mechanical and Manual Services



Households that provide manual desludging services themselves pay nothing, accounting for the high value for Manual at a price of 0. The modal price in the search market for manual and mechanical services is 15,000 CFA, but the mechanical price distribution first-order stochastically dominates the manual price distribution on all of the support.

Figure 2: Taxonomy of Household Types, Theoretical



The switching function, $\sigma(\eta, x)$, is the hazard rate of being on the boundary between buying a mechanical service on the platform rather than the manual service, $h_w[\eta|x]$, divided by the sum of the hazard rates of being on either of the two boundaries and reporting η, $h_w[\eta|x] + h_r[\eta|x]$.

Figure 3: Taxonomy of Household Types, Empirical

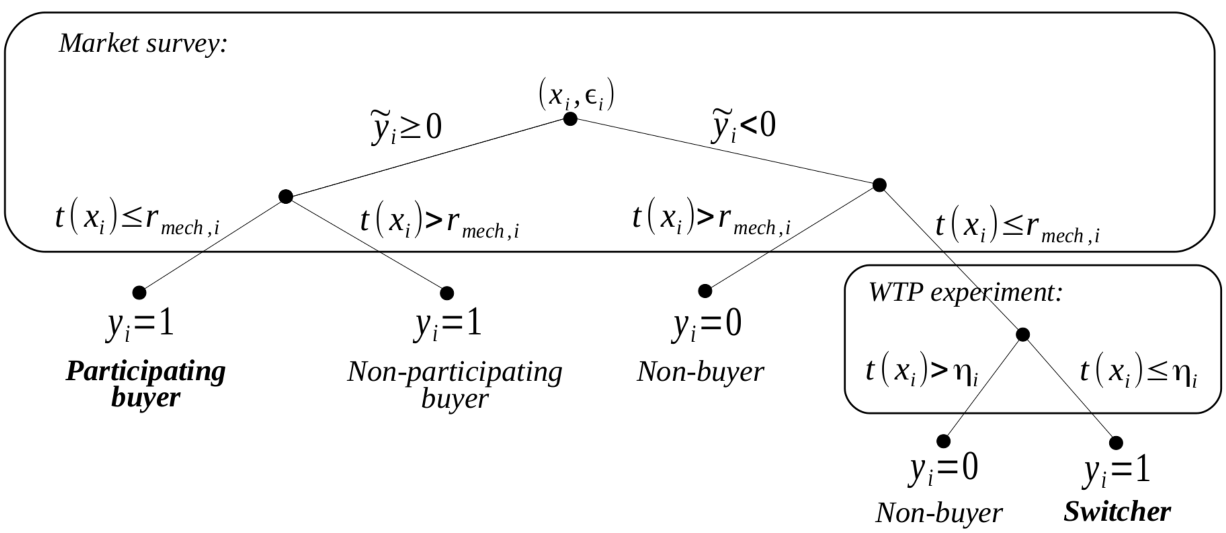


Figure 4: Offers

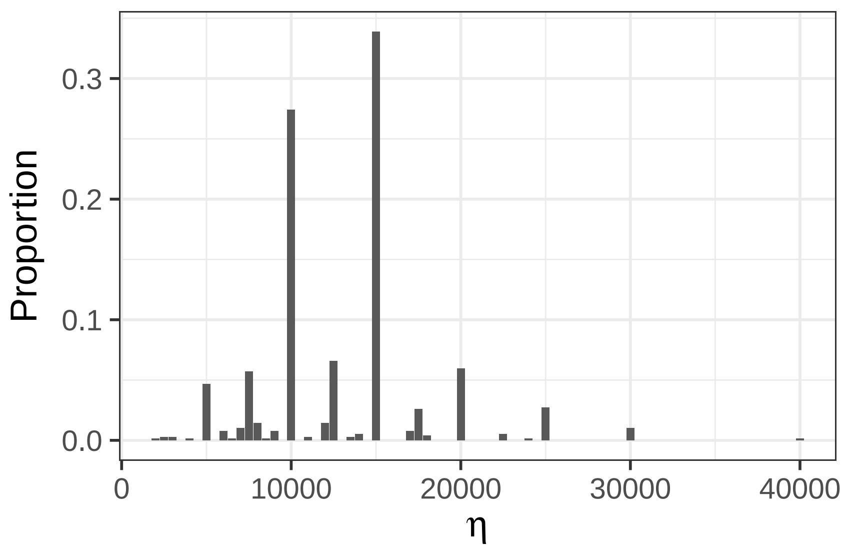
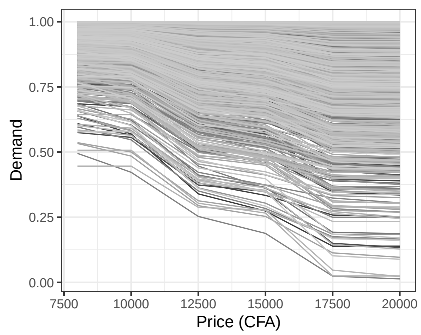
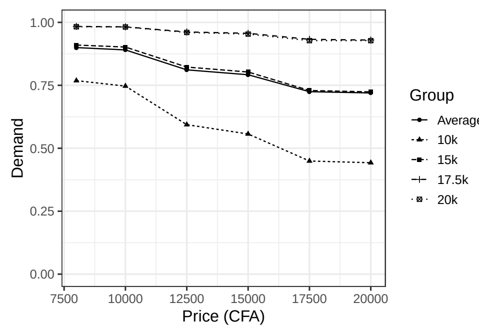


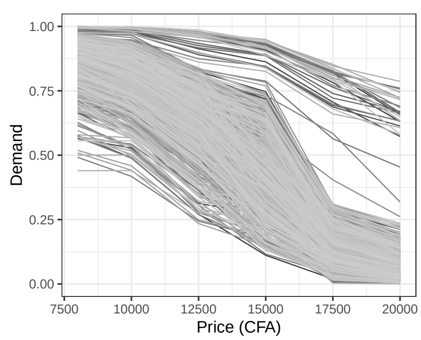
Figure 5: Demand



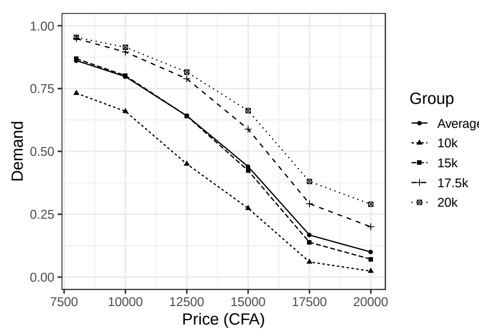
(a) Mechanical Demand by Household



(b) Average Mechanical Demand



(c) Platform Demand by Household



(d) Average Platform Demand

Figure 6: Supply-Side Auctions Average Clearing Prices by Round

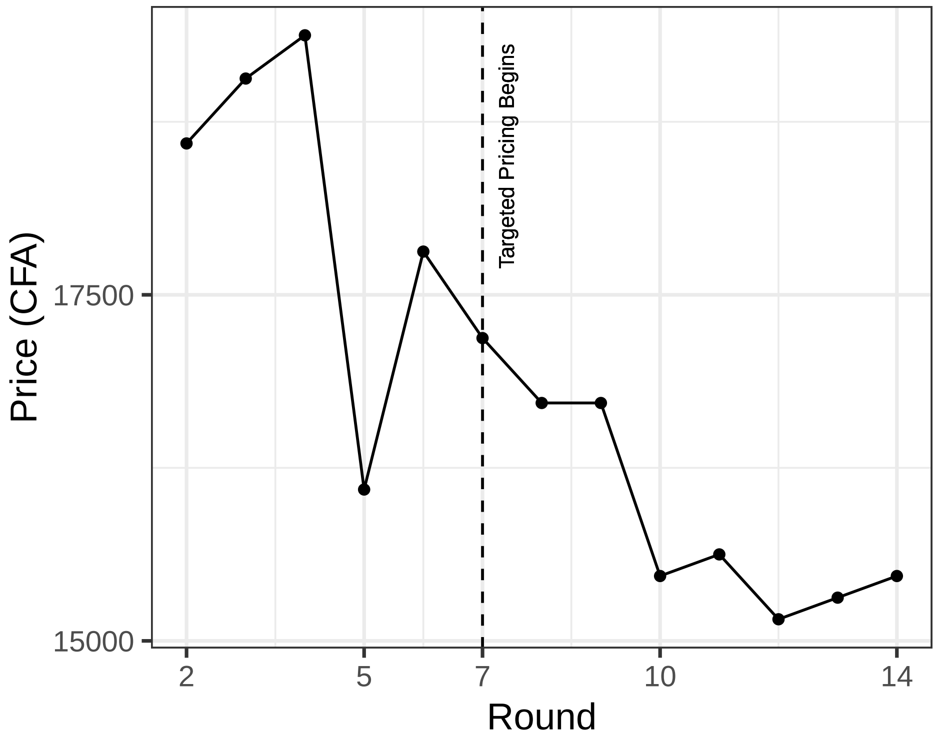
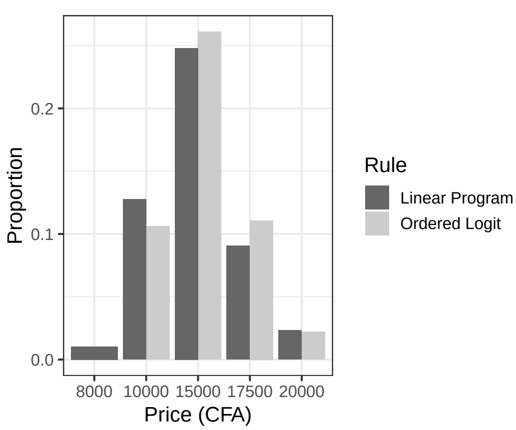
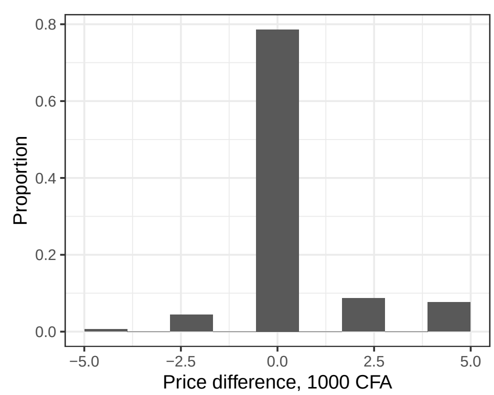


Figure 7: Pricing Rules

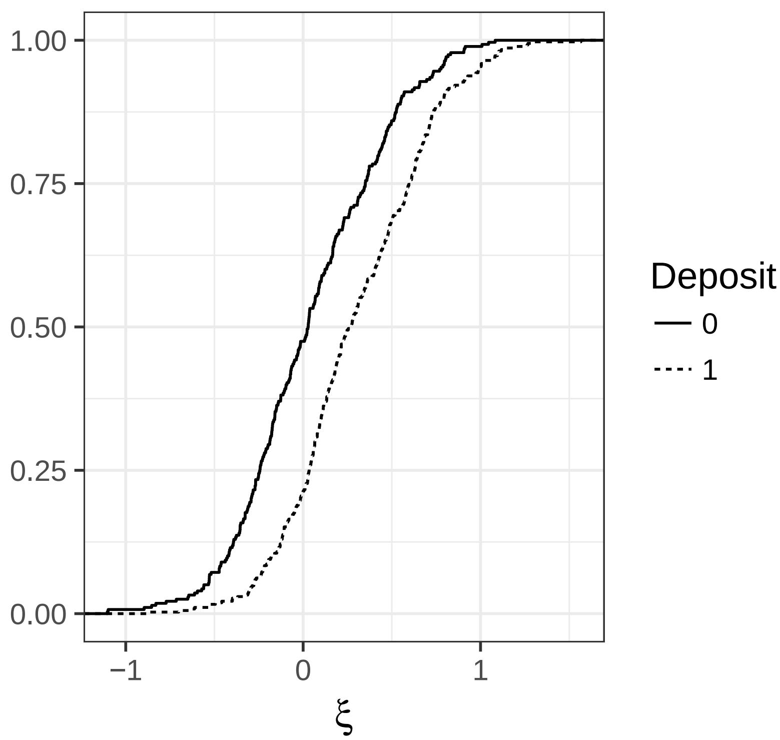


(a) Linear Programming and Ordered Logit Prices



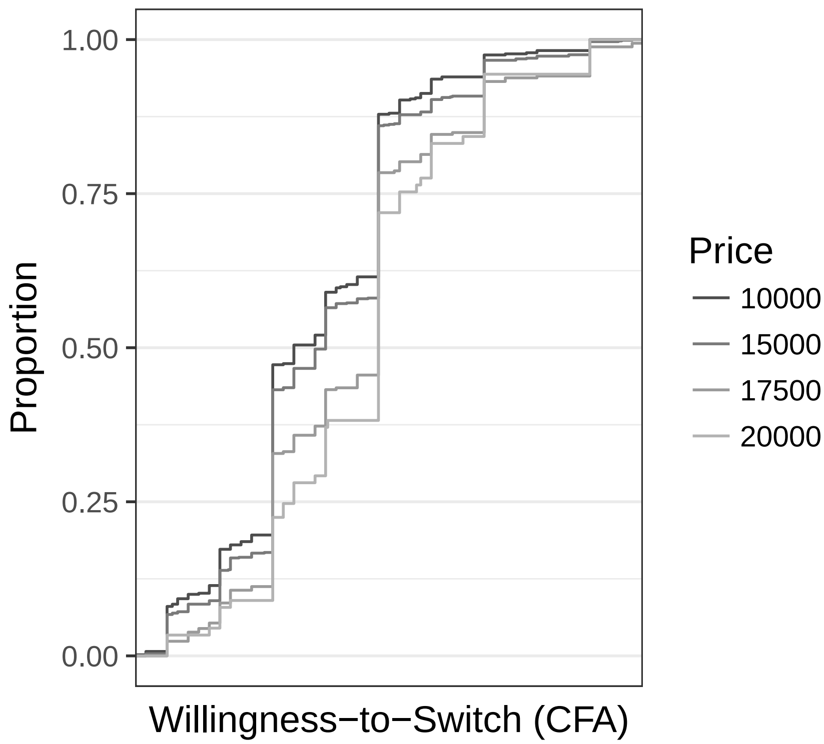
(b) Deviations of Ordered Logit from Linear Programming Price

Figure 8: ξ Conditional on Deposit Decision



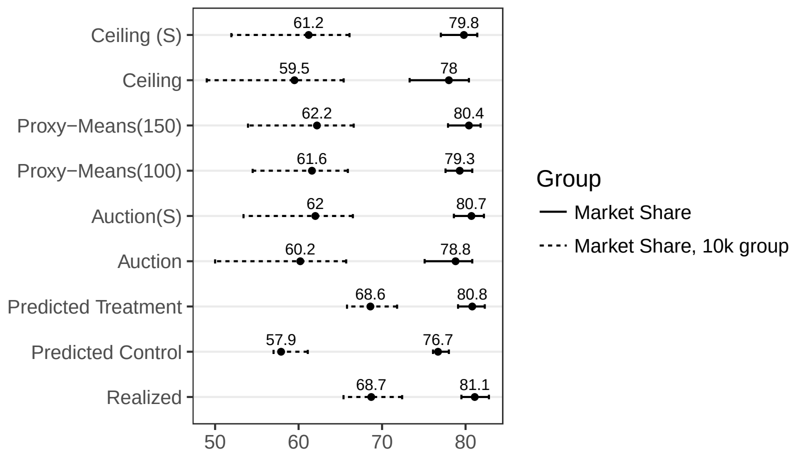
A Kolmogorov-Smirnoff test of equality of the two distributions is rejected at any conventional level of significance, with $D = .263$, corresponding to a p-value of 5.665×10^{-10}.

Figure 9: Willingness-to-switch values by price bin

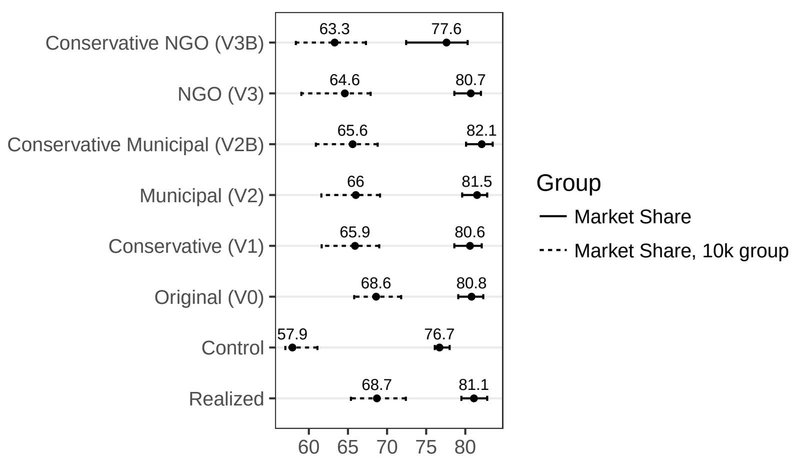


Note that while the supports of the willingness-to-switch values reported by the Demand Elicitation group have essentially the same support, the reports of the 20,000 CFA group first-order stochastically dominate those of the 17,500 CFA group, the 17,500 CFA group dominates the 15,000 CFA group, and the 15,000 CFA group dominates the 10,000 CFA group.

Figure 10: Counterfactual Market Shares

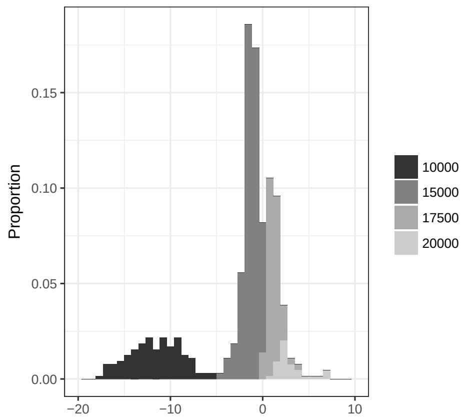


(a) Alternative Designs

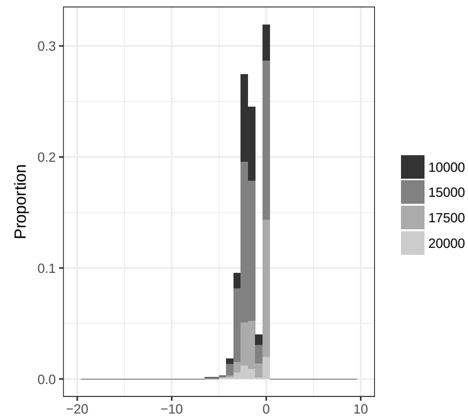


(b) Alternative Information

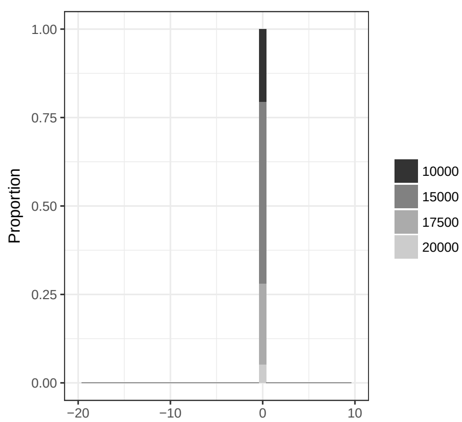
Figure 11: Alternative Designs: Subsidization Rates



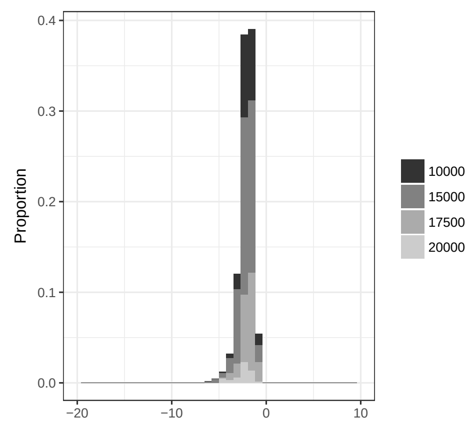
(a) Targeted Pricing



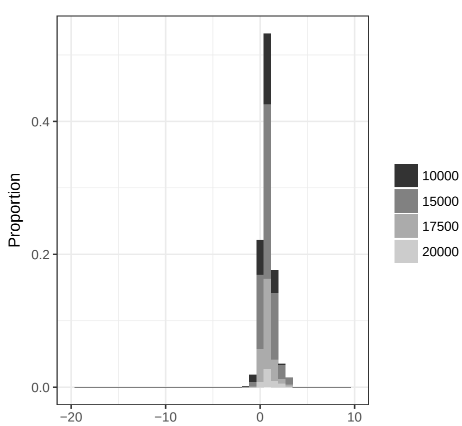
(b) Proxy-Means Testing



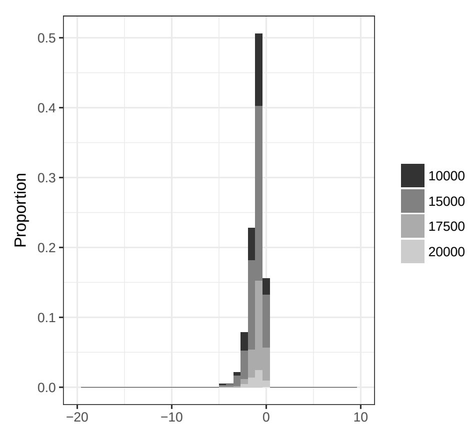
(c) Auctions



(d) Subsidized Auctions

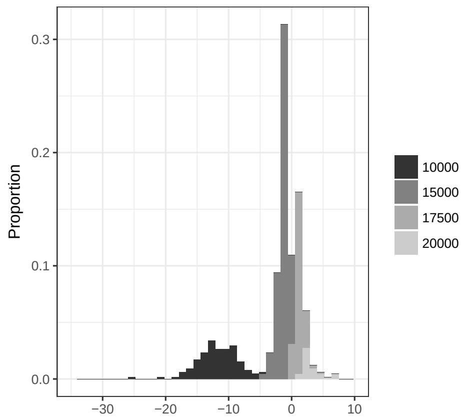


(e) Centralized Market

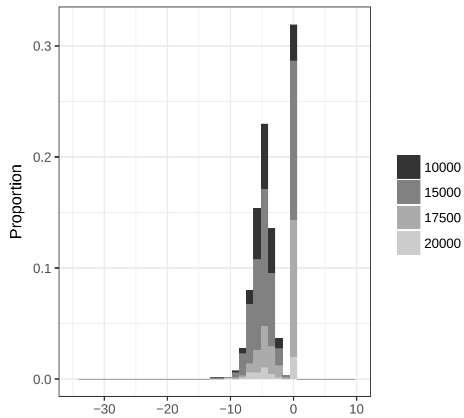


(f) Subsidized Centralized Market

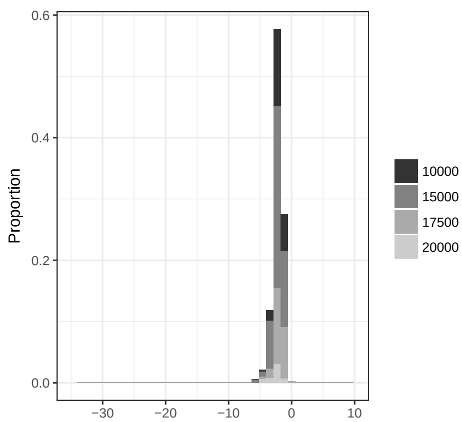
Figure 12: Design Variation: Subsidization Rates



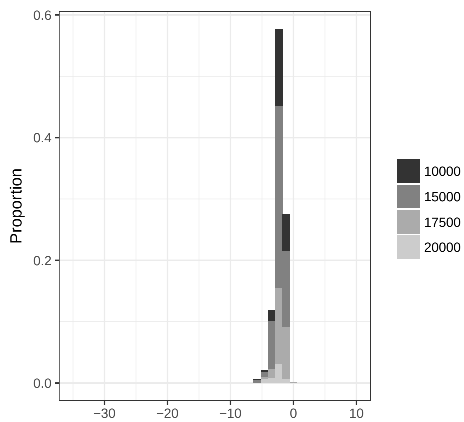
(a) Targeted Pricing



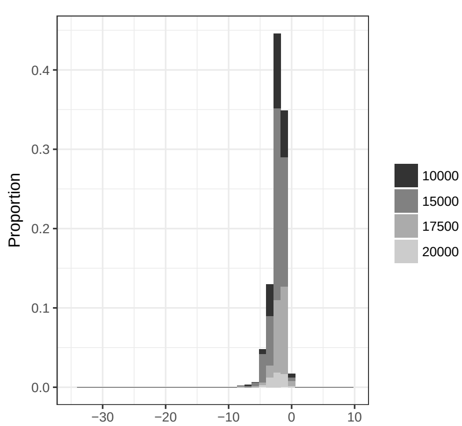
(b) Proxy-Means Testing



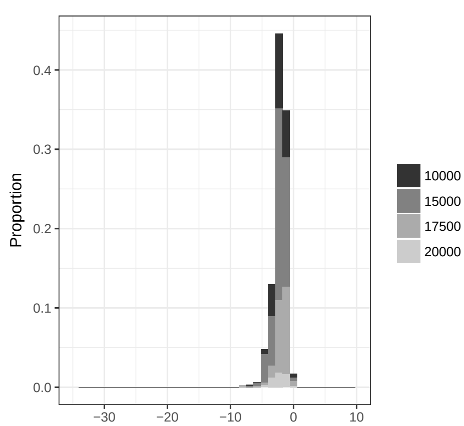
(c) Auctions



(d) Subsidized Auctions

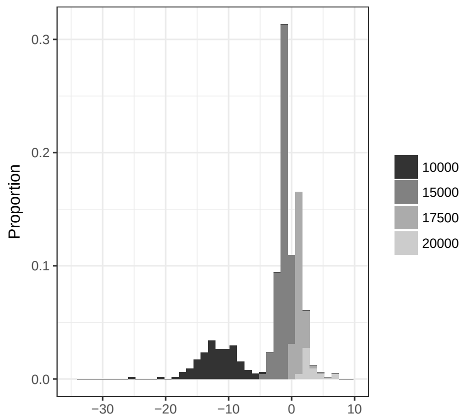


(e) Price Ceiling

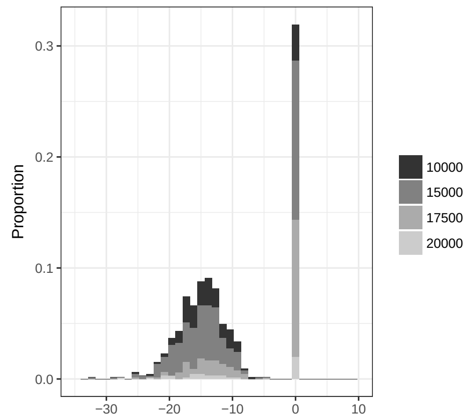


(f) Subsidized Price Ceiling

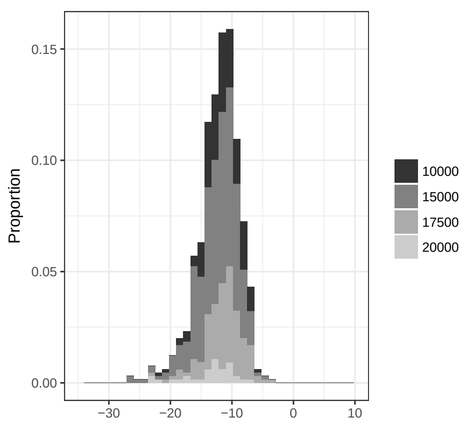
Figure 13: Design Variation, 10,000 CFA bin: Subsidization Rates



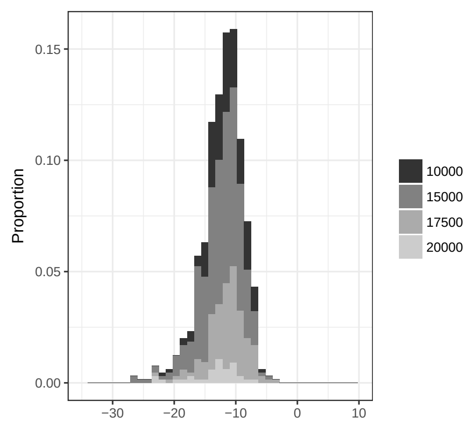
(a) Targeted Pricing



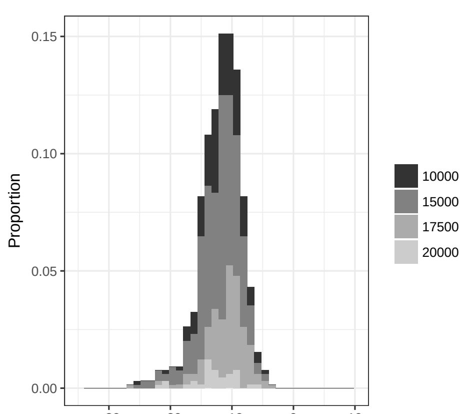
(b) Proxy-Means Testing



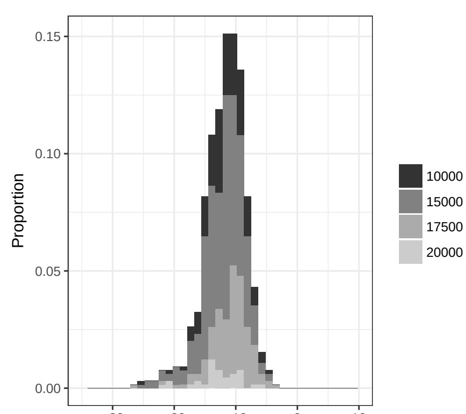
(c) Auctions



(d) Subsidized Auctions



(e) Price Ceiling



(f) Subsidized Price Ceiling

Figure 14: Information Structures

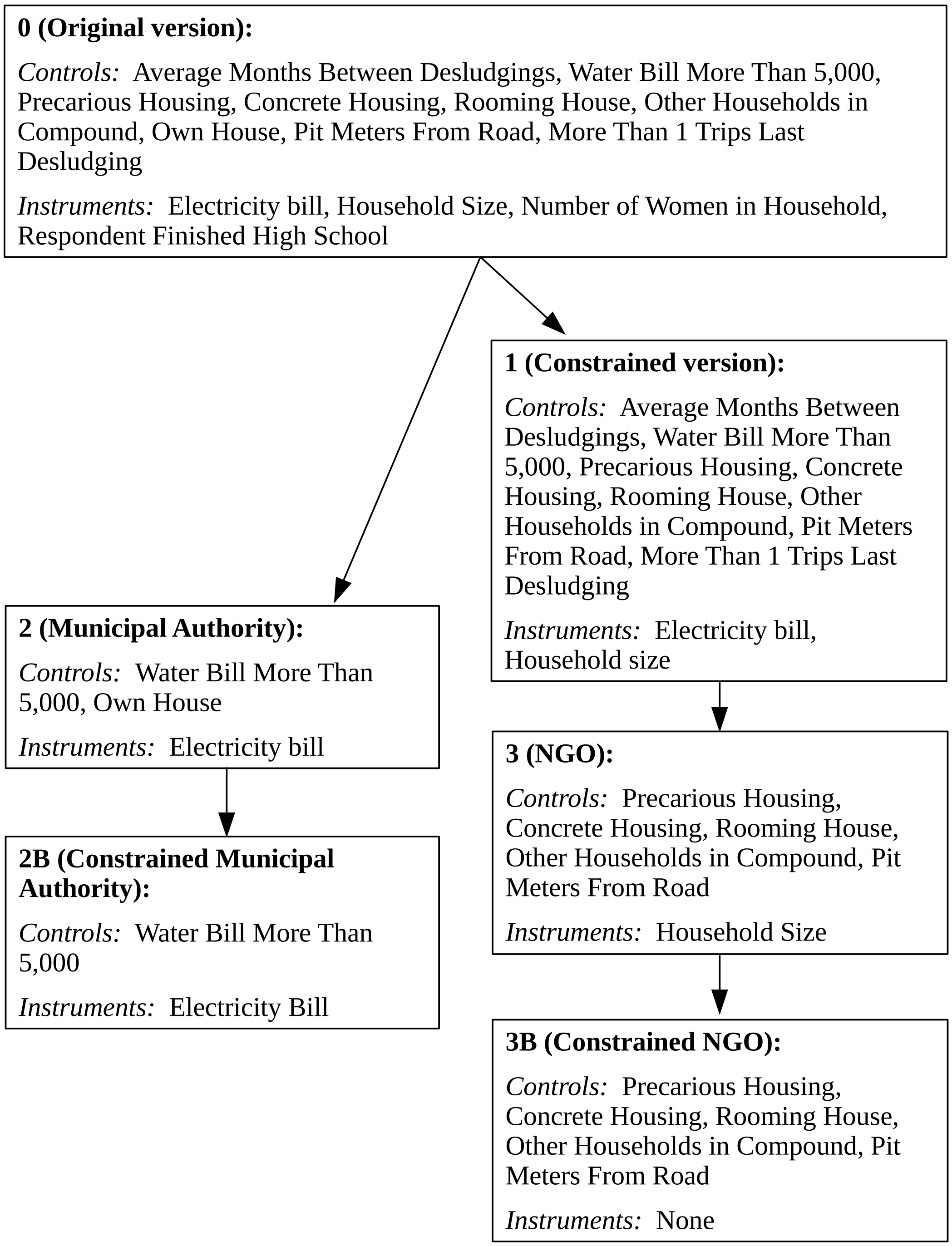
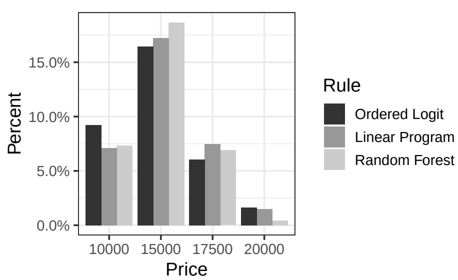
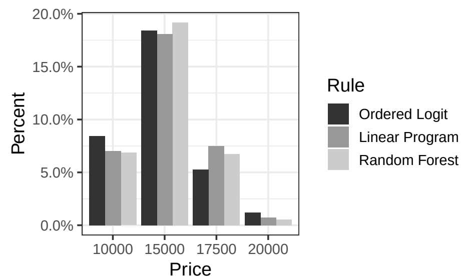


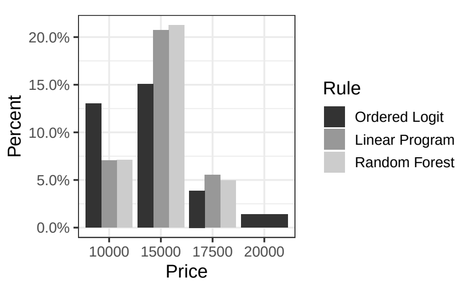
Figure 15: Alternative Information Design: Price Rules



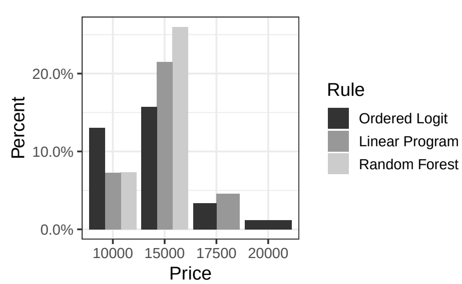
(a) Original (V0)



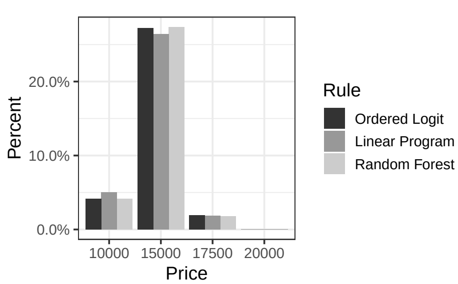
(b) Conservative (V1)



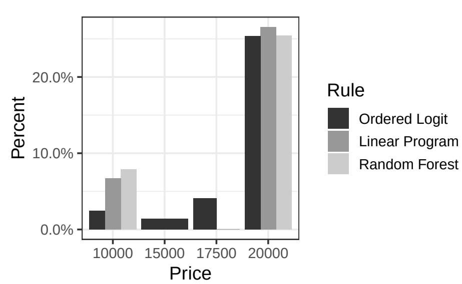
(c) Municipal Authority (V2)



(d) Conservative Municial Authority (V2B)

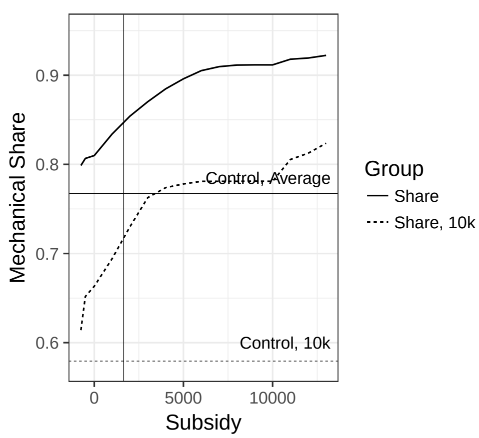


(e) NGO (V3)

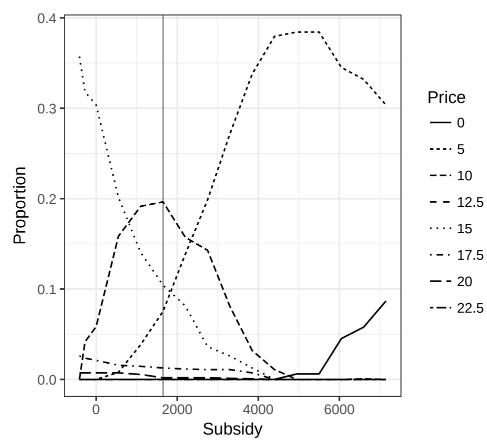


(f) Conservative NGO (V3B)

Figure 16: Sustainability Analysis



(a) Mechanical Market Share



(b) Price Offers

E Theory Appendix

Proof of Theorem 1:

Proof: We begin by providing a standard iff characterization of incentive compatibility in terms of the single-crossing property and the envelope theorem. We then solve the "relaxed problem" by dropping the monotonicity condition and investigating what sufficient conditions on primitives ensure that it fails to bind at the optimum.

Each household strategically chooses its report $\hat{\eta}$ to maximize

$$U(\hat{\eta}, \eta, x) = p(\hat{\eta}, x)(\eta - t(\hat{\eta}, x)), \tag{34}$$

and define the indirect utility function

$$V(\eta, x) = \max_{\hat{\eta}} p(\hat{\eta}, x)(\eta - t(\hat{\eta}, x)). \tag{35}$$

This characterization of the household's problem allows for a simple characterization of incentive compatibility:

Proposition 2 *A direct mechanism $\{p(\hat{\eta}, x), t(\hat{\eta}, x)\}$ is incentive compatible iff $\frac{\partial}{\partial \eta} V(\eta, x) = p(\eta, x)$ and $p(\hat{\eta}, x)$ is non-decreasing in $\hat{\eta}$.*

Proof: Assume the direct mechanism is incentive compatible. From the Milgrom-Segal envelope theorem (Milgrom and Segal (2002)), $V_\eta(\eta, x) = p(\eta, x)$. Taking two revealed-preference constraints

$$p(\eta, x)\eta - t(\eta, x) \geq p(\eta', x)\eta - t(\eta', x), \quad p(\eta', x)\eta' - t(\eta', x) \geq p(\eta, x)\eta' - t(\eta, x)$$

and re-arranging them yields

$$(\eta - \eta')(p(\eta, x) - p(\eta', x)) \geq 0,$$

so that if $\eta > \eta'$, $p(\eta, x) \geq p(\eta', x)$, and $p(\eta, x)$ is non-decreasing in η.

Now assume $V_\eta(\eta, x) = p(\eta, x)$ and $p(\eta, x)$ is non-decreasing in η. Then

$$\begin{aligned} U(\eta,\eta,x) - U(\eta',\eta,x) &= U(\eta,\eta,x) - U(\eta',\eta',x) + U(\eta',\eta',x) - U(\eta',\eta,x) \\ &= \int_{\eta'}^{\eta} p(z,x)dz + \int_{\eta}^{\eta'} p(\eta',x)dz \\ &= \int_{\eta'}^{\eta} p(z,x) - p(\eta',x)dz, \end{aligned}$$

where the second line follows from $V_\eta(\eta, x) = p(\eta, x)$. Now, since $p(\eta, x)$ is non-decreasing, the integrand on the third line is positive whenever $\eta > \eta'$, and negative whenever $\eta < \eta'$, so that the third line is always weakly positive. Therefore, the mechanism is incentive compatible. ■

We drop the constraint that $p(\hat{\eta}, x)$ be non-decreasing in $\hat{\eta}$ and solve the problem only requiring that $V_\eta(\eta, x) = p(\eta, x)$, and then determine sufficient conditions for $p(\hat{\eta}, x)$ to be non-decreasing. The logic of the relaxed solution is that the monotonicity condition is mathematically difficult to handle (e.g. Mussa and Rosen (1978), Myerson (1981), and Rochet (1987)) and is often satisfied at the optimum if a mild regularity condition is imposed.

If $V_\eta(\eta, x) = p(\eta, x)$, then its expected payoff must satisfy $V(\eta, x) = \int_{\eta_x^*}^{\eta} p(z, x)dz$ where η_x^* is the lowest type who trades with positive probability; note that the worst-off type $\underline{w}$ is quoted a price of $\underline{w}$ with probability zero in the market and there aren't enough subsidies to cover the whole market, so that $V(\underline{w}, x) = 0$. In any incentive compatible mechanism, this implies $p(\eta, x)(\eta - t(\eta, x)) = \int_{\eta_x^*}^{\eta} p(z, x)dz$, and a household of type (η, x) expects to pay

$$p(\eta,x)t(\eta,x) = p(\eta,x)\eta - \int_{\eta_x^*}^{\eta} p(z,x)dz. \tag{36}$$

Taking the expectation with respect to η and integrating by parts then yields

$$\int_{w_x^*}^{\overline{w}} p(\eta,x)t(\eta,x)dF_\eta[\eta|x] = \int_{w_x^*}^{\overline{w}} p(\eta,x)\eta - \int_{\eta_x^*}^{\eta} p(z,x)dzdF_\eta[\eta|x] = \int_{w_x^*}^{\overline{w}} p(\eta,x)\left\{\eta - \frac{1-F_\eta[\eta|x]}{f_\eta[\eta|x]}\right\} dF_\eta$$

This expresses the expected revenue from an x-type of household entirely in terms of the probability of trade, $p(\eta, x)$. Taking the expectation over x then yields expected total revenue. The preceding arguments establish equation (4).

Note that the distribution of η is

$$F_\eta[\eta|x] = (1 - F_w[\eta|x])F_r[\eta|x] + (1 - F_r[\eta|x])F_w[\eta|x] + F_w[\eta|x]F_r[\eta|x],$$

with density

$$f_\eta[\eta|x] = (1 - F_w[\eta|x])f_r[\eta|x] + (1 - F_r[\eta|x])f_w[\eta|x],$$

and virtual valuation

$$\psi_\eta[\eta|x] = \eta - \frac{1 - F_\eta[\eta|x]}{f_\eta[\eta|x]} = \eta - \frac{1}{\dfrac{f_w[\eta|x]}{1 - F_w[\eta|x]} + \dfrac{f_r[\eta|x]}{1 - F_r[\eta|x]}} = \eta - \frac{1}{h_w[\eta|x] + h_r[\eta|x]}.$$

So if the standard regularity condition that $1 - F_w[w|x]$ and $1 - F_r[r|x]$ are each log-concave, the associated hazard rates will be increasing, and $\psi_\eta[\eta|x]$ will be increasing in η.

Dropping the monotonicity condition that $p(\eta, x)$ be non-decreasing in η, the simplified problem is to maximize quantity

$$\mathbb{E}_{(w,r,x)}\left[p(\min\{w, r\}, x)b_x + (1 - p(\min\{w, r\}, x))\mathbb{I}\{w \geq r\}b_x\right]$$

subject to

$$\mathbb{E}_{(\eta,x)}\left[p(\eta, x)(\psi_\eta[\eta|x] - c_x)\right] + s \geq 0.$$

Consider the term $\mathbb{E}_{(w,r,x)}\left[(1 - p(\min\{w, r\}, x))\mathbb{I}\{w \geq r\}\right]$. Since $\eta = \min\{w, r\}$, the indicator function takes the value 1 only when $\eta = \min\{w, r\} = r$. Therefore, this term equals

$$\begin{aligned}
\int_r \int_w (1 - p(\min\{w, r\}, x))\mathbb{I}\{w \geq r\} dF_w[w|x] dF_r[r|x] &= \int_{\eta=\underline{w}}^{\eta=\overline{w}} \int_{w=\eta}^{\overline{w}} (1 - p(\eta, x)) dF_w[w|x] dF_r[\eta|x] \\
&= \int_{\eta=\underline{w}}^{\eta=\overline{w}} (1 - F_w[\eta|x])(1 - p(\eta, x)) f_r[\eta|x] d\eta \\
&= \int_{\eta=\underline{w}}^{\eta=\overline{w}} \frac{(1 - F_w[\eta|x]) f_r[\eta|x]}{f_\eta[\eta|x]} (1 - p(\eta, x)) dF \\
&= \mathbb{E}_{(\eta,x)} \left[\frac{(1 - F_w[\eta|x]) f_r[\eta|x]}{f_\eta[\eta|x]} (1 - p(\eta, x)) \right]
\end{aligned}$$

and note that

$$\frac{(1 - F_w[\eta|x]) f_r[\eta|x]}{f_\eta[\eta|x]} = \frac{(1 - F_w[\eta|x]) f_r[\eta|x]}{(1 - F_w[\eta|x]) f_r[\eta|x] + (1 - F_r[\eta|x]) f_w[\eta|x]} = \frac{h_r[\eta|x]}{h_w[\eta|x] + h_r[\eta|x]}.$$

The Lagrangian then is

$$\mathcal{L}(p, \lambda) = \mathbb{E}_{(\eta,x)} \left[p(\eta, x) b_x + (1 - p(\eta, x)) \frac{h_r[\eta|x]}{h_w[\eta|x] + h_r[\eta|x]} b_x \right] + \lambda \left(\mathbb{E}_{(\eta,x)} \left[p(\eta, x)(\psi_\eta[\eta|x] - c_x) \right] + s \right)$$

or

$$\mathcal{L}(p, \lambda) = \mathbb{E}_{(\eta,x)} \left[p(\eta, x) \frac{h_w[\eta|x]}{h_w[\eta|x] + h_r[\eta|x]} b_x + \frac{h_r[\eta|x]}{h_w[\eta|x] + h_r[\eta|x]} b_x \right] + \lambda \left(\mathbb{E}_{(\eta,x)} \left[p(\eta, x)(\psi_\eta[\eta|x] - c_x) \right] + \right.$$

expressing the problem entirely in terms of η, which is equation (6). The objective is linear in $p(\eta, x)$, and collecting terms multiplied by $p(\eta, x)$ yields

$$\beta(\eta, x, \lambda) = \sigma(\eta, x) b_x + \lambda (\psi_\eta[\eta|x] - c_x).$$

Now if $\beta(\eta, x, \lambda)$ has the single-crossing property in η for all (x, λ) and the crossing point is increasing in η, the monotonicity condition will be satisfied. Sufficient conditions for this to hold are that $\psi_\eta[\eta|x]$ and $\sigma(\eta, x)$ both be non-decreasing in η. This means that households with higher η types are more profitable to serve on the margin, and whenever a household reports a higher type, the platform infers it is more likely to switch *conditional on* x — i.e., that households of similar socio-economic observables who report higher η are more likely to face high prices in the market and be a switcher, conditioning on x. If either of these sufficient conditions is violated, $\beta(\eta, x, \lambda)$ might still be non-decreasing in η or satisfy the single-crossing property in η. If $\beta(\eta, x, \lambda)$ exhibits violations of the single-crossing property, the monotonicity condition binds, and an optimal control approach is required; the optimal mechanism will then involve a deterministic contract for high reports of η, and a series of contracts with lower probability of service and lower prices.

To characterize λ^*, note that the optimal allocation is a cut-off rule for every x, where the cutoff is given by

$$\sigma(\eta_x^*(\lambda), x)b_x + \lambda(\psi_\eta[\eta_x^*(\lambda))|x] - c_x) = 0.$$

By the implicit function theorem, $\eta_x^*(\lambda)$ is a continuous function, with derivative

$$\frac{d}{d\lambda}\eta_x^*(\lambda) = \frac{-(\psi_\eta[\eta_x^*(\lambda)|x] - c_x)}{\frac{d}{d\eta}\sigma(\eta_x^*(\lambda), x)b_x + \lambda\frac{d}{d\lambda}\psi_\eta[\eta_x^*(\lambda)|x]}$$

Now, since $\psi_\eta[\eta|x]$ is increasing in η and $w_x^*(\lambda)$ is weakly less than the monopoly cutoff where $\psi_\eta[\eta_x^m|x] - c_x = 0$, the numerator is positive, and monotonicity assumptions ensure the denominator is positive. Therefore, $w_x^*(\lambda)$ is a continuous and non-decreasing function.

The budget is then given by $\phi(\lambda) = \mathbb{E}_x\left[(\eta_x^*(\lambda) - c_x)(1 - F_\eta(\eta_x^*(\lambda))\right] + s$ with derivative

$$\begin{aligned}\phi'(\lambda) = \mathbb{E}_x &\left[\{(1 - F_\eta[\eta_x^*(\lambda)|x]) - f_\eta[\eta_x^*(\lambda)|x])(\eta_x^*(\lambda) - c_x)\}\frac{d\eta_x^*(\lambda)}{d\lambda}\right] \\ &= \mathbb{E}_x\left[-(\psi_\eta[\eta_x^*(\lambda)|x] - c_x)f_\eta[\eta_x^*(\lambda)|x]\frac{d\eta_x^*(\lambda)}{d\lambda}\right] \geq 0,\end{aligned}$$

because, again, $\psi_\eta[\eta|x]$ is non-decreasing and $\eta_x^*(\lambda)$ is weakly less than the monopoly solution, where $\psi_\eta[\eta_x^m|x] - c_x = 0$. Therefore, the budget is negative at $\lambda = 0$ since $\underline{w} + s < c_x$ for all x, non-decreasing, continuous, and strictly positive as $\lambda \to \infty$. This in turn implies there exists a finite $\bar{\lambda}$ for which it is strictly positive, allowing us to restrict attention to a compact interval $[0, \bar{\lambda}]$.

Therefore, by the intermediate value theorem, there exists a λ^* that balances the budget and characterizes the optimal mechanism. ■

F Attrition

One may be concerned that we could have differential attrition between the control and treatment group which could create bias in the estimation of our results. We estimated the effect of being assigned to a treatment on attrition. Results are shown in Table 34.

Table 34: Attrition

	(1) attrited	(2) attrited
Targeted Price group	-0.0101	
	(0.01)	
10k group		0.0526***
		(0.01)
15k group		0.0578***
		(0.01)
17k group		0.102***
		(0.02)
20k group		0.185***
		(0.03)
10k group × TP Group		-0.00971
		(0.02)
15k group × TP Group		0.0107
		(0.01)
17.5k group × TP Group		-0.0336
		(0.02)
20k group × TP Group		-0.0566
		(0.04)
Constant	0.0740***	
	(0.01)	

Column 1 shows that we find no evidence of differential attrition between the control and treatment group. Overall, our rate of attrition was 7.4%.

Treatment households were 1% less likely than control households to attrit, but this is not significant at conventional levels of significance.

In column 2, we test whether there is differential attrition between the control group and the treatment group across the different price bins. We suppress the constant so that we can also estimate the level of attrition in each of the price bins. We find that households who fit the qualification requirements for the 10,000 or 15,000 CFA prices (in both the treatment and control groups) had lower than average attrition of 5.3-5.8 percent, but that there is no differential attrition between the treatment and control group among these price bins.

Attrition is higher among the households who fit the qualification requirements for either 17,500 or 20,000 (10-19%), but again, there is no differential attrition between the treatment group and the control group in these price bins. Higher attrition among wealthier households who would receive higher prices is less of a concern for our estimates, as these households are also more likely to purchase mechanical desludgings on their own (in either the treatment or the control groups), and our main results are focused on the decisions of the households in the lowest price bins.

G Random Forests (Not for publication)

This discussion is adapted from (Hastie et al., 2017). The basic intuition for using algorithms like random forest instead of regression methods like OLS is that regression tends to condense the data around the mean, so that the extreme bins (say, 10,000 CFA and 20,000 CFA) get little to no coverage. Methods like ordered logit and random forest can do a better job at assigning observations to such bins.

The random forest algorithm for classification takes data $\{y_i, x_i\}_{i=1}^N$ where y_i takes values in a finite outcome set $Q = \{q_1, ..., q_K\}$ and x_i is a vector in $\mathbb{R}^N$, and provides a rule to impute an outcome $y_j \in Q$ to any vector x_j. The algorithm is a "forest" because it fits a large number of decision trees to the training data, and the trees are "random" in order to reduce the sensitivity of the resulting rule compared to a single tree. To make this more precise, we'll start with the analysis of a classification tree.

To be precise, this paragraph contains some graph theory terms, but the intuition can be grasped by referring to Figure 17. A graph T is a set of vertices V and edges E. A vertex v is *connected* to a vertex v' if the edge from v to v' is in the set E. A graph is a tree T if there is some path from every vertex to every other vertex (it is *connected*) but the path from any vertex v to another vertex v' is unique (it is *acyclic*). A vertex or node n is *terminal* if it has exactly one link: i.e., it is at the bottom of the tree.A graph is a *decision tree* if every non-terminal vertex v has a function $f_v(z) = v'$ selecting a successor vertex v'. A decision tree is illustrated in Figure 17, giving a simple decision tree for the assignment of households to prices based on the original information structure[57].

To build a decision tree, let S be the node size: the total number of vertices allowed. Since the outcomes $\{y_i\}_{i=1}^N$ are categorical, the algorithm does not use squared error, as in regression. Instead, for terminal node n, let

$$p_{nk} = \frac{\sum_{z_i \to n} \mathbb{I}\{y_i = q_k\}}{|z_i \to n|},$$

where $z_i \to n$ is the subset of data mapped to terminal node n. In words, p_{nk} for a terminal node n at the bottom of the tree is the fraction of observations mapped to outcome q_k. If all of the observations mapped to n take the value q_k, it will be one for k, and otherwise zero; generally it will be a fraction. Then it makes sense to map any input vector z assigned to n, $z \to n$, to the value

$$k(n) = \operatorname*{argmax}_{k \in \{1,2,\ldots,K\}} p_{nk},$$

the outcome that is most likely based on the data mapped to the terminal node n.

Intuitively, the worst case when $|Q| = 2$ is when p_{nk} is near a half: being

[57]This figure also illustrates why so few households are placed in the 20,000 CFA bin: the algorithm first splits on electricity bill, placing wealthy households in the 17,500 CFA bin, rather than the 20,000 CFA bin.

in this node reveals little about what value z should be mapped to, but it is decisively mapped to one value or the other. This motivates using a measure of terminal node impurity based on the Gini coefficient:

$$G_n(T) = \sum_{k=1}^{K} \sum_{k' \neq k} p_{nk} p_{nk'} = \sum_{k=1}^{K} p_{nk}(1 - p_{nk}).$$

This preserves the property that the worst case is when $p_{nk} = 1/2$, maximizing the indecisiveness of the decision. The idea is that if $G_n(T)$ is close to zero, then there is little "disagreement" about what value the terminal node n should be assigned to, $k(n)$, while if $G_n(T)$ is large, there is a lot of heterogeneity in the outcomes assigned to terminal node n. The objective function is then specified as

$$C_\lambda(T) = \sum_{m \in T} |z \to m| G_n(T) + \lambda |T|.$$

This trades off the benefit of a more complex model that better fits the data with a linear cost function. The parameter λ is generally found by cross-validation, similar to penalized regression methods like the LASSO: adding more nodes incurs a cost $\lambda|T|$, and λ is selected to minimize prediction error, so the data is split into many folds and λ is chosen to minimize expected prediction error.

Solving $\operatorname{argmax}_T C_\lambda(T)$ is not practical because trees grow exponentially with the input data. Instead, the decision tree is constructed through a greedy algorithm that focuses on reducing terminal node impurity G_n until the "budget constraint" on the number of terminal nodes available is reached:

i. For each terminal node n, and each variable $z_{j\ell}$, an ℓ-split s is a value that partitions the data in n into two new nodes,

$$n_1(s) = \{z_i \in n : z_{j\ell} \leq s\} \text{ and } n_2(s) = \{z_i \in n : z_{j\ell} > s\}.$$

Pick the ℓ-split that minimizes the sum of Gini coefficients in the new terminal nodes n_1 and n_2:

$$R^*_{n\ell} = \min_s \sum_{z_i \to n_1(s)} G_{n1} + \sum_{z_i \to n_2(s)} G_{n2}.$$

ii. Add the split to the tree T to get a new tree T' that yields the lowest value of $C_\lambda(T')$.

iii. If $|T'| = S$, so the tree has reached the budget constraint size S, stop; otherwise repeat steps i–ii.

Like all greedy algorithms, this is not guaranteed to yield the global maximum, and there are a number of potential objectives, including the algorithm to binary splits. A more subtle problem with a simple decision tree is that it is very sensitive to the data and the splits. In order to reduce this sensitivity, one can construct many trees and average over them to get a more robust decision rule. The cost is that it can no longer be presented as in Figure 17, because the outcome it prescribes is the majority vote over all of the individual trees (one could visualize each of the trees individually, but there are often hundreds of them).

Let B be a number of bootstrap samples of the data with replacement. The random forest algorithm is

i. For $b = 1$ to B,

   (a) Draw a bootstrap sample Z_b

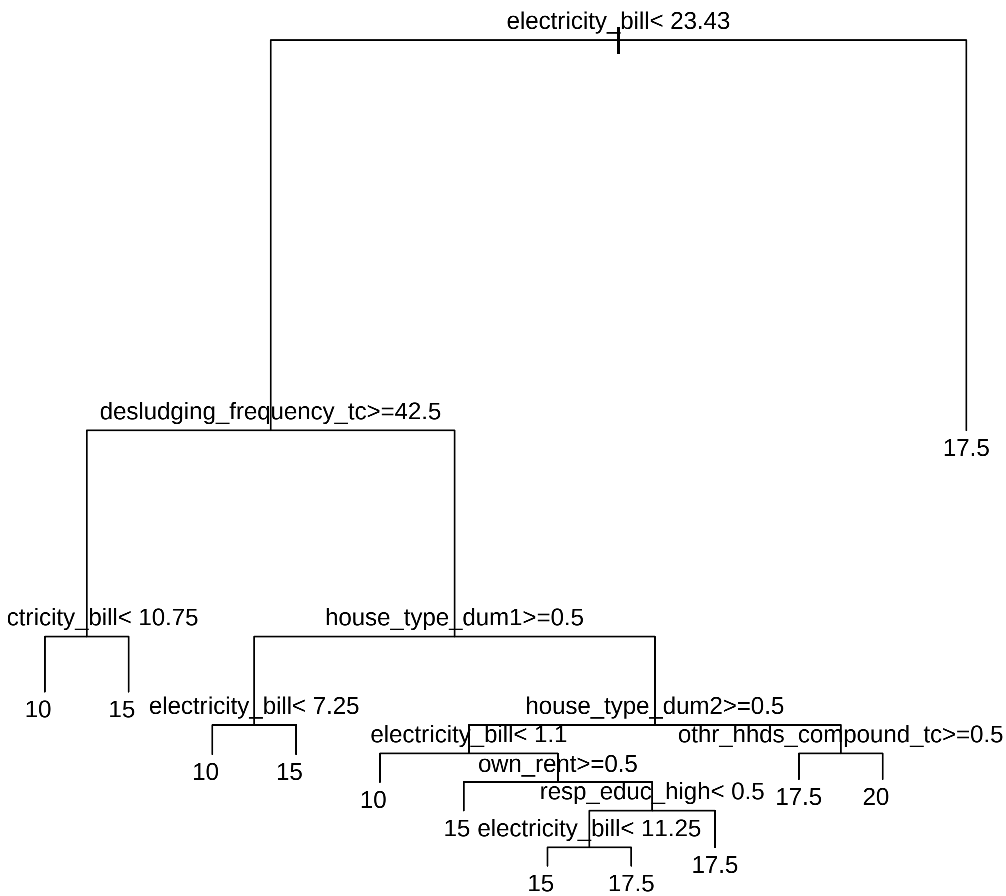
   (b) Draw a random subset $\lfloor \sqrt{L} \rfloor$ of variables where L is the total number of covariates, and grow a decision tree T_b from the bootstrap sample Z_b, but only using the randomly drawn variables.

ii. For covariates z_j, let π_k be the proportion of trees $\{T_b\}_{b=1}^B$ assigning z_j to q_k. Then z_j is assigned to $\text{argmax}_{k=1,..,K}\, \pi_k$.

Thus, the "random" in the random forest is that each tree is built using only a subset of the covariates available. The idea is that some variables with strong explanatory power will "crowd out" information from other variables, leading to trees that are very homogeneous across B draws of data. Whenever such a variable is drawn here, it will of course play a big role in determining the resulting tree, but when it is not selected, other variables might provide additional information that would otherwise be lost in the binary split structure. To get a prediction for data z_j, each of the B trees votes for one of the outcomes $\{q_1, ..., q_K\}$, and the modal q_j is selected.

How important or informative are individual variables? For a given covariate ℓ, one measure is to do to the following: sum the reductions in the Gini index in each of the trees in which ℓ appears, and divide by the number of trees in which ℓ appears; if a variable never appears, assign it an importance of zero. This delivers a measure of how much a given variable ℓ contributes to reducing mis-classification on average. Since the variables are drawn uniformly at random, if B is large, this should provide an estimate of the expected decrease in mis-classification that ℓ contributes to the random forest. We provide this for different information structures in 33.

Figure 17: Basic Decision Tree



H Other Tables (Not for publication)

Table 35: Counterfactual Subsidy Cost: Alternative Market Designs

	Realized	$\widehat{\text{Treatment}}$	Auction	Auction (S)	Proxy-Means(100)	Proxy-Means(150)	Ceiling	Ceiling(S)
Average	334	334	0	339	230	0	0	311
	(305,363)	(305,362)	(0,0)	(309,372)	(208,251)	(0,0)	(0,0)	(283,341)
10,000	496	512	0	328	275	0	0	305
	(445,550)	(462,566)	(0,0)	(263,410)	(216,343)	(0,0)	(0,0)	(235,394)
15,000	347	330	0	352	252	0	0	324
	(306,386)	(299,361)	(0,0)	(320,381)	(228,275)	(0,0)	(0,0)	(291,356)
17,500	190	211	0	308	140	0	0	279
	(153,231)	(184,245)	(0,0)	(241,385)	(106,171)	(0,0)	(0,0)	(228,338)
20,000	194	201	0	385	226	0	0	351
	(87,300)	(149,250)	(0,0)	(236,527)	(129,309)	(0,0)	(0,0)	(226,471)

Realized is the experimental value, $\widehat{\text{Treatment}}$ is the model predicted value, Auction is the predicted outcome when households are given the procurement auction prices, Proxy-Means(100) and Proxy-Means(150) refer to predicted outcomes using proxy-means testing at 100% and 150% of the poverty line, Cieling is a policy that constrains prices to be the average price in the search market, and (S) denotes an additional subsidy of \$3.00. Bootstrapped 90% confidence intervals reported below point estimate.

Table 36: Counterfactual Profits: Alternative Market Designs

	Realized	$\widehat{\text{Treatment}}$	Auction	Auction (S)	Proxy-Means(100)	Proxy-Means(150)	Cei
Average	-433	-433	0	-339	-230	-287	1
	(-489,-374)	(-488,-375)	(0,0)	(-372,-309)	(-251,-208)	(-313,-254)	(101
10,000	-1830	-1917	0	-328	-275	-317	9
	(-2027,-1635)	(-2114,-1732)	(0,0)	(-410,-263)	(-343,-216)	(-383,-249)	(66,
15,000	-231	-214	0	-352	-252	-313	1
	(-267,-189)	(-240,-185)	(0,0)	(-381,-320)	(-275,-228)	(-341,-276)	(102
17,500	156	193	0	-308	-140	-197	1
	(125,188)	(167,228)	(0,0)	(-385,-241)	(-171,-106)	(-240,-144)	(98,
20,000	495	495	0	-385	-226	-312	1
	(222,764)	(367,612)	(0,0)	(-527,-236)	(-309,-129)	(-425,-188)	(107

Realized is the experimental value, $\widehat{\text{Treatment}}$ is the model predicted value, Auction is the predicted outcome w given the procurement auction prices, Proxy-Means(100) and Proxy-Means(150) refer to predicted outcomes using p at 100% and 150% of the poverty line, Cieling is a policy that constrains prices to be the average price in the sea denotes an additional subsidy of $3.00. Bootstrapped 90% confidence intervals reported below point estimate.

Table 37: Design Variations: Budget Balance

	Realized	$\widehat{\text{Treatment}}$	Auction	Auction(S)	PMT(100)	PMT(150)	Market	Market(S)
Average	-101.63	-99.88	-10.87	-11.2	-275.55	-63.28	-12.12	-12.47
	(-150.24,-53.17)	(-161.81,-53.14)	(-130.4,24.8)	(-130.84,-0.02)	(-826.06,-136.23)	(-238.61,-0.01)	(-122.39,22.37)	(-123.75,21.9
10,000	-1334.07	-1404.44	-10.53	-10.86	-330.49	-69.86	-34.83	-35.16
	(-1476.81,-1189.47)	(-1582.31,-1275.55)	(-132.07,23.98)	(-133.41,-0.02)	(-1078.67,-154.57)	(-269.76,-0.02)	(-150.15,5.07)	(-153.85,4.5
15,000	116.28	116.18	-11.29	-11.64	-301.59	-68.97	-13.19	-13.55
	(90.51,143.73)	(91.37,135.85)	(-135.36,25.82)	(-135.96,-0.02)	(-887.99,-150.43)	(-260.44,-0.02)	(-128.47,22.17)	(-129.95,21.6
17,500	334.04	403.51	-9.88	-10.19	-169.2	-43.37	3.97	3.66
	(272.64,400.81)	(343.02,470.39)	(-117.85,21.94)	(-119.36,-0.01)	(-499.1,-77.4)	(-157.6,-0.01)	(-93.14,33.18)	(-93.88,32.6
20,000	688.77	695.92	-12.35	-12.73	-268.62	-68.5	17.07	16.69
	(308.13,1064.46)	(488.65,855.53)	(-147.85,24.95)	(-151.77,-0.02)	(-710.07,-103.86)	(-276.51,-0.02)	(-109.63,50.29)	(-111.17,49.8

Realized is the experimental value, $\widehat{\text{Treatment}}$ is the model predicted value, Auction is the predicted outcome when households are given the procurement auc prices, Proxy-Means(100) and Proxy-Means(150) refer to predicted outcomes using proxy-means testing at 100% and 150% of the poverty line, Cieling is a po that constrains prices to be the average price in the search market, and (S) denotes an additional subsidy of \$3.00. Bootstrapped 90% confidence intervals repo below point estimate.

Table 38: Design Variations, 10,000 CFA Group: Budget Balance

	Realized	$\widehat{\text{Treatment}}$	Auction	Auction(S)	PMT(100)
Average	-101.63	-99.88	-1418.99	-1419.51	-1259.78
	(-150.24,-53.17)	(-159.6,-52.95)	(-1802.73,-1103.99)	(-1803.39,-1104.39)	(-1938.78,-930.51
10,000	-1334.07	-1404.44	-1386.67	-1387.18	-1521.22
	(-1476.81,-1189.47)	(-1574.82,-1275.46)	(-1559.38,-1261.96)	(-1559.96,-1262.42)	(-2377.73,-1216)
15,000	116.28	116.18	-1463.3	-1463.84	-1370.85
	(90.51,143.73)	(92.26,136.24)	(-1814.53,-1158.73)	(-1815.2,-1159.17)	(-2114.18,-1019.0
17,500	334.04	403.51	-1313.77	-1314.25	-786.08
	(272.64,400.81)	(341.09,471.39)	(-1933.9,-853.28)	(-1934.6,-853.6)	(-1297.99,-497.51
20,000	688.77	695.92	-1569.44	-1570.01	-1211.28
	(308.13,1064.46)	(488.66,855.47)	(-2433.15,-830.5)	(-2434.06,-830.82)	(-2024.36,-615.05

Realized is the experimental value, $\widehat{\text{Treatment}}$ is the model predicted value, Auction is the predicted outcome v
Means(100) and Proxy-Means(150) refer to predicted outcomes using proxy-means testing at 100% and 150% of t
average price in the search market, and (S) denotes an additional subsidy of $3.00. Bootstrapped 90% confidence in

Table 39: Counterfactual Subsidization Rates: Alternative Information Designs

	Realized	Original(0)	Conservative(1)	Municipal(2)	Municipal(2B)	NGO(3)	NGO(3B)
Average	-2.66	-2.16	-2.16	-3.25	-4.21	-2.17	-0.39
	(-3.01,-2.31)	(-2.42,-1.88)	(-2.4,-1.92)	(-3.66,-2.79)	(-4.93,-3.49)	(-2.35,-1.95)	(-1.18,0.78)
10,000	-11.35	-8.28	-7.68	-7.92	-7.4	-5.86	-4.77
	(-12.58,-10.17)	(-9.2,-7.36)	(-8.56,-6.81)	(-8.83,-7.03)	(-8.26,-6.52)	(-6.51,-5.09)	(-5.63,-3.29)
15,000	-1.41	-1.35	-1.42	-3.24	-3.85	-1.47	0.61
	(-1.62,-1.16)	(-1.52,-1.16)	(-1.6,-1.24)	(-3.7,-2.67)	(-4.53,-3.09)	(-1.63,-1.31)	(-0.38,1.97)
17,500	0.98	0.65	0.26	-0.19	-2.84	-0.88	1.18
	(0.78,1.19)	(0.52,0.82)	(0.11,0.43)	(-0.4,0.03)	(-3.95,-1.92)	(-1.11,-0.68)	(0.49,1.99)
20,000	3.15	1.64	1.64	1.56	-1.25	-0.16	0.15
	(1.47,4.86)	(1.15,2.09)	(1.15,2.09)	(1.08,2.01)	(-1.77,-0.78)	(-0.46,0.14)	(-1.36,1.52)

Gives market shares for alternative information structures, defined in figure 14. Bootstrapped 90% confidence intervals reported below point estimate.

Table 40: Counterfactual Profits: Alternative Information Designs

	Realized	Original(0)	Conservative(1)	Municipal(2)	Municipal(2B)	NGO(3)
Average	-433	-352	-351	-528	-681	-352
	(-489,-374)	(-393,-306)	(-390,-311)	(-594,-450)	(-797,-563)	(-383,-315)
10,000	-1830	-1340	-1242	-1287	-1204	-950
	(-2027,-1635)	(-1488,-1192)	(-1383,-1101)	(-1431,-1140)	(-1342,-1057)	(-1056,-825)
15,000	-231	-221	-231	-524	-622	-239
	(-267,-189)	(-249,-189)	(-260,-200)	(-598,-432)	(-730,-499)	(-265,-211)
17,500	156	101	39	-33	-454	-143
	(125,188)	(81,127)	(15,67)	(-67,2)	(-630,-307)	(-181,-110)
20,000	495	258	258	246	-202	-27
	(222,764)	(181,329)	(181,329)	(170,317)	(-287,-126)	(-76,20)

Gives market shares for alternative information structures, defined in figure 14. Bootstrapped 90% confidenc
reported below point estimate.