Targeting subsidies through price menus: Menu design and evidence from clean fuels*

Sulemana Abubakari† K.P. Asante† Misbath Daouda‡
B. Kelsey Jack‡ Darby Jack§ Flavio Malagutti¶ Paulina Oliva‖

July 11, 2024

Abstract

We examine the feasibility of targeting subsidies through non-linear pricing. We consider a policy maker who wants to increase the use of cleaner cooking fuel, while keeping the subsidy budget low by targeting subsidies to more price sensitive individuals. Fuel is sold in cylinders of different sizes, allowing for size-differentiated pricing, but not limits on the frequency of purchase. We show that a separating equilibrium can be sustained without purchase limits. Next, we demonstrate implementation in two stages: (1) we gather the empirical inputs for a sufficient statistic based approach to menu design, and then (2) test the optimal non-linear price menu against counterfactual linear pricing. Relative to the counterfactual, allocating more of the subsidy to smaller purchases achieves the same level of LPG demand with 30% less in public expenditure. Because poorer households have a preference for smaller sized purchases, non-linear pricing is also progressive in our setting.

*We thank Iddrisu Seidu, Alex Appiah, Sule Awuni and the field officers at the Kintampo Health Research Centre for collaboration, and Alexander Abajian, Akanksha Arora and Linnea Graham for research assistance. Numerous seminar and conference audiences provided useful feedback. We are grateful for financial support from Columbia World Projects and the King Climate Action Initiative at J-PAL. This RCT was registered in the American Economic Association Registry under number AEARCTR-0008400. Authors are listed in alphabetical order. Additional details are provided in an online supplement available here.

†Kintampo Health Research Centre
‡UC Berkeley
§Columbia University
¶UC Santa Barbara
‖University of Southern California

1 Introduction

Governments in nearly every country subsidize goods and services that support redistribution goals or generate positive externalities (World Bank, 2022). Uniform subsidies are costly. In many cases, better targeting can lead to the same policy objectives at a lower cost. Perhaps unsurprisingly, large literatures in public and development economics study alternative targeting approaches (e.g., Nichols and Zeckhauser, 1982; Alderman and Lindert, 1998; Hanna and Olken, 2018). We focus on non-linear pricing as a means of targeting product subsidies to increase their cost effectiveness. Like the textbook case of non-linear pricing by a monopolist, we consider differentiating prices based on the size of the purchase. Unlike the textbook case, we consider the objective of maximizing social goals rather than profits.[1]

In principle, targeting subsidies through non-linear pricing is attractive for its scalability. Unlike targeting on observables, such as proxy means tests (e.g., Alatas et al., 2012), it does not require observing characteristics of all potential recipients. Unlike many other approaches to targeting on unobservables, such as ordeal mechanisms or bidding (e.g., Jack, 2013; Alatas et al., 2016; Dupas et al., 2016), it can be rolled out using existing market structures. However, designing the optimal non-linear price menu to maximize social surplus requires both new theory and new empirical approaches. In this paper, we adapt the theory of second degree price discrimination to a policy problem. We build on our theoretical framework to develop a sufficient statistic approach for designing the optimal price menu, thus providing a bridge between theory and implementation. Finally, we implement the optimal non-linear price menu in a field experiment on household purchases of clean energy in Ghana, where we benchmark its performance against a counterfactual linear price subsidy.

The standard case for second degree price discrimination separates consumers on heterogeneity in demand: rents are extracted from low demand types who would rather pay less in total even if that means paying more per unit. The existing literature assumes the monopolist can effectively restrict the quantity sold per consumer. This is a seemingly innocuous assumption in the case of, for example, health insurance, where there are strong disincen-

[1] Long literatures on non-linear pricing in taxation (Mirrlees, 1971) and public utilities (Baumol and Bradford, 1970) are motivated by welfare maximization. Non-linear pricing may, more generally, increase consumer welfare relative to linear pricing (Brown and Sibley, 1986).

tives to hold multiple policies at once and policies have a fixed one-year duration. In the context of our policy problem, we extend the theory to a more general case that allows for continuous purchase frequencies per consumer over an arbitrary period of time. This allows individuals to either consume large packages slowly (similar to Hendel and Nevo (2013) and Hendel et al. (2014)) or replicate the quantities in large packages through repeated purchase of small packages. Relaxing this assumption can undo separating equilibria that exist under restrictions on purchase quantities. We present a stylized model with discrete package sizes, based on our empirical context, to show that a separating equilibrium can be sustained in the more general case if households are heterogeneous in their savings cost or transaction (e.g., time, travel or hassle) costs. In our model, both characteristics are important for separation and uniqueness: high saving costs work against pooling on larger package sizes; high transaction costs work against pooling on smaller package sizes.[2] Whether bulkier purchases should then be marked up or discounted relative to smaller purchases is an empirical question that depends on the objective function of the policy maker and on the demand levels and price elasticities of different consumer types for different package sizes.

Our theoretical framework serves as an existence proof: under plausible conditions, a non-linear price menu that separates consumers on characteristics other than demand – savings and transaction costs in our model – can maximize the policy objective function. Rather than estimating the optimal non-linear prices using proxies for the parameters in the model, we adopt a more flexible sufficient statistic approach that is agnostic about the sources of heterogeneity that cause some consumers to choose larger or smaller purchases. Specifically, the sufficient statistic for selection at any quantity-differentiated price pair is how much more the consumer would be willing to pay for the large package compared to the small package.[3] This, together with information about demand for the selected types under any price pair, fully identifies the optimal price menu. By reducing the dimensionality of the problem, our approach has the advantage of circumventing the need to explicitly model and measure all potential sources of heterogeneity across consumers. We demonstrate the

[2]Specifically, only one characteristic needs to vary in the population; both need to be positive to sustain separation for any given price menu.

[3]Nikzad (2023) shows that in constrained optimization problems, the optimal number of contracts in a menu is related to the number of constraints. Here, we take this result as given and only consider menus with two package sizes.

feasibility of this approach in the context of clean cooking fuel in Ghana.

Our setting is well suited to testing non-linear pricing for several reasons. First, the status quo polluting fuel used by most households in our sample is charcoal, which low income households purchase in small quantities multiple times per week. This is an initial indication of heterogeneity in choice of purchase size for the incumbent cooking fuel.[4] Second, while a tax is presumably the first-best policy to address the negative consumption externalities from charcoal, effective taxation is difficult both for practical and political reasons. Instead, Ghana, like many countries around the world, promotes cleaner substitutes, namely liquid petroleum gas (LPG), which we calculate generates roughly one-fifth the greenhouse gas emissions as charcoal.[5] LPG prices are set by a national regulator and all transactions are through formal markets. Third, households purchase LPG in cylinders of fixed sizes, creating the opportunity for quantity-differentiated pricing. We introduce study-operated LPG depots, where study participants can exchange cylinders at prices that we control.[6]

We implement our approach in two stages: (1) design, then (2) test. Each stage uses a stratified random sample of charcoal users drawn from the same population. In stage 1, we gather data for the sufficient statistic that determines selection into cylinder size and the corresponding demand for LPG among the selected types. Specifically, we use a multiple price list (MPL) elicitation format to measure willingness to pay (WTP) for LPG in cylinders of different sizes. The sufficient statistic for each household is the difference in WTP across the two sizes, which captures the relevant heterogeneity for size selection – the consumer types. Conditional on responses to the MPL, participants receive a randomly selected size-price pair. They have three months to purchase LPG (i.e., exchange empty cylinders for full ones) at the selected price with no restriction on the frequency of purchase. The resulting demand data allow us to estimate price elasticities for different consumer types when exchanging each of the two cylinder sizes. Consumer types that opt for small packages

[4]A tendency for low income consumers to purchase in small package sizes has been documented by others (e.g., Pires and Salvo, 2015; Dillon et al., 2021), including Attanasio and Pastorino (2020) who study non-linear pricing in the presence of heterogeneous budget constraints that contribute to poor households' preferences for small purchase sizes.

[5]Subsidies for "greener" substitute goods are common around the world, in both developed and developing countries (Fowlie and Meeks, 2021; Blanchard et al., 2023).

[6]Households obtain cylinders on deposit, and stoves are often subsidized, so we supress the fixed cost of acquiring durables in our theory and implementation.

are more price sensitive, which implies a larger demand response from subsidizing LPG in the small cylinder than in the large one. After imposing a budget constraint, we solve for the quantity-differentiated price subsidies that maximize the policy objective function (i.e., demand). This objective function balances the health and environmental costs of charcoal use against the cost of public spending on subsidies.[7]

In stage 2, we test the optimal non-linear prices against linear prices (i.e., prices that do not vary with cylinder size) that are budget neutral in expectation. We randomly assign households to a non-linear price treatment or a linear price control. Households choose their preferred cylinder size at the treatment-specific prices, and – like in stage 1 – purchase LPG by exchanging their cylinders at a study-run depot for a period of three months. We predict that the treatment will increase the cost effectiveness of subsidies; either the policy maker will get more demand for her budget, or spend less for at least as much demand.[8]

We find that the non-linear price treatment delivers the same overall demand as linear pricing, but at a roughly 30% lower budget. This result is driven by a combination of selection across sizes and differential price elasticities.[9] As predicted by the stage 1 model, large cylinder types are relatively insensitive to price, so the higher per kg price under non-linear pricing does little to dampen their demand, while reducing public expenditures. Small cylinder types, on the other hand, increase their demand in response to the subsidy, while costing the policy maker comparatively less because small cylinder types consume less LPG. The additional subsidy on the small cylinder induces an additional 5.5% of households to choose the small cylinder. Finally, the sufficient statistic for consumer type is correlated with observable proxies for saving costs and wealth, and non-linear pricing effectively targets a larger share of the total subsidy value to poorer households.

The policy maker's ultimate goal is to reduce externalities from household energy use. While we cannot directly measure charcoal substitution, we use the relative greenhouse gas emissions from LPG and charcoal to calculate the minimum substitution rate that would

[7]A broader welfare maximization objective would also account for consumer surplus and non-monetary costs. We provide a back of the envelope calculation of the overall welfare effects of non-linear pricing.

[8]Our test could have imposed budget neutrality or demand neutrality; we chose the former as the basis for our design given our fixed sample size, though our results end up closer to the latter. We pre-specified testing for both alternatives.

[9]Take up of the program as a whole is very high: 94% of those offered the chance to enroll took it. We therefore see little effect on the extensive margin, i.e., choosing any cylinder vs. choosing none.

justify subsidies of the magnitude we implement. At a substitution rate of 0.38 or higher, even a low social cost of carbon ($51 per ton) implies benefits in excess of the subsidy cost. Our implementation offers a proof of concept for feasible design of non-linear pricing, with potential applications in a range of settings where consumer good subsidies can be applied to discrete package sizes, including health products, educational materials and other "green" goods.[10] In principle, policy makers could follow our two stage approach, with careful data collection for a representative sample informing pricing for the rest of the population. Scale up, however, requires extrapolation over both time and populations. We discuss this and several other practical considerations for policy, including supply costs, the cost of durable goods, and restrictions on the simultaneous exchange of multiple small cylinders.

We contribute to a large literature that investigates different approaches to targeting subsidies and transfers (Coady et al., 2004; Alatas et al., 2012; Jack, 2013; Cohen et al., 2015; Alatas et al., 2016; Johnson and Lipscomb, 2022; Ida et al., 2022; Rafkin et al., 2023). Most of these papers either test alternative targeting strategies in the field (e.g., Jack, 2013; Alatas et al., 2016) or use empirical inputs to design a theoretically optimal targeting scheme (e.g., Ito et al., 2023; Ida et al., 2022; Langer and Lemoine, 2022). A few do both, including Johnson and Lipscomb (2022) and Dubé and Misra (2023), both of which implement a two-stage approach in which the relationship between observables and willingness to pay is measured in a first stage, and subsidies are targeted based on observables in a second stage (i.e., third degree price discrimination). We, instead, design and test a targeting strategy that leverages heterogeneity in unobservable preferences.

We follow the literature on second degree price discrimination in setting up the price menu design as a screening problem, with heterogeneous consumer preferences.[11] A small literature examines second degree price discrimination as a policy tool for improving the cost effectiveness of public programs. Most notably, Dizon-Ross and Zucker (2023) test quantity

[10]One potential distinction between LPG and other socially desirable consumption is the nature of the transaction or hassle costs. Purchasing multiple small LPG cylinders at a time is potentially more costly than e.g., multiple small pill bottles, both because the cylinders themselves are a hassle to transport and because swapping cylinders in the home requires changing the cylinder-stove connection, which is more cumbersome than opening another bottle.

[11]This falls into a much larger class of non-linear pricing problems (see Armstrong (2016) for a review). Public utilities have long used non-linear tariffs for both cost recovery and redistribution (Borenstein, 2012; Szabó, 2015; Nauges and Whittington, 2017).

differentiated incentive contracts for completing a target number of steps per day against a flat incentive contract.[12] We complement their empirical evidence with a methodology for how to design the optimal price menu, and apply it to product pricing.[13] In doing so, we extend the standard model of second degree price discrimination by adding an intensive margin of demand, which – to our knowledge – has not been explicitly incorporated into the menu design problem. This extension allows theory on menu design to apply to a much wider set of empirical settings (including ours). A mostly theoretical literature in mechanism design also considers a social planner's objective function (e.g., Nikzad, 2023), and sometimes arrives at non-linear pricing as the optimal contract. Most relevant to our setting, Kang (2022) studies the case of non-linear pricing to indirectly tax externalities from gasoline consumption, and shows that the optimal contract could result in either quantity discounts or mark-ups depending on the relationship between price sensitivity and the externality.[14]

Finally, our application links us to a literature in development economics and public health on household cooking practices and indoor air pollution. Much of the prior work has focused on the health and environmental consequences of cooking with biomass in low and middle income countries (e.g., Smith et al., 2011; Bensch and Peters, 2015; Hanna et al., 2016; Berkouwer and Dean, 2022). Findings have generally been disappointing: even improved cookstoves that are taken up and used have modest (at best) impacts on health outcomes (Berkouwer and Dean, 2023). More recently, the emphasis has shifted toward cleaner fuels, where findings from large scale government programs are more positive (Gould et al., 2024). While these public programs have achieved some success, they have done so through large public expenditures and limited success in targeting subsidies to marginal users. We demonstrate a novel and scalable approach to targeting LPG subsidies to lower public expenditures while meeting clean energy goals.

[12] The setting in Dizon-Ross and Zucker (2023) resembles a larger literature on non-linear incentive contracts designed and tested in the personnel literature, many of which have social goals in the sense that they are applied to jobs like teaching or health care (e.g., Duflo et al., 2012).

[13] To our knowledge, the only other field experimental test of second degree price discrimination is by Levitt et al. (2016), who find no effect on demand or profits in an online retail setting.

[14] Kang (2022) does not, however, consider that high demand drivers could purchase multiple smaller quantities of gasoline or that low demand drivers could store across periods.

2 Background and context

We start by offering additional background on the cooking technologies and the policy problem we study, which underlies our theory and empirical strategy.

2.1 Pollution from household energy use

Pollution resulting from the burning of solid fuels or biomass (firewood, dung and charcoal) for cooking and heating is a leading cause of morbidity and mortality around the world (World Health Organization, 2023), and contributes to climate change through the release of greenhouse gases and black carbon and through degradation of forests (Bailis et al., 2015). Replacing biomass with cleaner burning fossil fuels, including LPG, has the potential to reduce emissions of both local pollutants and greenhouse gases (Floess et al., 2023).[15] LPG offers a scalable alternative to biomass, since it relies on decentralized distribution networks, and is relatively easy to transport and use (Gould et al., 2024).

Many governments have clean cooking targets that rely on the scale up of LPG as a cooking fuel, and many use subsidies to help reach these targets. For example, the World LPG Association estimates that India, Indonesia and Morocco together spent over $10 billion on residential LPG subsidies in 2022 (World LPG Association, 2023). In some cases, countries including Indonesia and Ivory Coast have tried to target subsidies by cylinder size, similar to the approach tested here, but with very large price differentials that have resulted in pooling on the cheaper size. A combination of subsidies and economic growth have contributed to the rise of LPG as a domestic cooking fuel: in the late 2000s, gas fuels overtook biomass as the leading cooking fuel in low- and middle-income countries (Stoner et al., 2021). Between 2010 and 2020, 75% of households that transitioned to cleaner fuels switched from solid fuels to LPG (Floess et al., 2023).

[15] While renewable energy sources offer an even cleaner alternative to biomass, current renewables technology cannot meet global cooking energy demand at a feasible cost.

2.2 LPG and charcoal in Ghana

The government of Ghana has a target of 50% LPG access by 2030, alongside a commitment to keep subsidization to a minimum. At present, biomass remains the primary cooking fuel in Ghana. The national census in 2021 found that around 54% of households used charcoal or firewood as their primary fuel. Only 37% of households used LPG as a primary fuel (Ghana Statistical Service, 2022).

LPG is purchased in cylinders and consumed by connecting a cylinder to a stove. Thus, LPG adoption requires both a one-time purchase of a stove and cylinder, and recurring expenditures of both money and time to acquire fuel.[16] As part of the policy push toward LPG, Ghana is transitioning from a refill model, in which households bring their cylinder to a filling station and choose how many liters to refill, to a recirculation model, in which households exchange empty cylinders for full ones. The cylinder recirculation model was announced in 2017 and launched in September 2023, at the end of our study. We design our data collection around a pilot version of cylinder recirculation (see Section 4).

LPG cylinders come in a range of sizes. For home cooking, 3, 6 and 14.5 kilogram cylinders are the most common. Under cylinder recirculation, suppliers in Ghana own the cylinders and are responsible for their maintenance; customers pay a one-time deposit for the cylinder and purchase gas by exchanging an empty cylinder for a full one. This represents an important change in how households pay for fuel: requiring purchase of a full cylinder reduces the flexibility in expenditures relative both to a refill model and to charcoal. However, the lumpiness in the quantities of gas available for purchase also represents an opportunity for pricing that is differentiated on cylinder size.

We gather data on status quo charcoal purchasing behavior from our study sample. Households purchase charcoal in small quantities and at high frequency, on average (Table 1). Around 40% of the sample reports buying charcoal in small bags of 1-2 kg. The first column shows the average of each characteristic for households that buy in large quantities and the second column shows the estimated coefficient on purchasing in small quantities in

[16] A 2-burner stove costs roughly the same as a full large cylinder. Numerous donors offer subsidized or free stoves, so we choose to ignore this aspect of the targeting problem and focus on the challenge of repeated purchases and use of clean fuels.

Table 1: Baseline charcoal purchasing patterns (stage 2)

	Large (mean)	Small (coef/se)
Number of monthly purchases	5.048	32.099***
	(14.037)	(1.041)
Implied monthly expenditure (GH¢)	298.451	-114.556***
	(840.885)	(37.487)
Implied monthly travel cost (GH¢)	5.939	-4.831**
	(40.343)	(2.113)
Implied monthly travel time (hours)	0.683	1.749***
	(2.342)	(0.220)
Wealth quintile	3.559	-0.234**
	(1.398)	(0.098)
Any formal savings	0.805	-0.126***
	(0.397)	(0.031)
Days to get 150 GH¢	5.890	0.346**
	(2.159)	(0.157)
High opportunity cost of time	0.326	0.032
	(0.469)	(0.034)
Any no cost transport access	0.301	-0.016
	(0.459)	(0.032)
Household size	5.108	-0.246
	(3.037)	(0.224)

Note: Column 1 shows means and (standard deviations). Column 2 shows coefficients and robust standard errors from separate linear regressions of each row variable on a small purchases indicator. 39% of households buy in small quantities at baseline. * p<0.1, ** p<0.05, *** p<0.01

separate regressions for each characteristic. Households that buy in large bags buy an average of just over five bags per month, versus 32 times more per month among those who purchase in small quantities. Households that purchase in small quantities spend less on charcoal in total and less on travel to buy it, but spend around three times more time buying it. They are also less wealthy, less likely to have formal savings and are more credit constrained. We observe no evidence of measurable differences in opportunity cost of time or access to transportation across purchasing types. These patterns may be shaped by other factors, including on the supply side, but suggest a revealed preference for small, high frequency purchases among poorer households, in spite of the time and hassle costs involved.[17]

[17]This echoes patterns of high frequency, small quantity purchasing behavior among low income households in other settings (Dillon et al., 2021; Pires and Salvo, 2015).

2.3 Externalities from charcoal and LPG

Both charcoal and LPG generate externalities when burned. The argument for policy intervention to switch households from charcoal to LPG hinges on higher relative externalities from charcoal. We do not measure these directly in our data collection, and so rely on the literature for a back of the envelope calculation. We focus on damages external to the household, though intrahousehold externalities or misinformation might also imply that some of the health damages accruing to household members are also inefficient.

Cooking with either charcoal or LPG emits both local air pollutants and greenhouse gases. The former imposes health damages on other households within an airshed, the magnitude of which depends on ambient pollution levels and the vulnerability of the exposed population. Given the complexity in these calculations, we construct a lower bound on the external damages based on greenhouse gas emissions only. We are primarily interested in the relative emissions from the two fuels. Floess et al. (2023) calculate the CO_2 equivalent emissions per MJ of usable energy for both charcoal and LPG, from both combustion and production. Per MJ, charcoal emits roughly 5 times the CO_2e as LPG, including production, combustion and transportation.[18] Using a social cost of carbon range from USD 51 to 204 (US EPA's two most recent numbers) the avoided damages range from USD 1.7 to 6.8 per kg of LPG that replaces the equivalent energy quantity of charcoal.[19]

The first best policy intervention to address external damages from charcoal would target charcoal directly, for example by taxing purchases at the marginal damage from charcoal. However, in spite of efforts to formalize the charcoal supply chain in Ghana, much of the market remains informal. This creates practical challenges for the feasibility of taxing the "bad." Instead, a second best policy, feasible in the nationally regulated LPG market, places a subsidy on the "good." The efficient level of the subsidy will depend on a number of factors, including the substitution rate between charcoal and LPG.

[18]Calculations can be found here. We adjust for the Floess et al. (2023) Ghana-specific assumed fraction of non-renewable biomass used in charcoal production of 27.7%.

[19]The ratio of usable energy per weight is around 1.6; for each kilogram of LPG used in cooking, around 1.6 kilograms of charcoal would be needed to generate the same energy.

3 Theory

This section has two purposes. First, to establish that plausible conditions are sufficient for heterogeneous consumers to sort across two different sizes of LPG cylinders. Second, to set up the policy maker's optimization problem, which internalizes this sorting.

Our empirical approach to designing the contracts we test in stage 2 does not rely on the specific functional form assumptions that we introduce in this section. However, we show suggestive empirical evidence that the sources of heterogeneity that drive separation in theory are also present in our empirical setting.

3.1 Separating equilibrium with unconstrained purchase frequency

We ground our theory in our study setting, and explore the advantages of lumpy purchases (i.e., a discrete number of fixed package sizes) of LPG for targeting government subsidies. Note, however, that lumpiness does not constrain the amount of LPG that can be purchased in a time period of a given length. Individuals that commit to a cylinder size can exchange that cylinder as many times as they want. This is, in fact, a general characteristic of non-perishable consumption goods that are sold in lumpy quantities. A household can consume the same amount of toilet paper per week by either buying a large family package that will be consumed over four weeks or a single small package each week. This possibility of making a continuum of purchases (including a fraction, e.g., a quarter of a family package if time is measured in weeks) over a single period of time rules out that the seller can constrain quantities purchased, a necessary assumption in most of the classic contract theory literature for a separating equilibrium. However, quantity discounts persist in many markets even in the absence of feasible constraints on the frequency of purchase. One could argue that supply cost reasons drive these differences: the marginal cost per unit is smaller when packages are larger. But supply cost differences cannot explain why, in equilibrium, different consumers pick different sizes in the presence of differential per-unit pricing. We propose a model that incorporates transaction costs and saving costs, and yields a separating equilibrium in the absence of constraints on purchase frequency.

We start with a parameterization of the consumer problem that allows us to incorporate sufficient dimensions of heterogeneity while retaining some tractability. We summarize the model here, and provide details and formal proofs in Appendix B. Modeling savings and purchases of durables (a full cylinder may last for several weeks) generally requires a dynamic framework. For simplicity, we set up a static consumer problem that accounts for the lumpiness of LPG purchases and the presence of saving costs that affect the total costs of consuming LPG. Importantly, we reflect the freedom consumers have in buying any amount of LPG in a given period by adjusting the frequency of exchanges. Thus, we allow the number of exchanges in a given period to be continuous, as opposed to binary.[20]

The consumer's maximization problem in a given period is given by:

$$\max_{c,q,x} u(c, qx; \theta) \tag{1}$$

$$\text{s.t. } c + p(q)x + D(\theta)x + 1(x > 0)\frac{p(q)}{x}A(\theta) = y(\theta)$$

where $q \in \{q_S, q_L\}$ is the size of the LPG cylinder in kg, c denotes consumption measured in GHc, x is the number of exchange trips and can take any non-negative value (including between 0 and 1), $p(q)$ is the price of an exchange as a function of cylinder size, $D(\theta)$ is the transaction cost per exchange, $1(x > 0)\frac{p(q)}{x}A(\theta)$ is the saving costs (which we explain in detail below), and $y(\theta)$ is income. Assume that type, θ, determines income, $y(\theta)$, saving costs, through $A(\theta)$, and transaction costs, through $D(\theta)$. Type may also determine preferences for LPG, such that $\frac{\partial u}{\partial qx \partial \theta} > 0$.

The term that governs the savings cost can be broken out as follows:

$$\underbrace{1(x > 0)}_{\substack{\text{if positive} \\ \text{exchanges}}} \quad \underbrace{p(q)}_{\substack{\text{amount to} \\ \text{hold-on to}}} \quad \underbrace{\frac{1}{x}}_{\substack{\text{\# of periods} \\ \text{until exchange}}} \quad \underbrace{A(\theta)}_{\substack{\text{cost per period} \\ \text{per dollar}}}$$

The logic behind this parameterization of saving costs is as follows: when the frequency of purchase is less than one (e.g., $x = 0.5$), then the household has to hold on to the quantity of money $p(q)$ for multiple periods until it is time to buy the next cylinder. $A(\theta)$ represents the cost of holding on to one dollar for a full period without spending it. Thus, if the frequency

[20]Consistent with our empirical setting, we abstract from the decision to acquire LPG-related durables and focus on cylinder size choice and demand.

of purchase is $x = 0.5$ per period, the number of periods you have to hold on to that dollar in order to purchase the cylinder is $\frac{1}{x} = \frac{1}{0.5} = 2$. If the cost of saving one dollar per period is 30 cents ($A(\theta) = 0.3$), then the additional cost that is incurred is $p(q) \times \frac{1}{x} \times A(\theta) = p(q) \times 2 \times 0.30$.

In order to make this problem tractable, we first analyze the optimization problem conditional on q, and then analyze the size choice across the discrete number of sizes. Consider the maximization problem conditional on q:

$$\begin{aligned} &\max_{c,x} u(c, qx; \theta, q) \\ \text{s.t. } & c + px + D(\theta)x + 1(x > 0)\tfrac{p}{x}A(\theta) = y(\theta) \end{aligned} \tag{2}$$

The first order conditions to (2) imply

$$\frac{u_2 q}{u_1} = p + D(\theta) - \frac{pA(\theta)}{x^2}, \tag{3}$$

where u_1 is the partial derivative of the utility function with respect to the first argument (non-LPG consumption), and u_2 is the partial derivative of the utility function with respect to the second argument (LPG in kg). Thus, the expression in (3) shows that the optimal number of exchanges is such that the marginal utility of an exchange (left hand side) is equal to the marginal cost of an exchange (right hand side). Note that the marginal cost of an exchange depends on the number of exchanges because of the non-linearity introduced by saving costs. Denote the solution to (2) as

$$x^*(p; \theta, q), \;\; c^*(p; \theta, q). \tag{4}$$

We now turn to the optimal choice across sizes. To do so, we plug (4) back to the utility function in (1) in order to obtain the indirect utility function conditional on size:

$$\equiv u(c^*(p(q); \theta, q), x^*(p(q); \theta, q)q).$$

The condition that governs the choice of the large cylinder over the small one is given by

$$v(p(q_L); \theta, q_L) - v(p(q_S); \theta, q_S) > 0. \tag{5}$$

Single-crossing and separating equilibrium

The two conditions we need for establishing the existence of a separating equilibrium are the satisfaction of the single crossing property (SCP) and indifference between sizes by individuals characterized by a type within the continuum of types.

Proposition 1 We assume that the marginal utility of consumption, u_c, is constant, that $x^*(p;\theta,q)$ is non-decreasing in q, and that $p(q)$ is increasing in q.[21] The last assumption amounts to assuming that exchanging a larger cylinder is more expensive than exchanging a small one. We also assume that $A(\theta) \geq 0$, $D(\theta) \geq 0$, $\frac{\partial A(\theta)}{\partial \theta} = A'(\theta) \leq 0$ and $\frac{\partial D(\theta)}{\partial \theta} = D'(\theta) \geq 0$. Then, $v(p(q);\theta,q)$ is single crossing in (q,θ) and $q^*(\theta)$ is non-decreasing.

The proof can be found in Appendix B. The intuition for the role of costs $D(\theta)$ and $A(\theta)$ is as follows: as θ increases, purchasing in large quantities becomes more attractive because transaction costs are larger ($D'(\theta) \geq 0$) and saving costs are lower ($A'(\theta) \leq 0$).

The SCP guarantees that that indifference conditions cross at most once. If they never cross, then there is a pooling equilibrium (everyone prefers a single size). Proposition 2 discusses how the existence of both a transaction cost, $D(\theta) > 0$, and a savings cost, $A(\theta) > 0$, partitions the support of θ by preference over cylinder size.

Proposition 2 In addition to the assumptions under Proposition 1, assume that agents cannot borrow and that income is non-decreasing in type ($y'(\theta) \geq 0$). Assume also that there is a finite lower bound to transaction cost $D(\theta)$ and a zero lower bound to saving costs $A(\theta)$. Finally, assume that exchanging the large cylinder is unaffordable for some types in the lower part of the support of θ, who can still afford the small cylinder.[22] Then, a separating equilibrium is guaranteed to exist under linear prices or bulk discounts, and can still exist under bulk markups as long as the transaction cost, $D(\theta)$, is sufficiently large.

The proof of this proposition can be found in Appendix B. Again, it is helpful for intuition to think about the role of transaction and saving costs. First, note that in the absence of

[21] In Lemma 1 of Appendix B we show that a $x^*(q,\theta)$ that is non-decreasing in q follows from the utility function being single crossing in $(x,-q)$.

[22] These assumptions are formally described in Appendix B.

transaction and saving costs – i.e., $D(\theta) = A(\theta) = 0$ – individuals will be indifferent between the large and the small cylinder under linear pricing. This is because the price per kg of LPG is the same for both sizes and there is no other cost of purchasing LPG. Next, introduce a transaction cost, $D(\theta) > 0$, still under linear pricing. Now, everyone experiences a higher cost of purchasing a small cylinder compared to a large cylinder. This would result in a pooling equilibrium where everyone chooses the large cylinder, even if there is heterogeneity in $D(\theta)$. Now, introduce $A(\theta) > 0$. This cost penalizes the purchases of the large cylinders, and more so for those with small θ. Because this cost moves in the opposite direction as $D(\theta)$ as θ increases, there will be a type, θ', for whom the transaction costs and saving costs will exactly balance out. Because of SCP, we know that types above this indifferent type, θ', will prefer the large cylinder and types below θ' will prefer the small cylinder.

Although this intuition is simplest for the case of linear prices, it also applies to non-linear prices. Once we have established the indifferent type for linear prices, we can think of changes in pricing that yield non-linear schedules as changing the value of the indifferent type. E.g., if we make the large cylinder more expensive, the indifferent type will have a lower value of θ compared to the case with linear pricing.

3.2 Policy maker's optimization problem

Propositions 1 and 2 establish sufficient conditions under which a separating equilibrium will exist in a market with two different cylinder sizes. They show that, when different types face different saving costs and transaction costs, each cylinder size will appeal to a different subset of the type support even if the per-kg price differs across them. Separation opens the door for the policy maker to target subsidies through size-specific pricing.

We assume that the policy maker cares about increasing demand for LPG (in kg) at the lowest public cost possible. Thus, the policy objective function can either maximize demand subject to a given budget, or minimize the budget, subject to a given demand target. Next, we define what the demand and budget functions look like in order to discuss the trade-offs faced by the policy maker when changing the prices of each size of cylinder.

Per-household average demand for LPG in kg is given by

$$
\begin{aligned}
Q\left(p(q_L), p(q_S)\right) \quad = \quad & \underbrace{F(\theta')}_{\substack{\text{share that} \\ \text{chooses } S}} \quad \underbrace{q_S}_{\substack{\text{converts} \\ \text{into kg}}} \quad \underbrace{\int_{-\infty}^{\theta'} x^*(p(q_S);\theta, q_S)\frac{f(\theta)}{F(\theta')}d\theta}_{\substack{\text{avg. \# of } S \text{ exchanges} \\ \text{for those that select into } S}} \\
+ \quad & \underbrace{(1-F(\theta'))}_{\substack{\text{share that} \\ \text{chooses } L}} \quad \underbrace{q_L}_{\substack{\text{converts} \\ \text{into kg}}} \quad \underbrace{\int_{\theta'}^{\infty} x^*(p(q_L);\theta, q_L)\frac{f(\theta)}{1-F(\theta')}d\theta}_{\substack{\text{avg. \# of } L \text{ exchanges} \\ \text{for those that select into } L}}
\end{aligned}
\tag{6}
$$

where $F(\theta)$ and $f(\theta)$ are the CDF and PDF of θ. The average public expenditure per household are given by an expression with a similar structure:

$$
\begin{aligned}
B\left(p(q_L), p(q_S)\right) \quad = \quad & F(\theta') \quad (r \times q_S - p(q_S)) \quad \int_{-\infty}^{\theta'} x^*(p(q_S);\theta, q_S)\tfrac{f(\theta)}{F(\theta')}d\theta \\
+ \quad & (1-F(\theta')) \quad (r \times q_L - p(q_L)) \quad \int_{\theta'}^{\infty} x^*(p(q_L);\theta, q_L)\tfrac{f(\theta)}{1-F(\theta')}d\theta
\end{aligned}
\tag{7}
$$

where r is the cost per kg of LPG to the policy maker.

Equations (6) and (7) illustrate the tradeoffs that the policy maker balances when choosing $p(q_L)$ and $p(q_S)$ optimally. For example, assume a fixed public budget. As more of that budget is allocated towards subsidizing gas in the small cylinder relative to the large cylinder, two margins will change. First, the number of large cylinder exchanges will fall, while the number of small cylinder exchanges will increase. If the low θ types are more price-responsive than the high θ types when exchanging small cylinders, this will increase average demand. Second, more individuals will select into the small cylinder (θ' will increase). The sign of the effect of this margin on demand depends on the sensitivity of the switchers to prices, transaction costs and saving costs. For example, if switchers have high transaction costs, their demand level will fall when they switch to a small cylinder.

An advantage of restricting the policy maker to care about (6) and (7) is that both of these quantities are empirically observable outcomes in stage 2 of our study. However, these functions ignore the consumer surplus and deadweight loss associated with a given menu of prices. In Section 6 we provide a back of the envelope estimate of the broader welfare impacts of non-linear pricing.

3.3 A sufficient statistic approach

Finding the $p(q_L)$ and $p(q_S)$ that solve the policy maker's optimization problem described by (6) and (7) based on observable individual characteristics is complicated by the potentially numerous sources of heterogeneity summarized in θ. While Section 3.1 shows one plausible parameterization that is sufficient to induce separation, other parameterizations could also be sufficient. Thus, instead of estimating a fully specified demand model, our empirical approach is agnostic about the sources of heterogeneity that drive cylinder size choices.[23]

We rely on two shortcuts. First, we directly elicit a sufficient statistic for the relevant heterogeneity that governs selection into cylinder size. Second, we estimate a flexible average demand function conditional on selection. This approach still allows us to evaluate (6) and (7) for each price pair, $(p(q_L), p(q_S))$, and conduct a numerical search for the optimal menu. Here, we summarize the theory behind the first shortcut (see Appendix C for additional detail). Specifically, we show how all the relevant information for selection across cylinder sizes can be summarized by an intuitive economic quantity: the difference in maximum willingness to pay between the two sizes. Our strategy for reducing the dimensionality of the relevant heterogeneity is similar to the aggregation technique described by Rochet and Stole (2003). The second shortcut is described in Step 4 of Section 5.1.

The sufficient statistic for consumer type emerges from the selection condition in (5). We approximate the left hand side of this inequality using a Taylor expansion, which delivers:[24]

$$-\frac{\left.\frac{\partial v(p(q);\theta,q))}{\partial q}\right|_{q_S}}{\left.\frac{\partial v(p(q);\theta,q))}{\partial p(q)}\right|_{q_S}} > \frac{p(q_L) - p(q_S)}{q_L - q_S}. \tag{8}$$

This condition is intuitive: it states that for the individual to choose the large cylinder, the marginal willingness to pay (MWTP) evaluated at the quantity in the small cylinder needs

[23] A disadvantage of this approach, relative to a full structural model, is that it cannot be used to evaluate whether separation is guaranteed for a different population, based on its sample characteristics. In addition, this lack of structure limits the counterfactual analyses we can perform.

[24] A second-order Taylor expansion delivers similar empirical results so we implement a simpler first-order expansion.

to be larger than the per kg difference in prices between the two sizes.

Next, we show that the MWTP in (8) can be expressed as the difference in WTP for exchanging the large and the small cylinders. This yields an even simpler approximation to the choice condition in (5):

$$p(\theta, q_L) - p(\theta, q_S) > p(q_L) - p(q_S). \tag{9}$$

Inequality (9) states that individuals will choose the large cylinder when this choice delivers a larger surplus compared to the surplus from choosing the small. Given this, the share of individuals who choose the large cylinder under (exchange) prices $p(q_L), p(q_S)$ is given by

$$S_L = \int_{-\infty}^{\infty} \mathbf{1}\left(p(\theta, q_L) - p(\theta, q_S) > p(q_L) - p(q_S)\right) f(\theta)\theta.$$

Note that this expression depends on θ only through the difference in WTP across sizes, $p(\theta, q_L) - p(\theta, q_S)$. Thus, this difference is a sufficient statistic for selection. To simplify the discussion below, we denote this sufficient statistic as

$$\tilde{\theta} \equiv \tilde{\theta}(\theta) = p(\theta, q_L) - p(\theta, q_S).$$

4 Sample, data and implementation

In this section we discuss the sample, data and implementation, much of which is common to both the design (stage 1) and evaluation (stage 2) stages of our field experiment.

4.1 Sample

Our study is implemented in Techiman, a city of around 250,000 in the Bono East region of central Ghana. Techiman is the site of a Health and Demographic Surveillance System (KHDSS) run by the Kintampo Public Health Research Centre (KHRC). KHRC is one of three national health research centers under the Ghana Health Services, Ministry of Health. The KHDSS covers over 500,000 individuals across six districts, and collects regular (approximately every six months) household data.

Prior to stage 1, we drew samples for both stages using KHDSS data. First, we constructed a sampling catchment around four selected LPG cylinder exchange depot sites and included households more than 150 meters and less than 1 kilometer away from the depot. Second, we restricted eligibility to households that were primary charcoal users. Third, we stratified the sample on wealth (a durable assets index) and exchange depot and assigned households to the two stages of the experiment.[25]

During the baseline surveys (one per stage), we further excluded households that failed to meet the following screening criteria: (a) financial decision maker not available, (b) planning to move in next six months, (c) household contains more than 9 members, (d) household already uses LPG as their primary fuel, (e) household prepares food commercially.

4.2 Data and implementation

Study data come from the following sources.

KHDSS household data We use 2021 KHDSS data for for sampling and randomization. The dataset is a census of households in Techiman that includes information on household size, a wealth index, and primary cooking fuel.

Household surveys During each stage of the study, we gather baseline survey data as part of the enrollment process. Surveys were conducted with the head of household or another financial decision maker. The questionnaire was administered by a trained enumerator, and covered household demographics, finances, time and travel costs, and fuel use and expenditures. Cylinder size choices were elicited at the end of the household survey.

The surveys include questions corresponding to specific parameters in the theoretical model (see Section 3) and defined in our pre-analysis plan. Specifically, we measure the opportunity cost of time, saving costs as proxied by liquidity and credit constraints, the baseline propensity toward small purchases of the status quo cooking fuel (charcoal), and a

[25] The original stratification was on wealth quintile and depot. After revised power calculations based on data from stage 1 indicated that we required an expanded sample size for stage 2, we repeated the stratification combining quintiles 1 and 2 and quintiles 3 and 4 into larger wealth groupings. This was justified by data from stage 1 that showed a relationship between wealth and LPG demand only in the top quintile. Thus, stage 1 has 20 strata (4 depots and 5 wealth quintile) while stage 2 has only 12 (4 depots and 3 wealth groups).

durable asset index as a proxy for wealth or income (see Appendix I.4).

At the end of the three month exchange window, we administer a short endline survey, which measures self-reported fuel use and expenditures.

Cylinder size choices The enrollment process involved a decision of whether to take up an LPG "starter kit" (stove, empty cylinder with an ID code and regulator/hose). Households were required to make a deposit for these items, which did not vary with cylinder size.[26] In stage 1, households were randomly assigned to a deposit of 50 or 100 GH¢; in stage 2, it was set to GH¢ 50 for everyone.[27] At the time of the take up decision, participants received information about the cylinder exchange program and their subsidized LPG offer. They were also told that they would have a chance to switch cylinder sizes at the end of the 3-month exchange window if they wished to keep the starter kit and forgo their deposit.[28] Households were given up to two weeks to gather their deposit before the starter kit was delivered.[29]

Stage 1 choices Designing the optimal non-linear subsidy requires estimates of demand, accounting for selection, at each size-specific (small, large) price pair. We gather the necessary data for this exercises from two main sources. First, a baseline survey for stage 1 included a multiple price list (MPL) WTP elicitation that also introduced random variation into cylinder sizes and prices. Second, we observe household level LPG purchases for three months, at assigned sizes and prices.

The multiple price list was implemented as follows (see Online Supplement I for additional

[26] In the market, at the start of stage 1, the 3 and 14.5 kilogram cylinders cost GH¢ 145 and 245, respectively. A 2-burner stove cost around GH¢ 122. We implemented a deposit that did not depend on size to avoid confounding selection based on durable good costs with selection based on repeated expenditures.

[27] The variation in stage 1 was implemented as a source of exogenous variation in cash on hand liquidity, but we observe little effect on cylinder size preferences. The lower deposit increased compliance with the take up decision (stove delivery success), so we used it for all offers in stage 2. We also randomly assigned households in stage 1 to either their closest or second closest depot to generate variation in transaction costs; this too was dropped in stage 2 in favor of assigning everyone to their closest depot.

[28] Allowing participants to switch sizes at the end of the exchange window ensured that choices were based on preferences over sizes at study prices, rather than on the long run value of the asset or preferences over sizes at future market prices.

[29] The fact that households could renege on their initial cylinder choices at the time of delivery (by not being home, for example) has the potential to interfere with the study design. In both stages, small cylinders have a slightly higher likelihood of delivery failure. The randomly assigned LPG price (stage 1) and treatment (stage 2) are statistically unrelated to delivery failures (see Table A.1). In our main results, we show robustness to including or excluding households that reneged on their choices.

detail). Each participant completed three different MPLs: Small cylinder vs Nothing (MPL A), Large cylinder vs Nothing (MPL B) and Small cylinder vs Large cylinder vs Nothing (MPL C).

MPLs A and B each included eight binary choices between one of the cylinders and nothing, where the choices varied in the price of a cylinder exchange at the depot. The first four choices were the same for all participants; responses determined values for a second set of four choices, which used smaller intervals for a more precise measure of WTP. Choices in MPLs A and B determined the content of MPL C, which elicited preferences across three options: small, large, and nothing. Prices of the large cylinder exchange were anchored at 80-100% of the maximum WTP from MPL B and the exchange price of the small cylinder varied around the maximum WTP from MPL A. MPL C included only four rows.

After the subject made all choices in all three MPLs, one row was drawn for implementation. This introduces random variation in the cylinder size and price combination, conditional on choices. When combined with the exchange data, this variation allows us to estimate demand at a wide range of price pairs, accounting for selection.

We observe a high degree of non-switching behavior on the extensive margin: 43% of participants prefer a cylinder over nothing in all choices. This is not altogether surprising; the price variation in the MPL was constrained by the market price of LPG since filling stations continued to operate during the LPG exchange phase of the project.

Stage 2 choices Data for stage 2 are considerably simpler than for stage 1. Cylinder selection decisions comprise a single choice between the large and the small cylinder, where the prices differ by treatment, but other aspects of the choice do not.

LPG cylinder exchanges Households that acquired a cylinder could access LPG through their assigned exchange depot at their assigned price for a period of three months. Depots were supplied by our LPG partner, Andev, and consisted of a metal cage where both full and empty cylinders were stored. Depot managers completed exchanges, collected payments, tracked inventory and liaised with the study team. Managers used tablets to record purchase details, including participant study ID, cylinder ID, price and time of the transaction.

During the three month study window (per stage), households could make as many exchanges as they wished, paying cash to a depot manager. They were, however, limited to exchanging a single cylinder at a time and to the cylinder size that they selected during enrollment.[30] Households were not explicitly prohibited from using study cylinders to buy LPG at filling stations (at market price) instead of using the exchange depots. To measure the extent of refilling – unmeasured demand – we conducted random spot checks on participants who had not yet exchanged a cylinder (see Section 5.3).

LPG prices proved to be very volatile during implementation of stage 2. The design of the optimal price menu relied on prices used in stage 1. By the time stage 2 launched, in February 2023, LPG prices had risen by around 25%. At the start of stage 2, we therefore inflate the prices from stage 1 to account for both real and nominal price increases (see Online Supplement III for additional detail on our adjustments and implications). Alternative strategies for adjusting the subsidy budget have little effect on the optimal contract. After stage 2 launched, prices fell to the point that the price of LPG in the large cylinder in the treatment arm exceeded the price in the market. To mitigate the risk of purchases outside of the exchange depots, we deflated prices back to stage 1 levels, following prices in the LPG market. In our analysis, these changes primarily affect the subsidy budget, so we show robustness to several alternative cost calculations. Our main analysis uses "nominal" prices as experienced by participants.[31] In Section 6, we discuss how a policy maker could accommodate price volatility and other time varying determinants of demand and costs.

4.3 Summary statistics and randomization

Over a half of households in the sample are female headed and the mean age of the household head is around 49 years. Households contain an average of 4.9 individuals. These characteristics are all similar across the two stages (see Table A.2). Wealth and wealth-related

[30] These implementation details increase our ability to generate separation in the stage 2 treatment group. Allowing households to exchange multiple cylinders simultaneously would make it more attractive to select the smaller cylinder and receive the larger subsidy. We discuss the potential to place restrictions on cylinder ownership in Section 6.

[31] We multiply stage 1 prices by 1.338 and apply to all transactions from February to June, then return to stage 1 prices for all stage 2 purchases from July to September. This affects both the prices charged to consumers and the calculation of the subsidy expenditure per exchange.

characteristics (liquidity constraints and baseline charcoal purchasing patterns) do, however, differ between the two stages, with a slightly higher average wealth in the second stage due to the change in how strata are defined between stages (see footnote 25).

Next, we examine randomization outcomes from stages 1 and 2. In stage 1, cylinder size and LPG prices are assigned randomly through the MPL. By design, the MPL choice set updated dynamically. We thus focus our balance test on the subset of choices viewed by all participants, which we refer to as the "static" choices. We see imbalance in one pairwise comparison of choices in MPL A, but no other significant differences (pairwise or joint) in the randomization outcomes in stage 1 (see Figure A.1).[32] In stage 2, households were randomly assigned to treatment (non-linear price menu) and control (linear price menu), stratified by LPG exchange depot and wealth. Conditional on strata fixed effects, all differences in means are small and insignificant (see Table A.3).

5 Design, then test

5.1 Stage 1: Design

In the first stage of the experiment, we design the optimal non-linear price menu following six steps. Notation follows that of the theoretical model, except that, from here on, we denote $p(q_S)$, the exchange price for the small cylinder, as p_S; and $p(q_L)$ as p_L.

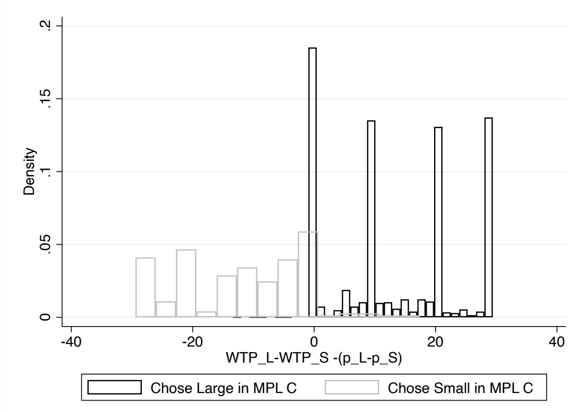
Step 1: Compute the sufficient statistic for each household. In the first step, we use the MPL data to approximate the difference in willingness to pay across cylinder sizes, which is the sufficient statistic for selection, $\tilde{\theta}_i = p(\theta_i, q_L) - p(\theta_i, q_S)$, for each household in the data. We denote this approximation as $\hat{\tilde{\theta}}_i$. Specifically, when we observe switching in MPLs A and B and/or MPL C, the WTP is interval-identified: $p(\theta_i, q_S)$ from MPL A, $p(\theta_i, q_L)$ from MPL B and $p(\theta_i, q_L) - p(\theta_i, q_S)$ from MPL C. In other words, we have two separate ways of identifying $\tilde{\theta}_i$. This redundancy is important in cases when there is non-switching

[32]Figure A.2 shows the frequency distribution of all prices and cylinder sizes, which is a function of both the randomized MPL draws and participant preferences. We conduct robustness checks on the demand elasticities resulting from the price variation to assess the effect of selection in the MPL.

behavior in either MPL A or MPL B, which leaves $p(\theta_i, q_S)$ or $p(\theta_i, q_L)$ unidentified. In these cases we approximate $\tilde{\theta}_i$ using data from MPL C. See Appendix I.3 for additional detail.

This procedure yields $\hat{\tilde{\theta}}_i$ for all observations, including those with never switching behavior or choices that are inconsistent across MPL modules.[33] Figure 1 shows the distribution of $\hat{\tilde{\theta}} - (p_L - p_S)$, or the approximation of $p(\theta_i, q_L) - p(\theta_i, q_S) - (p_L - p_S)$. This also represents

Figure 1: $\hat{\tilde{\theta}} - (p_L - p_S)$ and MPL choices



Note: Horizontal axis is the surplus from choosing the large cylinder size, or $\hat{\tilde{\theta}} - (p_L - p_S)$. Black bars are participants who chose the large cylinder in MPL C at the relevant price difference; gray bars are participants who chose the small cylinder.

the relative surplus associated with choosing the large cylinder for different price-type combinations in MPL C. The figure shows a clear separation: most large cylinder choices are associated with a positive surplus and small cylinder choices with a negative surplus. Given that we use MPL A and B to proxy $p(\theta_i, q_L) - p(\theta_i, q_S)$ in cases where $p(\theta_i, q_L) - p(\theta_i, q_S)$ from MPL C is also available, this result is not mechanical.

All else equal, households who gravitate towards large cylinder (have a higher $\hat{\tilde{\theta}}$) are younger and wealthier, less liquidity constrained, and have a lower baseline likelihood of purchasing charcoal in small quantities (see Table A.4). These patterns are consistent with the theoretical model's heterogeneity parameters, however – and more importantly – even with detailed survey data and geographic fixed effects, we can explain only around 10% of the variation in the WTP responses as measured by the R-squared from an OLS regression.[34]

[33] We omit from the analysis 5 households for which the WTP difference is not available from MPLs A and B and who present odd switching behavior in MPL C. Jack et al. (2022) provide a more general discussion of diagnosing and accommodating subject errors and truncated values in MPL data, consistent with the approaches adopted here.

[34] We have also explored prediction using a random forest model, which also provides limited explanatory

This limits the feasibility of targeting on observables using, for example, a two-stage approach similar to Johnson and Lipscomb (2022). It also suggests that a fully structural approach, which would rely on measuring different sources of heterogeneity, would perform poorly relative to our sufficient statistic approach.

Step 2: Estimate logit for choice of size. We observe 31 cases where choices are inconsistent either within or between MPLs. To allow for errors, and rationalize these cases, we add a random disturbance to the true indirect utility. We use a logit model to approximate the probability that a given individual chooses the large cylinder, conditional on his or her sufficient statistic for selection type, $\tilde{\theta}$, and the price menu she faces, (p_L, p_S). We fit this logit model to the observed choices in MPL C, under the assumption that the marginal utility of income is constant. The probability of choosing the large cylinder can then be written as:

$$\Pr\left(v(p_L;\theta,q_L) - v(p_S;\theta,q_S) + \varepsilon_L - \varepsilon_S > 0|\hat{\tilde{\theta}}_i\right) = \frac{\exp\left(\gamma\left(\hat{\tilde{\theta}}_i - (p_L - p_S)\right)\right)}{1+\exp\left(\gamma\left(\hat{\tilde{\theta}}_i - (p_L - p_S)\right)\right)} \tag{10}$$

where $\gamma = \frac{\partial v(p(q);\theta,q))}{\partial p(q)}\Big|_{q_S} \frac{1}{\sigma_\varepsilon}$.

Step 3: For a grid of potential price pairs, create selection weight, $\hat{w}_i$. We create selection weights to reflect the probability of choosing the large cylinder as a function of $\hat{\tilde{\theta}}_i$ over a grid of price pairs. This generates a smoothed predicted selection as $(p_L - p_S)$ varies.

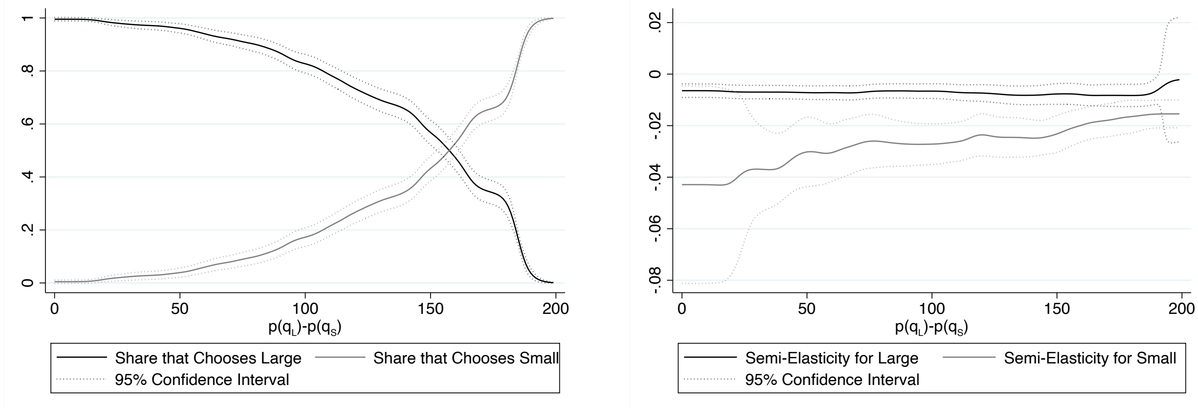
Specifically, using the estimate for γ that accounts for deviations in the indirect utility over time, and the formula in (10), we create weights for each observation in the sample, for each value of $(p_L - p_S)$ in the grid:

$$\hat{w}_i\left(p_L - p_S\right) = \frac{\exp\left(\hat{\gamma}\left(\hat{\tilde{\theta}}_i - (p_L - p_S)\right)\right)}{1+\exp\left(\hat{\gamma}\left(\hat{\tilde{\theta}}_i - (p_L - p_S)\right)\right)}.$$

Figure 2a shows the selection into sizes as a function of the price difference $(p_L - p_S)$ by averaging (5.1) over the whole sample.

power (see Figure A.4).

Figure 2: Selection and elasticities



(a) Weighted selection probabilities (b) Predicted semi-elasticities of demand

Note: Panel A - Predicted selection probabilities as a function of prices. The vertical axis is the share of the sample choosing each cylinder size. Panel B - Predicted semi-elasticities of demand, incorporating selection. The vertical axis is the semi-elasticity for each cylinder size, evaluated for the sample that selects into each cylinder size at the corresponding price difference on the horizontal axis. 95% confidence intervals computed using bootstrapped standard errors from 1,000 random samples.

As the exchange price of the large cylinder grows, relative to the small (i.e., moving right on the horizontal axis), a smaller share of the sample selects into the large cylinder.

Step 4: Use WLS to estimate LPG demand at each price pair. Next, we use these weights to incorporate selection into predicted demand at each price pair, assuming a quadratic fit, consistent with our data (see Figure A.3).[35]

Denote the number of observations for which we observe x_{Si} as N_S and the number of observations for which we observe x_{Li} as N_L. We index each value of $(p_L - p_S)$ in the grid by k. We then estimate a different set of demand parameters for the large cylinder for each k, $\boldsymbol{\beta}_k^L = [\beta_{0k}^L, \beta_{1k}^L, \beta_{2k}^L]$, by weighted least squares (WLS):

$$\min_{\boldsymbol{\beta}_k^L} \sum_{i=1}^{N_L} \hat{w}_{ik}(x_{Li} - \beta_{0k}^L - \beta_{1k}^L p_{Li} - \beta_{2k}^L p_{Li}^2)^2.$$

We use an analogous procedure to estimate the demand parameters of the small cylinder,

[35] Table A.6 decomposes the demand elasticities, inclusive of selection, into the extensive and intensive margins. We test whether price sensitivity in LPG demand depends on whether the price was assigned using endogenous prices (prices from the dynamic part of the MPL). Neither average price sensitivity nor size-specific price sensitivity depends on whether the participant received their cylinder through the static or dynamic portion of the MPL. This implies that the additional selection imposed by the dynamic design does not meaningfully affect the demand estimates used in the menu design.

$\boldsymbol{\beta}_k^S$, where weights are given by $1 - \hat{w}_{ik}$.

WLS should deliver a consistent estimate of the average of price coefficients across the selected sample (Solon et al., 2015).[36] The average predicted demand for cylinder exchanges of size $Z \in \{S, L\}$, reflecting the selection weights, is then

$$\hat{x}_k^Z(p_Z) = \hat{\beta}_{0k}^Z + \hat{\beta}_{1k}^Z p_Z + \hat{\beta}_{2k}^Z p_Z^2$$

The first thing to notice in Figure 2b is that semi-elasticities of demand are lower (more price sensitive) for the small cylinder than for the large. This is true for all price differences, but particularly so when the exchange price difference between them is small, i.e., a higher price per kg in the small cylinder. The small number of participants who still choose the small cylinder are very price sensitive. As $(p_L - p_S)$ increases (moves right along the horizontal axis), the share of the sample choosing the large cylinder declines and becomes more selected. The relatively flat line plotting the semi-elasticity among those selecting into the large suggests that selection has a relatively minor effect on semi-elasticities of demand for the large cylinder. On the other hand, moving right along the horizontal axis moves a larger share of less price sensitive people into preferring the small cylinder, decreasing price sensitivity, on average, among those selecting into the small. This indicates that subsidizing the small cylinder – up to a point – may increase demand because of the greater price sensitivity among small cylinder types for most price pairs. The total quantity demanded at each price pair also depends on how cylinder size specific non-monetary costs affect demand levels.

Step 5: Evaluate the policy maker's objective function and budget constraint at each price pair. We evaluate the policy maker's objective function (aggregate demand) for each point of the $[(p_L - p_S), p_S]$ grid as

$$\left(\left(\frac{1}{N}\sum_{i=1}^{N}\hat{w}_{ik}\right) \times \hat{x}_k^L(p_L) \times q_L\right) + \left(\left(\frac{1}{N}\sum_{i=1}^{N}(1 - \hat{w}_{ik})\right) \times \hat{x}_k^S(p_S) \times q_S\right) \tag{11}$$

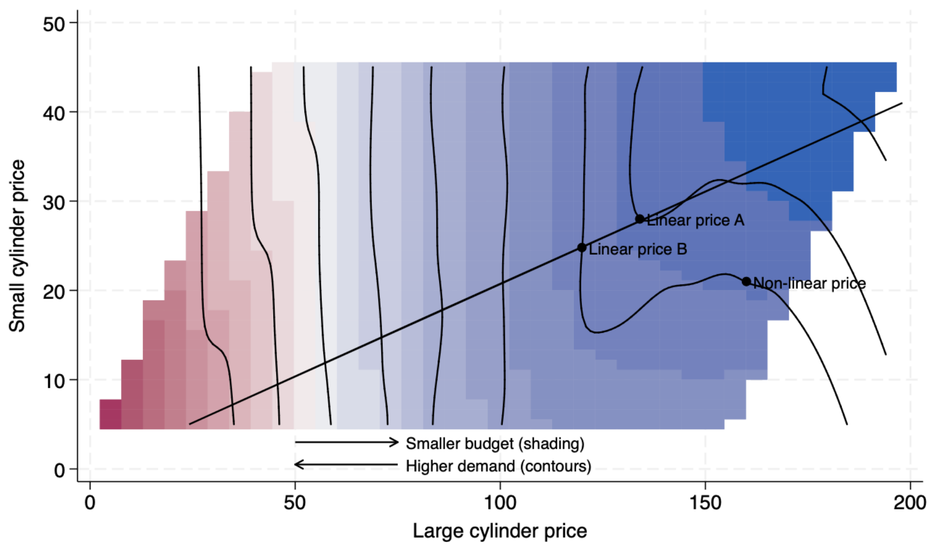
Then, we evaluate the policy maker's budget at each value of the $[(p_L - p_S), p_S]$ grid by

[36] Solon et al. (2015) specify that a condition for consistency is that the variance of the independent variables is constant across selected samples. In principle, this should be the case given that prices were randomly assigned. We examine this in Table A.7 and discuss implications in Section 5.

substituting q_L with $(q_L r - p_L)$ and q_S with $(q_S r - p_S)$ in (11), where r is the cost per kilogram of LPG for the policy maker.

Figure 3 shows how cylinder exchange prices affect the subsidy budget and objective function. As prices increase (move away from the origin), the budget (heat map) and demand (contour lines) both decrease, on average. The figure shows that (a) at high budgets (low

Figure 3: Welfare and budget by price pair



Note: Shaded heatmap shows subsidy budget, which falls as prices move away from the origin. Contour lines show the objective function – total demand – which is decreasing as contours move outward from the origin. The straight line shows the linear price (equal price per kg) points. Linear pricing A and B show alternative counterfactuals for non-linear pricing that hold fixed the budget and demand, respectively.

prices) the policy maker typically cannot do better than the linear contract (contours and budget shading are parallel) and (b) with smaller budgets, there exist some points where a non-linear menu can deliver the same budget, but higher predicted demand. The small budget (modest demand increases) scenario is more relevant for our specific context.

Step 6: Choose non-linear price menu and linear prices that deliver same budget level and calculate sample size needed to detect differences in demand The non-linear pricing contract we implement is labeled in Figure 3's lower right quadrant. Two counterfactual linear price menus are also shown. Linear pricing A shows a budget-neutral alternative to our chosen non-linear pricing menu. At the same budget, it is predicted to deliver lower demand (contour lies to the right). This is the linear price "control" that we implement in stage 2. Linear pricing B shows a higher-budget alternative, which delivers the same level of demand (same contour).[37]

Table 2: Stage 2 study arms and menus

Study Arm	Cylinder	Exchange Price	Kg Price
Control (Linear Price)	3kg	GH¢ 28	GH¢ 9.3
Control (Linear Price)	14.5kg	GH¢ 134	GH¢ 9.3
Treatment (Non-Linear Price)	3kg	GH¢ 21	GH¢ 7.0
Treatment (Non-Linear Price)	14.5kg	GH¢ 160	GH¢ 11.0

Note: Unsubsidized per kg price = GH¢ 13.38; 1 USD = 8 GH¢.

Our chosen menu corresponds to a budget of GH¢ 14 per participant.[38] The predicted mean difference in LPG demand per household per month is 0.59 kg. Power calculations also depend on the variance of demand, accounting for selection, under each treatment arm. Online Supplement II provides the details on how we estimate the variance using an algorithm similar to that used for demand. To detect a 0.59 kg difference in demand requires a sample size of 828 households. We overshoot this target with 915 households in our stage 2 sample to ensure adequate power.

5.2 Stage 2: Test

Stage 2 tests the effect of a non-linear price menu for LPG on demand, relative to linear pricing designed to be budget neutral in expectation. From the policy maker's perspective, the objective of the separating contract is to maximize LPG demand, conditional on a fixed budget. Our primary outcomes are therefore (1) LPG demand per month and (2) subsidy budget per month, which is a function of demand and the exchange prices.[39]

Since the policy maker is interested in both the costs and benefits of subsidies, and since

[37]Results from stage 1 predict that non-linear pricing increases demand by 17.5% or reduces the budget by 29.4% relative to linear pricing counterfactuals A and B, respectively.

[38]To facilitate exchanges using cash, exchange prices were set to whole numbers, resulting in a small difference in the kg price across cylinder sizes in the linear price control: 9.33 vs 9.24 for the small and large cylinder, respectively.

[39]In our demand measure, we include all LPG purchased during the exchange window, regardless of whether it was consumed. This implicitly advantages the demand outcome for large cylinder owners since each exchange is approximately 5x the quantity of the small cylinder, leaving large cylinder owners with a higher stock of LPG (conditional on exchange date) at the end of their eligibility window.

differences in implementation, sample or time may break the budget-neutrality of the design, we define two joint hypothesis tests for the effect of non-linear price treatment relative to the linear price control:

	Budget	Demand
H1:	$= 0$	> 0
H2:	< 0	≥ 0

Specifically, we predict that non-linear pricing will increase cost effectiveness either through (H1) an increase in demand, holding the budget fixed, or (H2) a decrease in the budget, with the same or greater demand.[40] Table 3 implements the hypotheses tests in an OLS regression specification, with a single cross sectional observation per participant.[41] Columns 1-3 show

Table 3: Stage 2 subsidy budget and LPG demand

	Subsidy budget (GH¢/month)			LPG demand (kg/month)		
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment effect	-3.333***	-3.391***	-3.209***	-0.153	-0.166	-0.105
	(0.837)	(0.824)	(0.824)	(0.259)	(0.254)	(0.253)
Control mean	11.585	11.585	11.585	3.089	3.089	3.089
Strata FE		×	×		×	×
Surveyor FE			×			×
Enrollment week FE			×			×
Observations	915	915	915	915	915	915

Note: Columns 1-3 show treatment effect on subsidy per month; columns 4-6 on LPG demand per month. Columns 2 and 5 include strata fixed effects; columns 3 and 6 include strata, surveyor and enrollment week fixed effects. * p<0.1, ** p<0.05, *** p<0.01

the effect of treatment assignment on the per person-month subsidy budget and columns 4-6 show the effect on kilograms of LPG purchased. We observe outcomes consistent with H2: similar levels of demand in treatment and control, but a public budget 29% smaller when

[40] Our predictions follow our pre-analysis plan and are asymmetric in the strictness of equality. In particular, in H1 we test the stage 1 model predictions, while H2 is more flexible. We implement these tests with and without strata fixed effects and controls for implementation details.

[41] We show results that include all households regardless of starter kit delivery outcomes (see Table A.1); delivery failures are coded as zeros both for demand and the subsidy budget. This is in contrast to the menu design in stage 1, which omits delivery failures. Table A.5 shows that results are largely unaffected by whether these households are in the sample or not.

prices are differentiated by quantity (column 1, no controls).[42]

Next, we analyze treatment effects on LPG demand by cylinder size, which is a function of both selection and demand. Stage 1 results predict that households that select into the small cylinder will be more price sensitive regardless of treatment. The selection response to the non-linear prices will dampen this effect – high transaction costs of households on the margin will lead to lower demand frequencies – but the overall effect will still be positive. We do not have a prediction on demand for the large cylinder. The price effect will lower demand while the selection effect will increase it. The net effect is ambiguous. As shown in Table 4, our empirical results are consistent with these predictions: LPG demand among small cylinder owners is around 50% higher in treatment than control; among large cylinder owners, demand is statistically unaffected by treatment. Also consistent with stage 1 predictions, treatment

Table 4: Stage 2 demand by cylinder size and size selection

	LPG demand (kg/month)				Small cylinder	
	(1)	(2)	(3)	(4)	(5)	(6)
Sample:	Small	Small	Large	Large	All	All
Treatment effect	0.532**	0.548**	-0.310	-0.171	0.054*	0.055*
	(0.218)	(0.238)	(0.337)	(0.339)	(0.030)	(0.030)
Control mean	1.105	1.105	3.854	3.854	0.266	0.266
Strata FE	×	×	×	×	×	×
Surveyor FE		×		×		×
Survey week FE		×		×		×
Observations	267	267	641	641	915	915

Note: Columns 1-4 show results from linear regression of LPG demand per month on indicator for treatment assignment. Columns 1-2 show results for sample of households with small cylinder, columns 3-4 for sample with large cylinder. Households that chose no cylinder (N=7) are omitted from columns 1-4. Columns 5-6 show results from linear regression of indicator for small cylinder take up on indicator for treatment assignment on full sample. All columns include stage 2 strata fixed effects, even columns also include surveyor and survey week fixed effects. * p<0.1, ** p<0.05, *** p<0.01

increases take up of the small cylinder is around 5.5 percentage points, or 21%.

While the results support our main hypothesis that non-linear pricing can increase the cost effectiveness of subsidies, the pattern of results – a decrease in public spending, with

[42] The treatment effect on demand is also negative but much smaller in magnitude (<5% lower than the control group mean) with a 95% confidence interval of (-0.66, 0.36) kg in column 4.

no change in demand – deviates from the stage 1 predictions. We compare the predictions with the results in Table 5, to better understand the discrepancies.[43]

Table 5: Model predictions vs stage 2 results

	Model Predictions		Empirical Results	
	(1) Control	(2) Treatment	(3) Control	(4) Treatment
Subsidy per month	14.000	14.000	13.309	9.789
	(1.788)	(1.216)	(0.978)	(0.666)
Small cylinder share	0.195	0.340	0.266	0.318
	(0.018)	(0.022)	(0.021)	(0.022)
Small exchanges per month	0.789	1.206	0.380	0.610
	(0.121)	(0.127)	(0.048)	(0.063)
Large exchanges per month	0.248	0.284	0.272	0.262
	(0.036)	(0.038)	(0.021)	(0.023)

Note: Columns 1-2 show model predictions and columns 3-4 show results from stage 2 data. Standard errors are bootstrapped in columns 1-2. See footnote 43 for details on the sample.

By design, stage 1's predicted subsidy per household per month is the same in treatment and control. The subsidy budget is lower in the stage 2 results, and to a greater extent for the treatment group. This is our main result, as shown in Table 3. The rest of Table 5 helps explain the discrepancy between stage 1 predictions and stage 2 results. First, in the predictions, non-linear pricing increases demand both by increasing take up of the small cylinder and increasing demand among those who take up the small cylinder. In the results, the effect on take up is muted because of higher than predicted take up of the small cylinder in the control. For the small cylinder, observed demand is proportionately lower – by 48% and 50% in control and treatment, respectively – than predicted demand. Consequently, the predicted demand boost from the small cylinder is largely undone. This, however, substantially lowers the subsidy budget in the treatment group without negatively impacting demand. Large cylinder exchanges per month more closely match predictions, though the slight demand boost predicted in large cylinder exchanges is not observed in the results from stage 2.

[43]To facilitate a cleaner comparison, Table 5 includes only households with a successful delivery and stage 2 exchange data pre-price adjustment, i.e., imposes the same sample restriction as used in the stage 1 menu design. See also Table A.5.

While we cannot determine the precise cause of the divergence between predicted and observed take up and demand in the small cylinder, we discuss several possibilities. First, inflation between stages 1 and 2 may have increased demand for the small cylinder in the control group and depressed the demand of low income consumers, who are more likely to select the small cylinder. Second, sample differences between the two stages (see Section 4.2 and Table A.2) may also have contributed to the divergence, along with differences in how households chose their cylinder size (i.e., in an MPL in stage 1 and in a take-it-or-leave-it contract in stage 2). Finally, as noted in footnote 36, consistency of the demand parameters via WLS requires that the variance of prices is constant across types. However, Table A.7 shows higher variance of prices for higher types, which might contribute to a divergence between the stage 1 predictions and stage 2 results.

The public budget savings in the treatment group would support expanding program coverage by roughly 30%, without increasing the LPG subsidy budget. In other words, "demand neutrality" can be converted into budget neutrality by expanding program coverage. However, without a pure control group that received subsidized durables, but not subsidized LPG, we cannot assess the effect of program expansion on overall LPG demand.

5.3 Additional results and confounds

In this section, we consider several additional impacts and implications of non-linear pricing.

Who gets subsidized? In our theoretical model, saving costs and transaction costs are sufficient for separation. However, our empirical approach abstracts from observable characteristics. Following our pre-analysis plan, we construct three indices, proxying for time costs (part of transaction costs), savings or liquidity costs, and small purchases/wealth, along with a single aggregate index that combines all three. All are scored such that a higher index value is associated with a theoretically lower θ (low opportunity cost of time, high savings cost and a preference for small purchases). Index construction is described in greater detail in Online Supplement I.[44]

[44] We deviate from the pre-analysis plan in one way: we exclude access to transport from the opportunity cost of time, since it has ambiguous effects: it reduces the hassle cost of repeated trips but also makes transportation of bulky large cylinders less difficult. Table A.9 shows results for each pre-specified

Table 6: Stage 2 targeting on observables

Index:	(1) Willing to wait	(2) Savings cost	(3) Small purchases	(4) Aggregate
Panel A: Outcome: Small cylinder				
Treatment effect	0.050*	0.051*	0.044	0.049
	(0.030)	(0.030)	(0.031)	(0.030)
Index	-0.020	0.106***	0.070***	0.082***
	(0.021)	(0.020)	(0.022)	(0.020)
Treatment × Index	0.067**	-0.043	-0.041	-0.013
	(0.030)	(0.039)	(0.031)	(0.030)
Observations	915	915	915	915
Panel B: Outcome: Subsidy budget (GH¢/month)				
Treatment effect	-3.353***	-3.319***	-3.012***	-3.157***
	(0.836)	(0.831)	(0.806)	(0.822)
Index	0.497	-1.401**	-2.745***	-1.942***
	(0.646)	(0.606)	(0.720)	(0.663)
Treatment × Index	0.225	0.691	1.820**	1.506*
	(0.784)	(0.714)	(0.866)	(0.795)
Observations	915	915	915	915

Note: Columns 1-4 show treatment effects for all households. Each column controls for index specified in column header and its interaction with treatment predictor. * p<0.1, ** p<0.05, *** p<0.01

The results in Table 6 align closely with the theoretical model presented in Section 3.1. First, recall that most households prefer the large cylinder, even with the cost per kilogram of LPG is higher. This is consistent with the transaction cost $D(\theta)$ being positive on average. Recall that Proposition 1 requires heterogeneity in either $A(\theta)$ or $D(\theta)$ but not both. We observe that households with a higher opportunity of time are not more likely to take up the small cylinder in the control group (panel A, column 1), but households with higher saving costs are (column 2). This is consistent with heterogeneity in saving costs $A(\theta)$ rather than transaction costs $D(\theta)$ mattering in this context. Then, for separation, we need that both savings and transaction costs are positive. In other words: all households would prefer to avoid the transaction costs associated with the small cylinder, but those with particularly high saving costs cannot. In the non-linear price menu, the price in the large cylinder goes

characteristic.

up, increasing the savings cost and pushing some households who otherwise would have taken up the large cylinder to choose the small. The result in panel A, column 1 suggests that these households tend to have relatively low opportunity cost of time, though other transaction costs may drive down their purchases when they switch to the small cylinder size.

Panel B of of Table 6 tests whether the non-linear price menu targets more of the subsidy budget toward households with these characteristics. The coefficient on the index term in the control group is negative for three out of the four indices, consistent with these types of households having lower overall demand for LPG (and therefore receiving a smaller share of the linear per kg subsidy). However, the non-linear price menu increases their share of the pie: the interaction terms are all positive and the small purchase index interaction is significant at $p = 0.053$.

Inframarginal subsidies and impacts on charcoal use The primary justifications for subsidizing LPG use – reductions in particulate and greenhouse gas emissions – depend on displacement of charcoal use. With neither a treatment effect on LPG demand, nor a pure control group that received subsidized durables but no LPG subsidy, this is difficult. Here, we present what we do observe about counterfactual LPG and charcoal use.

We start with counterfactual LPG demand. By construction, the study sample consists of households that did not use LPG as a primary cooking fuel at baseline, which helps mitigate concerns about inframarginal subsidies. We consider three different approaches to assessing counterfactual LPG demand. First, we use the empirical model with data from stage 1 to simulate demand at zero subsidy. Based on observed demand and the quadratic functional form used in the stage 1 model, demand is predicted to be zero in the absence of any LPG subsidy even if durables were subsidized, i.e., none of the subsidies were inframarginal. Second, we use survey data at baseline and endline to measure stated use of LPG before and after the introduction of subsidies. At baseline, 23% of the sample reported using any LPG in the past three months. At endline, 61% of the sample report using any LPG in the past three months (which coincides with the exchange window).[45] Without information on the intensive margin of demand for users who already had LPG durables at baseline, these

[45] This latter number is higher than the share of the sample that made any exchanges at the study depots because some households presumably used their own cylinders outside of the study.

survey responses offer a conservative upper bound of 38 to 80% inframarginal subsidies. Third, in a separate project that worked with a smaller sample of households from the same sampling frame, households were randomly assigned to the non-linear price treatment (N=69) or a control (N=80) that received neither subsidized durables nor subsidized LPG (Daouda et al. 2024). At endline, only 10% (1.3%) of the control group reported any (primary) LPG use versus 94% (74%) of the treatment group. This suggests a much lower degree of inframarginality (10.6%) than the before-after comparison.

Next, we examine charcoal displacement, starting with a self-reported before-after comparisons. Between baseline and endline, average household monthly spending on charcoal fell from around 400 GH¢/ to 127 GH¢, a drop of more than 50% in spite of fairly stable prices. Given that the average amount of LPG purchased under the program is less than 50% of a typical household's cooking needs, the implied substitution rate is >1. Alternatively, we can quantify how much displacement is needed to justify subsidizing LPG, using the calculation of the external costs of greenhouse gas emissions presented in Section 2. We calculate the cost per ton of avoided CO_2e under the linear price control and non-linear price treatment, for different charcoal displacement rates (see Figure A.5). If displacement is assumed to be 1:1, the subsidies in both treatments are highly cost effective at 10 and 14 USD/ton under non-linear and linear pricing, respectively. As the displacement rate falls, the social cost of carbon (i.e., benefit of avoided emission) required to justify the subsidies (i.e., cost of avoided emissions) increases, and targeting using non-linear prices can be justified under a lower charcoal displacement rate or a lower social cost of carbon. Net emissions are positive for displacement rates less than 0.22 (i.e., each kg of LPG purchased displaces less than 0.22 energy equivalent units of charcoal). This degree of fuel stacking is highly unlikely at the subsidy levels that we test, i.e., the income effect from subsidies is very small.

Potential confounds and robustness checks We next discuss potential confounds and examine robustness to alternative assumptions. Online Supplement III provides additional details.

First, our calculation of the subsidy is based on the nominal prices paid by study participants, which closely follow LPG prices at the pump (see Section 4.2). We consider three alternative calculations of the subsidy budget (see Table A.10): LPG market retail prices at the time of purchase, the time varying refinery price, and inflation-adjusted nominal prices. Treatment effects are of similar or larger magnitudes with these alternative calculations. Second, our analysis assumes that we observe all LPG demand during the three month exchange window. Households were not prohibited from visiting filling stations (prohibitions would have been unenforceable). We conduct random spot checks to assess filling outside of the exchange depots. In 79 unannounced visits, we observe no evidence of unmeasured LPG demand (Figure II). Third, to check adherence to study protocols, we conducted "mystery shopper" visits to the exchange depots. We find no cases in which depot managers allowed someone to purchase at a price that differed from their assigned price.

6 Discussion and conclusion

6.1 Cost effectiveness and scalability

Our main results offer a proof of concept that non-linear pricing can make subsidies more cost effective, however, scale up requires several additional considerations.

Supply costs, durables and limits on exchange Several implementation details pose potential departures from a scaled up version of non-linear pricing. First, we suppressed size-specific differences in supply costs. One argument for bulk discounts is that supplying goods in smaller packages is more expensive per unit. In our setting, the cost per kg to supply gas via a small cylinder is nearly 2.5 times the cost to supply it in a large cylinder. If we add this to the policy maker's subsidy budget, the total subsidy cost per household per month in the non-linear price treatment increases from 8.3 to 14.4 GH¢ per household per month, but the magnitude of the treatment effect changes only slightly, from -3.21 GH¢ (column 3 of Table 3) to -2.70 GH¢. We calculate that the supply cost difference between a small and large cylinder would need to more than triple to eliminate the budget savings from non-linear pricing.

Second, to avoid selection based on differential cost of durables, we held the starter kit deposit price fixed across cylinder sizes.[46] Subsidized durables may have contributed to high participation in the exchange program (see footnote 26). Asking households pay market prices would presumably change selection, both into LPG and across cylinder sizes. Analysis of these effects is beyond the scope of this study; however, we note that cost effectiveness gains from non-linear pricing may be undone by large fixed costs, which may reduce heterogeneity in size preferences among those who take up any LPG.

Third, we limited participants to a single project cylinder, and required that their previous cylinder was returned (empty) in order to acquire a new (full) cylinder. Thus, households could not exchange multiple cylinders at the same time.[47] In a scaled up market implementation of non-linear pricing, one challenge for maintaining separation across cylinder sizes arises if those with a preference for larger purchases to take advantage of the small cylinder subsidies by exchanging multiple small cylinders at a time. A straightforward approach to minimizing this behavior is a limit on the quantity of simultaneous purchases (perhaps on the small cylinder only). These types of "one offer per customer" restrictions are common for other discounts. While customers could return for multiple sequential purchases in a short time period, the added hassle cost would reduce the incentive to pool on the more heavily subsidized small cylinder.[48]

Policy implementation Our two stage approach to designing optimal non-linear pricing is, in principle, scalable. The policy designer would need to collect experimental data on cylinder size preferences and price elasticities for a representative subset of the population of interest and plug these into the sufficient statistic-based price menu. These prices could then be rolled out to a much larger population. However, this approach presents at least two important challenges.

First, holding the population fixed, package size choices are likely to change over time for a variety of reasons, including financial access, income growth, transportation costs, and

[46] Recall that we did vary the deposit price randomly in stage 1. Lower deposit prices increased delivery success (see Table A.11) but had little effect on WTP in the MPL.

[47] Relative to other goods, this concern may be less important for LPG, where some of the hassle cost of small package sizes is associated with replacing a cylinder in the home.

[48] Alternatively, cylinders could be registered per household to limit ownership, such as many countries do with SIM cards; this, however, might reduce overall LPG use if households like to have a spare (full) cylinder available at home.

other factors. In addition, macroeconomic conditions – including inflation and LPG prices – affect both consumer prices and the subsidy cost. Addressing inflation, when LPG prices increase in lockstep with other inflation indicators, is straightforward. However, when LPG prices change, because of the global gas market or local policy, in ways that diverge from inflation, the subsidy cost may grow more or less quickly than nominal prices over time. This results in a new optimal price menu. If consumer demand is assumed to be unchanged, this can be identified using the same model and data as used to solve the policy maker's problem prior to the price change. Changes to consumer demand are more challenging and require new data inputs to solve for optimal prices.

Second, collecting experimental data to design optimal price menus has advantages in that the relevant variation is available by design. However, it requires on-the-ground infrastructure for implementation and is costly if the results have limited external validity over time or across populations. In principle, one could use non-experimental sources of variation in cylinder size selection and LPG demand. Recall that the sufficient statistic approach requires observing $p(\theta_i, q_L) - p(\theta_i, q_S)$. This may be identifiable with cross sectional or panel data from a market with naturally occurring price variation and observations of cylinder size choices. However, to separately identify selection and demand elasticities, additional variation – in, for example, the size-specific price of durables – may be necessary. We leave for future work the exercise of extending the menu design framework to accommodate non-experimental data. Alternatively, structural estimation of a demand model could be used to design a generalizable price menu, based on modeled relationships between observables, size preferences and demand. Our results caution against this approach. In practice, this mapping may be unstable over time or across populations. In our data, we also find that it provides low explanatory power even in-sample.

6.2 Welfare effects of non-linear pricing

As discussed in Section 3, the policy maker's objective function ignores: (1) consumer surplus, (2) the level and distribution of non-monetary costs, i.e., transaction costs and saving costs, (3) the deadweight loss (DWL) from subsidies, (4) the marginal cost of public funds and (5) external costs or benefits from pollution. We conduct an empirical back of the envelope exercise using data from stage 2 to gauge the magnitude of these effects (see Table A.12).[49]

First, the difference in consumer surplus among "stayers" (those who always choose small or large) can be approximated from the area under the demand curve for cylinders of each size and price menu. We add to this bounds on the non-monetary costs to "switchers" (those whose preferred cylinder size depends on treatment). The population weighted average change in consumer surplus is (-3.65, -3.42). Second, the DWL can be calculated as the complement of the consumer surplus result. Because the increase in subsidy goes to the relatively price sensitive group (small stayers), their DWL increases by more per household than the reduction for large cylinder types, where negative values correspond to an increase in the DWL of subsidies. The population weighted average change in DWL is (-0.56, 0.045).

Finally, we consider two scenarios for the external benefits of non-linear pricing. First, holding the population fixed, the policy maker saves 3.34 GH¢ in subsidy budget per household-month. At a marginal cost of public funds of 1.17 (Auriol and Warlters, 2012), this implies a public benefit of 7.23 GH¢ per household-month. Second, if instead the policy maker uses the savings to expand program coverage by 29%, the additional households covered under non-linear pricing provide additional greenhouse gas reductions, following the estimates in Sections 2. We use the lower end of the social cost of carbon ($ 51/ton) and assume full displacement of charcoal per kg of LPG purchased. This implies an 11.49 GH¢ external benefit from avoided climate damages per household-month.[50]

Combining the impact on consumer surplus, DWL and external benefits gives an overall welfare estimate under each scenario. Under scenario 1, this ranges from 3.02 to 3.86 GH¢ per household-month. Under scenario 2, it ranges from 6.06 to 7.14 GH¢ per household-month. Both calculations ignore the redistributional effect of targeting, which increases the subsidy allocation to relatively poor households (see Table 6). Finally, by restricting the pollution benefits to avoided climate damages, our estimates do not account for any health gains from improved air quality.

[49] The stage 1 model can also be used to calibrate consumer welfare, but given the divergence between stage 1 and 2 results, we focus on the stage 2 data here.

[50] Specifically, this assumes that each household under non-linear pricing provides 1.29 times the climate benefits as does a household under linear pricing. We also multiply the consumer surplus and deadweight loss results by 1.29 in this scenario.

6.3 Conclusion

We contribute to a large literature on the targeting of subsidies for social goals. By leveraging sources of heterogeneity in the population, we show that non-linear prices can improve cost effectiveness, increasing demand for a clean fuel per dollar spent. Our two-stage experiment to design and test non-linear pricing against a counterfactual of linear prices offers a proof of concept for our approach. While further refinement is needed to make the approach scalable, our results demonstrate both the potential for second degree price discrimination to increase welfare and the value of experimental validation of theoretical predictions of optimal contracts and mechanisms.

Our application is to clean fuels in Ghana, but our approach can be extended to a range of consumer goods and services where demand is socially valuable, including health products and other environmental goods. Whether the population exhibits the heterogeneity necessary to induce separation and support non-linear pricing is an empirical question. While our theoretical model establishes sufficient conditions for separation using very specific sources of heterogeneity, preferences over small versus large purchase quantities are (anecdotally) common and leveraging these for targeting subsidies can be both progressive and efficient. We leave as future work generalizing our model to derive the necessary conditions for separation under a more general structure, as well as extending our empirical framework to accommodate non-experimental data.

References

Alatas, V., A. Banerjee, R. Hanna, B. A. Olken, and J. Tobias (2012). Targeting the poor: evidence from a field experiment in Indonesia. *American Economic Review 102*(4), 1206–1240.

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

Alderman, H. and K. Lindert (1998). The potential and limitations of self-targeted food subsidies. *The World Bank Research Observer 13*, 213–229.

Armstrong, M. (2016). Nonlinear pricing. *Annual Review of Economics 8*, 583–614.

Attanasio, O. and E. Pastorino (2020). Nonlinear pricing in village economies. *Econometrica 88*, 207–263.

Auriol, E. and M. Warlters (2012). The marginal cost of public funds and tax reform in africa. *Journal of Development Economics 97*(1), 58–72.

Bailis, R., R. Drigo, A. Ghilardi, and O. Masera (2015). The carbon footprint of traditional woodfuels. *Nature Climate Change 5*, 266–272.

Baumol, W. J. and D. F. Bradford (1970). Optimal departures from marginal cost pricing. *American Economic Review 60*(3), 265–283.

Bensch, G. and J. Peters (2015). The intensive margin of technology adoption: Experimental evidence on improved cooking stoves in rural Senegal. *Journal of Health Economics 42*, 44–63.

Berkouwer, S. and J. Dean (2023). Private actions in the presence of externalities: The health impacts of reducing air pollution peaks but not ambient exposure. *Working Paper*.

Berkouwer, S. B. and J. T. Dean (2022). Credit, attention, and externalities in the adoption of energy efficient technologies by low-income households. *American Economic Review 112*, 3291–3330.

Blanchard, O., C. Gollier, and J. Tirole (2023). The portfolio of economic policies needed to fight climate change. *Annual Review of Economics 15*(1), 689–722.

Borenstein, S. (2012). The redistributional impact of nonlinear electricity pricing. *American Economic Journal: Economic Policy 4*, 56–90.

Brown, S. J. and D. S. Sibley (1986). *The theory of public utility pricing*. Cambridge University Press.

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

Cohen, J., P. Dupas, and S. Schaner (2015). Price subsidies, diagnostic tests, and targeting of malaria treatment: Evidence from a randomized controlled trial. *American Economic Review 105*, 609–645.

Dillon, B., J. De Weerdt, and T. O'Donoghue (2021). Paying more for less: Why don't households in Tanzania take advantage of bulk discounts? *The World Bank Economic Review 35*(1), 148–179.

Dizon-Ross, R. and A. Zucker (2023). Mechanism design for personalized policy: A field experiment incentivizing exercise. *Working Paper*.

Dubé, J.-P. and S. Misra (2023). Personalized pricing and consumer welfare. *Journal of Political Economy 131*(1), 131–189.

Duflo, E., R. Hanna, and S. P. Ryan (2012). Incentives work: Getting teachers to come to school. *American Economic Review 102*, 1241–1278.

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

Floess, E., A. Grieshop, E. Puzzolo, D. Pope, N. Leach, C. J. Smith, A. Gill-Wiehl, K. Landesman, and R. Bailis (2023). Scaling up gas and electric cooking in low-and middle-income countries: climate threat or mitigation strategy with co-benefits? *Environmental Research Letters 18*(3), 034010.

Fowlie, M. and R. Meeks (2021). The economics of energy efficiency in developing countries. *Review of Environmental Economics and Policy 15*(2), 238–260.

Ghana Statistical Service (2022). Ghana 2021 population and housing census - volume 3k: Housing characteristics. Technical report, Ghana Statistical Service.

Ghana Statistical Service (2022-2023). Price indices: Consumer price index bulletin. Accessed on January 23, 2023.

Gould, C. F., R. Bailis, K. Balakrishnan, M. Burke, S. Espinoza, S. Mehta, S. B. Schlesinger, J. R. Suarez-Lopez, and A. Pillarisetti (2024). In praise of cooking gas subsidies: transitional fuels to advance health and equity. *Environmental Research Letters*.

Hanna, R., E. Duflo, and M. Greenstone (2016). Up in smoke: The influence of household behavior on the long-run impact of improved cooking stoves. *American Economic Journal: Economic Policy 8*, 80–114.

Hanna, R. and B. A. Olken (2018). Universal basic incomes versus targeted transfers: Anti-poverty programs in developing countries. *Journal of Economic Perspectives 32*, 201–226.

Hendel, I., A. Lizzeri, and N. Roketskiy (2014). Nonlinear pricing of storable goods. *American Economic Journal: Microeconomics 6*(3), 1–34.

Hendel, I. and A. Nevo (2013). Intertemporal price discrimination in storable goods markets. *American Economic Review 103*(7), 2722–2751.

Ida, T., T. Ishihara, K. Ito, D. Kido, T. Kitagawa, S. Sakaguchi, and S. Sasaki (2022). Choosing who chooses: Selection-driven targeting in energy rebate programs. *Working Paper*.

Ito, K., T. Ida, and M. Tanaka (2023). Selection on welfare gains: Experimental evidence from electricity plan choice. *American Economic Review 13*(4), 2937–2973.

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

Jack, B. K., K. McDermott, and A. Sautmann (2022). Multiple price lists for willingness to pay elicitation. *Journal of Development Economics 159*, 102977.

Johnson, T. R. and M. Lipscomb (2022). Pricing people into the market: Targeting through price discrimination. *Working Paper*.

Kang, Z. Y. (2022). Optimal indirect regulation of externalities. *Working Paper*.

Langer, A. and D. Lemoine (2022). Designing dynamic subsidies to spur adoption of new technologies. *Journal of the Association of Environmental and Resource Economists 9*(6), 1197–1234.

Levitt, S. D., J. A. List, S. Neckermann, and D. Nelson (2016). Quantity discounts on a virtual good: The results of a massive pricing experiment at King digital entertainment.

Proceedings of the National Academy of Sciences of the United States of America 113, 7323–7328.

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *The Review of Economic Studies 38*(2), 175–208.

National Petroleum Agency (2022-2023). Indicative prices. Accessed on January 23, 2023.

Nauges, C. and D. Whittington (2017). Evaluating the performance of alternative municipal water tariff designs: Quantifying the tradeoffs between equity, economic efficiency, and cost recovery. *World Development 91*, 125–143.

Nichols, A. L. and R. J. Zeckhauser (1982). Targeting transfers through restrictions on recipients. *The American Economic Review 72*(2), 372–377.

Nikzad, A. (2023). Constrained majorization: Applications in mechanism design. *Working Paper*.

Pires, T. and A. Salvo (2015). Cash-constrained households and product size. *Working Paper*.

Rafkin, C., A. Solomon, and E. Soltas (2023). Self-targeting in US transfer programs. *Working Paper*.

Rochet, J.-C. and L. A. Stole (2003). The economics of multidimensional screening. *Econometric Society Monographs 35*, 150–197.

Smith, K. R., J. P. McCracken, M. W. Weber, A. Hubbard, A. Jenny, L. M. Thompson, J. Balmes, A. Diaz, B. Arana, and N. Bruce (2011). Effect of reduction in household air pollution on childhood pneumonia in Guatemala (RESPIRE): A randomised controlled trial. *The Lancet 378*, 1717–1726.

Solon, G., S. J. Haider, and J. M. Wooldridge (2015). What are we weighting for? *Journal of Human resources 50*(2), 301–316.

Stoner, O., J. Lewis, I. L. Martínez, S. Gumy, T. Economou, and H. Adair-Rohani (2021). Household cooking fuel estimates at global and country level for 1990 to 2030. *Nature communications 12*(1), 5793.

Szabó, A. (2015). The value of free water: Analyzing South Africa's free basic water policy. *Econometrica 83*, 1913–1961.

World Bank (2022). *Poverty and shared prosperity 2022: Correcting course.* The World Bank.

World Health Organization (2023). Global health observatory, household air pollution data. Technical report, World Health Organization.

World LPG Association (2023). Defining and measuring LPG subsidies. Technical report, WLPGA.