Environmental externalities and intrahousehold inefficiencies*

Kelsey Jack†
Seema Jayachandran‡
Sarojini Rao§

Preliminary and Incomplete

Abstract

When consumption generates negative externalities, the preferred policy solution is to set a price that reflects the social cost of consumption. In some cases – for example, household water and electricity use – consumption is susceptible to a second externality problem: each individual enjoys the private benefits of consumption but shares the costs with other household members, leading to overconsumption even from the household's viewpoint. We test the prediction that intrahousehold inefficiency dampens price sensitivity in the context of water use in Zambia, combining billing records, randomized price variation, and lab-experimental measures of intrahousehold efficiency. We find that households with above-median measures of intrahousehold efficiency have a short-run price elasticity of -0.463, while those with below-median efficiency have an elasticity of -0.147. These results suggest that the required Pigouvian price when usage is billed at the household level, yet own and other household members' consumption is difficult to observe, will need to correct both the environmental and the intrahousehold externalities. Alternative policies such as individual-level price incentives or access to real-time data on household water consumption (which would improve the enforceability of intrahousehold agreements) could also be useful.

*We thank the Southern Water and Sewerage Company for their cooperation and collaboration on this project. We have benefitted from comments made by audience members at the University of Chicago, the World Bank and Tufts University. Flavio Malagutti, Lorenzo Uribe, Lydia Kim, and Alejandro Favela provided outstanding research assistance. We also thank Amanda Kohn for designing the flyer with water pricing information that was used in the study. The International Growth Centre and J-PAL's Urban Services Initiative and Governance Initiative provided financial support. All errors are our own.

†Department of Economics, Tufts University
‡Department of Economics, Northwestern University
§Department of Economics, University of Chicago

1 Introduction

The standard policy prescription for managing consumption externalities is to change prices to internalize the externality (through a Pigouvian tax, for example). The effectiveness of the pricing approach depends on consumers observing and responding to the price. However, when "the consumer" is a group of individuals – often a household – a second externality problem may arise: each individual enjoys the private benefits of consumption but shares the costs with other members of the group, which means that households might overconsume even relative to their first best. The empirical implication is that households in which individuals fail to internalize the costs they impose on other members of the household are less price sensitive. The price insensitivity of these households will drive up the required Pigouvian price.

We test the relationship between the degree to which individuals within the household internalize the utility of other members of the household – which we refer to as intrahousehold efficiency – and price sensitivity in the context of water usage in Livingstone, Zambia.[1] We combine monthly billing data for 1,282 customers of the regional water utility with lab-experimental measures of intrahousehold cooperation and a randomized intervention that generates exogenous variation in prices. We separately survey both members of married couples that receive monthly water bills reflecting household consumption. As part of the survey, each respondent plays a modified dictator game with his or her spouse, which provides an incentive compatible measure of intrahousehold efficiency. In addition, randomly selected households are given a financial incentive to conserve water. If household water use, as measured by the monthly bill, falls below a specified consumption target, the household is entered into a lottery with a 1 in 20 (or better) probability of winning a substantial payout. This financial incentive to reduce consumption is akin to raising the price of water over some range of consumption. The incentive treatment is given to the couple together in one-third of the treated households, and to the man or woman alone in each of the other thirds. We follow households for three to nine months after treatment and observe both consumption and bill payment behavior.

In a simple household model, adapted from a non-cooperative household decision framework, individuals within the household make water use decisions taking others' decisions as given. Because the individual retains the full private benefits of consumption while sharing the costs (via the water bill) with the rest of the household, consumption will tend to be

[1]Water consumption in developing countries is a candidate for pricing policies to address externalities. Seasonal scarcity results in periods of water shortage in many developing country cities, including Livingstone (NWASCO 2015). The externalities that users impose on one another are exacerbated by poor infrastructure and increased weather volatility due to climate change (Jacobsen et al. 2013; Van der Bruggen et al. 2010).

above the household's Pareto optimal level.[2] We introduce heterogeneity across individuals and households by allowing each decision maker to internalize some share of others' payoffs. In the limit, each member of the household fully internalizes the payoffs to others, and the household and individual optimum converge. The model, along with our empirical set up, is agnostic about whether individuals internalize each other's payoffs because of altruism or monitoring and enforcement. Our set up generates three main predictions that we test with our data. First, we predict that more efficient households will be more price sensitive. Second, the effect of a change in individual-level prices within the home depends on whether the person who experiences the price change is already the residual claimant on the water bill. If not, the price change will have a relatively greater effect on that individual's – and therefore the household's aggregate – water use. Third, the individual-level price change will also be more effective if it is directed to the bigger water user in the household.

We observe an average decrease in monthly water use of 6.3 percent in response to the financial incentive to reduce consumption, which corresponds to a short run price elasticity of around -0.275. Consistent with the predictions from our model, the average response is driven almost entirely by more efficient households. Couples that share more with each other in a modified dictator game — our proxy for intrahousehold efficiency — reduce their water use more in response to the incentive treatment. We also observe heterogeneity based on which member of the couple is given the incentive to conserve. Specifically, the incentive treatment targeted at the woman in the household leads to reductions in water use that are twice as large as when the incentive is targeted toward the man or to the couple, though effects across incentive sub-treatments are not significantly different from one another. This greater sensitivity to a change in price targeted to the woman is consistent with our household model given two features of our sample: in the majority of households, the man is the residual claimant on the water bill and the woman uses more water than her husband. We test the last two predictions directly and observe that the individual price incentive is more effective if is directed to the individual who is not otherwise the residual claimant on the water bill and to the individual who uses more water. These two characteristics are highly correlated and, when tested together, residual claimant status is a more important driver of the effect.

Families in Livingstone devote a relatively large share of their incomes to water: an average monthly bill in our study setting is around USD 10.[3] Median monthly income among

[2]Our set up parallels a moral hazard in teams model, in which each individual shares the income from production effort – in our case, water conservation – with other team members but bears effort costs privately. Each player will therefore free ride on the effort of other team members. As Holmstrom (1982) points out, one solution to the free riding Nash equilibrium is a third party principal who acts as a budget breaker. This means that teams of two (a married couple, for example) may be least equipped to overcome the free riding problem.

[3]We use an exchange rate of 10 Kwacha / USD and adjust for inflation to 2015 USD values throughout.

households with piped water in Livingstone is around 220 USD (2015 equivalent; LCMS 2010), so households allocate roughly 5 percent of income toward their water bill.[4] As a result, the incentive for households in our sample to conserve water, even absent our intervention, is strong relative to that in many developed countries where water or energy expenditures are a relatively small share of income. In spite of the substantial cost, consumption levels are high; households in our sample consume roughly half the monthly average of US households. Two features of the billing environment interfere with this incentive and may also affect households' response to our price incentive treatment. First, water is priced on an increasing block tariff, which results in a poor understanding of water prices and how consumption decisions affect the bill. We elicit price perceptions using a carefully framed procedure that elicits beliefs over quantities that are salient to water use decisions. While the median response is close to the truth, we observe a wide variance in beliefs across respondents and wide confidence intervals around individual beliefs. All households that receive the price incentive also receive information about the actual price of water. In addition, a sub-sample of household receive information about prices but no incentive to conserve. Second, a lack of trust in the water provider or a misunderstanding of the billing process might undermine the association between consumption and the monthly bill. In our survey, when respondents are asked whose water use is to blame for a high monthly bill, 43 percent put the blame on the water provider. This may be in part because of poor access to information about consumption: only 47 percent of men and 21 percent of women were able to correctly identify the consumption quantity on their monthly bill. We introduce a cross-cutting "provider credibility" treatment that explains how bills are generated. Neither the price information nor the credibility treatment result in measurable impacts on water use, even when prior beliefs about prices or the provider are taken into consideration.

Our empirical strategy, which relies on existing variation in intrahousehold decision-making across households, is susceptible to concerns about unobservables that may be correlated with both measured heterogeneity in decision-making (specifically, giving in the dictator game and residual claimant arrangements) and household-level price sensitivity. We address these concerns in two ways. First, and most importantly, we develop multiple predictions about the types of households that should respond most to our incentive intervention, including differential predictions based on which member of the couple receives the incentive. Support for each of these predictions makes it less likely that omitted variables correlate with multiple intrahousehold characteristics. Second, we investigate observable sources of

[4] Note that this exceeds the U.S. EPA's water affordability threshold of 2 percent of income and UNDP's threshold of 3 percent of income. In the United States, around 12 percent of households are estimated to be below the affordability threshold (Mack and Wrase 2017).

heterogeneity, such as household size or wealth, that have intuitively plausible effects on both intrahousehold decision-making and price sensitivity. While the line between controlling for spurious correlation and eliminating relevant variation is somewhat arbitrary, our main results stand up to a series of robustness checks.

The study links two previously unconnected strands of literature. First, the household economics literature studies deviations from the unitary model of the household, where even non-unitary households (i.e. different preferences or payoffs across members of the household) may be efficient if members of the household internalize each others preferences or are able to contract over bargaining surplus (Browning and Chiappori 1998; Lundberg and Pollak 2008). However, the presence of intrahousehold externalities, limited commitment or asymmetric information can result in Pareto inefficiencies (Lundberg and Pollak 1994; Bloch and Rao 2002; Chen 2013). Most of the existing empirical literature on non-unitary decision-making has failed to reject Pareto efficiency, with exceptions primarily associated with production or income pooling (e.g., Udry 1996; Duflo and Udry 2004; Ashraf 2009). Our focus is on consumption decisions, where evidence of non-cooperative outcomes are relatively scarce (examples include Bobonis 2009; Attanasio and Lechene 2014), though the individual consumption of water or electricity may be considerably more difficult to monitor than other types of consumption.[5] Our design is based on heterogeneity in in the degree of intrahousehold frictions, which has been addressed in only a limited number of studies in the intrahousehold literature (Hoel 2015; Angelucci and Garlick 2016, Angelucci and Garlick (2016)). Our findings highlight a channel through which intrahousehold inefficiencies generate externalities outside of the household.

Second, the literature on corrective pricing in environmental economics highlights a number of reasons that consumers may fail to respond to prices, including misperceptions of price (Ito 2014; McRae and Meeks 2016), a lack of information about the price (Jessoe and Rapson 2014; Kahn and Wolak 2013), or a lack of salience (Allcott 2011). The incentive mis-alignment within the household that we study resembles the incentive problem between landlords and tenants, which leads to over-consumption of electricity (Levinson and Niemann 2004) and underinvestment in efficient durables (Myers 2015). Intrahousehold inefficiencies may be relatively important in developing countries, where a substantial share of income goes toward energy and water and where gender roles are often imbalanced.

As with any friction that distorts price sensitivity, the welfare implications of addressing the consumption externality directly versus first addressing the friction depends on both the cost and effectiveness of each type of intervention. Our results imply that the price increase

[5] A survey of 116 online respondents supports this: 82 and 72 percent of respondents said that electricity and water were among the most difficult items to to track own and spouse's consumption, respectively.

needed to internalize consumption externalities will be larger in the face of intrahousehold inefficiencies. The welfare implications of our findings mirror those of Allcott et al. (2014), where correcting the externality also corrects the intrahousehold "internality". However, the loss to households – money that could have been spent on other goods – under a corrective tax is increasing in the size of the intrahousehold distortion. Poorer households are disproportionately prone to intrahousehold inefficiencies in our sample, and may have a relatively high marginal utility of consumption of other goods. Given the potentially high consumer welfare cost of increasing the price of water for low income households, further attention to intrahousehold decision frictions and low cost policy interventions designed to address them offers a promising direction for future research. We show suggestive results that better information about consumption increases price sensitivity: households in which both spouses can read their water bills or are aware of one anothers' consumption are more responsive to the price incentive treatment.

Methodologically, our approach has parallels in other studies that collect experimental evidence on household and other small group decisions around efficient resource use (Ostrom 2006; Velez et al. 2009; Mani 2011; Kebede et al. 2013). Few, however, have linked these measures directly to decisions outside of the experimental game. In this way, our proposed approach resembles that of Karlan (2005) and others, where experimental measures in the field are used to explain heterogeneity in real world decisions.

The paper proceeds as follows. The next section presents a simple model of water use in the household. Section 3 describes the experimental design and implementation. Section 4 presents the results, and Section 5 concludes.

2 Water use within the household

We describe a stylized model of individual water use within the household to generate predictions about price sensitivity at the household level. The household's total water consumption W is the sum of the consumption decisions w_i of each individual $i = A, B$, which we will take to be husband and wife.

2.1 Set up

The water utility observes the household's total water consumption $W = w_A + w_B$ and charges the household pW, where p is the average marginal price of water for the household. Individuals observe their own consumption w_i but not that of their spouse w_{-i}. Individual bargaining weights $\lambda_i > 0$ determine ex post division of the income Y that remains after the household pays the water bill: $\lambda_i(Y - pW)$. Bargaining weights sum to one ($\lambda_A + \lambda_B = 1$) and aggregate water expenditure cannot exceed income ($Y \geq pW$).[6]

2.2 Individual best response

We assume non-cooperative decision making in which i chooses her water consumption, w_i, taking w_{-i} as given.[7] Each individual receives utility from water use and disutility from paying for water: $v_i = f(w_i/\gamma_i) + \lambda_i(Y - pW)$ where $f' > 0$ and $f'' < 0$. The parameter γ_i measures heterogeneity across individuals in their utility from water consumption (or disutility from conserving); those with higher γ_i consume more.

Individuals may also internalize some share $\alpha_i \leq 1$ of their spouse's utility, with person i's utility function given by $u_i = v_i + \alpha_i v_{-i}$. For the time being, we refer to α_i as a measure of i's altruism toward his or her spouse. Person i chooses w_i to satisfy the first order condition:

$$f'(w_i^*/\gamma_i)/\gamma_i = p(\lambda_i + \alpha_i\lambda_{-i}). \tag{1}$$

For $\lambda_i = 1$ or $\alpha_i = 1$, person i fully internalizes the household's cost of water consumption. Equation (1) shows that w_i^* is decreasing in p, λ_i, and $\alpha_i\lambda_{-i}$ (given $f' > 0$ and $f'' < 0$), and increasing in γ_i. Intuitively, i will use less if she is the residual claimant on any money saved, or if she is altruistic toward the savings that will accrue to her spouse, and will use more if she has greater utility from water use.

2.3 Effect of a price change

Our experimental treatments are designed to make water use more costly to the household, effectively increasing the price. From equation (1), an increase in the price decreases individual water use. Water use is observed at the household level, so the testable prediction is: $\frac{\partial W}{\partial p} < 0$, i.e., the demand curve is downward-sloping.

Our focus, though, is on how this price sensitivity differs based on the inner workings of the household. At the individual level, $\left|\frac{\partial w_i}{\partial p}\right|$ is increasing in person i's altruism toward his or her spouse, α_i. This is easy to see from equation (1), where α_i enters the righthand side multiplicatively with p. The testable predictions, denoting the average α_i in the household as $\overline{\alpha}$, are as follows.

[6]Note that households have an additional margin of adjustment in the timing and completeness of bill payment. We assume for the moment that bills are paid in full and on-time.

[7]w_{-i} is not observed but can be inferred, potentially with some error. In contrast to this non-cooperative solution, a cooperative household would jointly choose aggregate consumption to equate the weighted marginal benefit and the price. We discuss observability versus contractibility as the cause of the non-cooperative solution in Section 2.4.

Result: $\frac{\partial^2 W}{\partial p \partial \bar{\alpha}} < 0$. In other words, households that are more altruistic, on average, are more price sensitive.

While individual water use is not observable, we note that individual use will contribute disproportionately to aggregate water use when the more intensive water user's behavior changes. This can be seen by moving the γ_i from the denominator on the lefthand side of equation (1) to the righthand side, where it enters multiplicatively with λ_i and $\alpha_i \lambda_{-i}$. Consequently, price sensitivity will be increasing in the degree to which incentives are aligned within the home, specifically, where the intensity of water use (γ_i) is aligned with residual claimant status (λ_i) or with internalizing spouse's savings on the bill ($\alpha_i \lambda_{-i}$).

Result: $\frac{\partial^2 W}{\partial p \partial \gamma_i \lambda_i} < 0$. Households are more responsive to a price change if the more intensive water user is residual claimant on the bill.

Result: $\frac{\partial^2 W}{\partial p \partial \gamma_i \alpha_i \lambda_{-i}} < 0$. Households are more responsive to a price change if the more intensive water user is residual claimant on the bill.

In our experiment, in addition to manipulating the household price p, we also give financial rewards for a reduction in household consumption to a specific individual, which one can think of as manipulating a person-specific component of the price, which we denote P_i. The individual utility function then becomes $v_i = f(w_i/\gamma_i) + \lambda_i(Y - pW) - P_i W$. The effect of such a manipulation depends on i's intensity of water use, or γ_i, and her existing incentives to conserve water, λ_i.

Result: $\frac{\partial^2 W}{\partial P_i \partial \gamma_i} < 0$. Household consumption is more sensitive to a person-specific price targeted at a more intensive water user. This can be seen by via the revised first order condition, in which P_i and γ_i enter multiplicatively.

$$f'(w_i^*/\gamma_i) = \gamma_i(p(\lambda_i + \alpha_i \lambda_{-i}) + P_i) \tag{2}$$

Result: $\frac{\partial^2 W}{\partial P_i \partial \lambda_i} > 0$. Household consumption is less sensitive to a person-specific price aimed at someone who has more residual claimancy (higher λ_i) in the household. This is because this person already feels a larger proportion of the water price ($p\lambda_i$), such that the change in P_i is smaller relative to existing incentives around water use.

2.4 Discussion of assumptions

What makes water special A key feature of water consumption implicit in our setup is that the household – not the individual – pays for water. Household utilities such as water or electricity tend to have this feature in contrast to, for example, clothing, where a couple could divide up income and make individual purchases. Note that this point is distinct from saying water is a public good; individuals can consume individually but not purchase individually.

For other goods, such as food, spouses could make individual purchases but do not typically do so; it is natural that ingredients for cooking are not assignable, but, say, snack food is. This raises the other key feature of water assumed in this setup: lack of observability. A spouse's water use is difficult to observe. First, it is hard to match water quantities to activities (e.g., how many gallons used in a 5 minute shower, how many gallons used to wash dishes). Second, feedback on consumption is infrequent since it typically arrives once a month with the water bill. This compounds the observability problem. Consider another type of sometimes secretive consumption where feedback is more straightforward: if you notice that the number of cookies in the cookie jar has decreased since the last time you were in the kitchen, you know one of your family members stole a cookie from the cookie jar. If water meters were more accessible and easier to interpret, an individual could check the meter before and after a spouse's shower to observe consumption.[8] Adding to these observability challenges, knowing one's own consumption is often difficult. Even ex post, if i can only observe own consumption with some error ϵ, then she can only infer w_{-i} from the total bill with error: $\hat{w_{-i}} = W - (w_i + \epsilon)$. Moreover, some part of water consumption is a public good at the household level (e.g., washing the family's dinner dishes) but is still subject to some "private" consumption if doing dishes in a way that minimizes waste is privately more costly.

Altruism versus enforcement Our modeling set up abstracts from enforcement and takes as given that the spouse's water use is not observable or not verifiable. In practice, water use might be partly observable, in which case monitoring and enforcement of intrahousehold agreements becomes relevant. Even if water use were observable, difficulty enforcing intrahousehold agreements is sufficient to result in inefficient levels of aggregate consumption, and might lead to variation across individuals and households in α_i. Going forward, we refer to α_i and $\bar{\alpha}$ as measures of intrahousehold efficiency, to accommodate the possibility that either higher levels of altruism of better monitoring and enforcement might drive individuals within a household to consume closer to the household optimum. While we do not explicitly model the nature of the intrahousehold friction, we consider empirically that households in which more water use is more observable (i.e. $\bar{\alpha}$ is higher due to monitoring and enforcement) are also more price sensitive.

[8]This improvement in intrahousehold observability may explain part of the decline in electricity use associated with the introduction of smart metering (Jessoe and Rapson 2014).

3 Experimental design and data

An empirical test of our predictions requires three inputs: (1) a measure of aggregate household water consumption W, (2) exogenous variation in water prices p and P_i, and (3) measures of λ_i, γ_i and α_i. We describe how we operationalize each of these in turn. Our experimental design is summarized in Figure 1.

3.1 Water use

We partnered with the private regulated utility, the Southern Water and Sewerage Company, that provides piped water to residents in Livingstone, Zambia. Households are billed based on monthly meter readings, and charged according to an increasing block tariff.[9] Our main outcome measure is household water use per month. We obtain monthly billing records, based on meter readings, for January 2012 through September 2016. For each household in our sample, we create a panel that extends four months after treatment and 20 months before treatment, ensuring that observations for all households cover a two year window, regardless of when they were surveyed.[10] Water use is measured in cubic meters based on physical water meter readings collected monthly between the 20th and 25th of each calendar month. We keep only successful meter readings (i.e. drop the months in which meter readings were estimated or a meter was reported as broken or disconnected). We log transform the consumption outcome variable, which drops a small number of zero reading months (which are likely billing errors in any case). We generate an indicator for the month following a zero or missing observation to account for the fact that the first actual reading after an estimated reading or month with a broken meter may also reflect only part of that month's consumption.

The tariff schedule for 2015, when our study took place, is shown in Appendix figure A.1. The average price in the pre-intervention period (2013-15) among households that we survey is 4.36 Kwacha, or around 0.44 USD, per cubic meter. Average household consumption is around 19 cubic meters per month, a little under half of typical US household consumption, resulting in monthly consumption charges of around 85 Kwacha or 8.50 USD per month.[11] While we do not have household level monthly income or expenditure measures for our sam-

[9] Tariffs are regulated by the National Water Supply and Sanitation Council, and are intended to recover operating and maintenance costs, with cross subsidization from high to low tariff blocks and across customer types (NWASCO 2014).

[10] Households received up to 8 months of treatment, so we discard some treated months in favor of allowing all study households to contribute equally to the estimated treatment effect. As a robustness check, we include all treatment months in the analysis.

[11] Customers are charged for meter rental at a rate of 5 Kwacha per month and for sanitation and sewerage as a fixed proportion of monthly water use.

ple, we use the 2010 wave of the Living Conditions Monitoring Survey (LCMS), restricted to households with piped water in urban Livingstone, to calculate a median monthly expenditure of 192 USD (CPI adjusted to 2015 USD) and a median monthly income of 220 USD.

In addition to the measure of consumption, we estimate the impacts of the intervention on other customer outcomes including payment behavior and missing meter readings. Our outcome measures and other relevant statistics related to the monthly bill are shown in the top panel of Table 1, which also tests for balance across the treatments as discussed below.

3.2 Price shock

Randomly varying water prices was infeasible in our setting. Instead, we manipulate the household's experienced water price by increasing the returns to water conservation. The treatment is implemented in conjunction with a household survey, run between May and December 2015. Households in the *incentive treatment* are provided with a monetary incentive to reduce water use. In the months following treatment, households are entered into a lottery for 300 Kwacha (30 USD) for reductions in billed consumption, relative to an account-specific reference level. Conditional on qualifying for the lottery, households had a 1 in 20 (or better) chance of winning. To qualify for the lottery, households had to reduce their consumption by at least 30 percent relative to their average water usage in a two-month reference window, which resulted in a mean reduction target of 5.8 (median 4.95) cubic meters. The reference window was updated twice over the course of roughly eight months of fieldwork.[12]

In the notation of the household model, the incentive treatment adds a term to the indirect utility function: $v_i = f(w_i/\gamma_i) + \lambda_i(Y - pW + R \times \mathbf{1}(W \leq \bar{W}))$, where R is the expected value of the lottery payout. With 1 in 20 probability of winning, the expected value of the lottery incentive is around 1.5 USD or 15 Kwacha.[13] The average price associated with the reference window is 5.1 Kwacha / cubic meter. With a target reduction of 5.8 cubic meters, the expected value of the (net) price shock is 2.06 Kwacha per cubic meter or a roughly 40 percent increase in the average price.[14]

[12]The first reference period (March-April 2015) was used for households surveyed in May-early August; this was updated using a June-July reference window for households surveyed through late September. In September, we expanded the sample; new households had July-August 2015 as their reference period. This sample was used from late September through December, when fieldwork was completed.

[13]The probability of winning could be higher than 1 in 20 since we drew one winner for every 20 eligible households. Thus, if 21 households were eligible, we drew 2 winners. This was explained to households. At the same time, eligibility was made somewhat more difficult by the fact that the utility bills in round numbers; a household with a reduction target in fractions of a cubic meter would have to cut back to the nearest whole cubic meter to qualify.

[14]Note that this calculation accounts for the increasing block tariff which causes the average price to fall from 5.1 to 4.6 Kwacha per cubic meter. We therefore calculate the "net" price shock after accounting for

Our incentive treatment consists of three sub-treatment arms. In the first, both spouses learn about the lottery, and know that information is provided to both. In this case, the intervention is analogous to an increase in p. The second and third sub-treatments provide only the wife or only the husband with information about the lottery and its outcome. These individual sub-treatments move the payoff from the lottery to outside of the λ_i term: $v_i = f(w_i/\gamma_i) + \lambda_i(Y - pW) + R \times \mathbf{1}(W \leq \bar{W})$, so are analogous to an increase in P_i in the model. This increases i's unilateral payoff from water savings, which has the greatest effect on overall household consumption if $\gamma_i > \gamma_{-i}$ and if $\lambda_i < \lambda_{-i}$. In other words, the relative effect of the individual incentive is largest if it is targeted to the spouse with the greatest contribution to aggregate usage or to the spouse with the lowest returns to conservation in the absence of the incentive. Of course, individuals could share the information with their spouse or the spouse could find out out about it, but the individual-specific treatment comes closer to an individual price than does the joint treatment.

3.3 Other experimental manipulations

We introduce two additional sources of variation through our experimental design. The first, which provides price information, both acts as a potential additional source of price variation and as a control for beliefs. The second, which provides information about the credibility of the water provides, was intended to address misperceptions about the billing process.

3.3.1 Price beliefs

We leverage the fact that most households are unaware of their marginal price to generate additional variation in prices. As part of the survey, we elicit price beliefs for both spouses then, in a *price information treatment*, provide accurate information about water prices. All incentive treatment households are also in the price information treatment, as shown in Figure 1. A challenge in communicating price information to households is that the units of consumption are difficult to map on to consumption. We both elicit price beliefs and provide price information in units of usage rather than in cubic meters. Specifically, the survey asked "*suppose you wanted to save 20 Kwacha from your monthly bill; then, by how many minutes would your household as a whole have to reduce the use of the tap each day?*"[15] Treated

this mechanical reduction in the average price associated with lower consumption. Our elasticity calculation, which we describe below, does not require that we directly measure the magnitude of the price change.

[15]The question text included clarifications that we meant running the tap at a normal rate, as they would for daily activities like washing their hands, and also that we were asking them to think about the minutes that the tap was running during the various activities they did, and not the overall time spent doing chores. If the respondent said they did not know and could not provide an estimate, the questions was repeated once and they were given a second chance to respond. If they were still unable to answer the question, we asked

households received information that cutting back by 20 minutes per day would save the average household 20 Kwacha on their monthly bill.

The effect of the price information depends on prior beliefs about water prices. Specifically, if households fully update, then the "price treatment" associated with the information is just the difference between the true price and the prior. Thus, individuals with a prior below the true price receive a positive price shock and vice versa. We categorize individuals into beliefs above and below the price information, and construct a household level measure that equals one if either spouse underestimated the price. The price information intervention also serves to remove variation in price beliefs, which allows us to calculate price elasticities associated with the incentive treatment. Thus, all households enrolled in the incentive treatment also receive the price information treatment.

3.3.2 Provider credibility

In a cross-cutting *provider credibility treatment*, households were offered assurance that water bills are based on actual water use, to address concerns that both the lottery and price information treatments will be ineffective if households do not believe that bills reflect consumption. Treated households were given information about the timing of the billing cycle and how their bill is calculated in the event that a meter reader is unable to read the meter, either because it cannot be accessed (the gate is locked) or because it is too unclear to be read. The information was paired with a reassurance that the provider is committed to honest billing practices and try to ensure that households are only charged based on their actual water usage. [16]

them about a series of narrowing intervals, e.g. less than 20, 20-40, 40-60, more than 60, and conditional on the first interval, we asked about 5 minute intervals within that first interval, and then re-asked the main price belief elicitation question. 81% of men and 83% of women answered the question the first time it was asked, and an additional 10% of respondents answered the question in the second or third attempt. We then asked them about the highest and the lowest that they thought the number of minutes could be, and then asked them again for a best guess, giving them a chance to revise their previous answer if they wanted. The price belief elicitation question was asked after a series of questions on their own and their spouse's water use, as we found during piloting that thinking about water intensive chores beforehand made it easier for respondents to understand the question. The price elicitation module was piloted with almost 300 households, who were then excluded from our sample.

[16]The script for the provider credibility treatment is as follows: We have collected this information purely for research and will not share any details with SWSC. However, we want to provide you with a little bit of extra information about how SWSC calculates your bill. SWSC tries to ensure that bills are accurate by reading your meter monthly and using the amount of water consumption shown on your meter to calculate your bill. That is, the amount that you are charged is based on the amount of water you use. The meter readings taken this month measure your usage since the time when last month's reading was taken. Once SWSC has collected all the readings for this month, this is used to calculate the bill that will be given to you next month. For example, when you received your water bill in March you were charged for the water your household used between the 21st of January and the 20th of February, roughly speaking. When you received your water bill in April, you were charged for the water your household used between the 21st of February

3.4 Intrahousehold measures

We measure γ_i, λ_i and α_i through a household survey conducted separately (and simultaneously) for the husband and wife. A series of questions documents water use (γ_i), bill payment responsibilities (λ_i) and intrahousehold cooperation, enforcement and altruism. In addition, we conduct modified dictator games between spouses as part of the survey visit. Both spouses played the game concurrently and in private and the game was run by a trained surveyor who was of the same gender as the respondent. The game proceeded as follows.

The respondent is asked to pick one of two sealed envelopes and open it; the envelope contains either 20 or 30 Kwacha, and respondents only learns the value of their own draw, not the distribution. The surveyor explains that the money is theirs to keep and that they will be asked to make decisions using this money, but that they are under no expectation to share the amount. The respondent is then asked how much money they would send to their spouse, and separately, how much money they would send to a water conservation NGO. Before asking for their responses, the surveyor informs them that in each case, any money the respondent chooses to send will be doubled by the experimenter. The surveyor clarifies that the two decisions are mutually exclusive since the recipient of any money sent will be randomly selected after the respondent has made their decisions, with equal probability on each outcome, and that the respondent cannot influence which recipient is chosen at the end. The random endowments as well as the random selection of recipient ensures that respondents can conceal their own earnings from their spouse and, thus, her actions are based more closely on her own preferences rather than concern about retribution from her spouse. If the respondent chooses to send some money, and the spouse is randomly chosen as the recipient, the spouse will know how much money was sent (since both spouses played the same game), but not how much money the respondent started with and hence how much the respondent chose to keep for herself. Similarly, the respondent can also choose to send nothing to the spouse and claim that the NGO was chosen as the final recipient. The surveyor explains these aspects of the game to the respondent and asks questions during the explanation to check for respondent comprehension, so that respondents know what information can and cannot be hidden from their spouse.

The game provides a revealed preference measure of α_i. The more the respondent sends to her spouse, the more she cares about his financial resources. The game simultaneously measures enforcement-based income sharing within the household, which may or may not

and the 20th of March, and so on. If there are some months that they cannot get a meter reading, then you are charged an estimate based on your previous consumption, and they try to get meter readings again as soon as possible. Then the next time they read your meter, they adjust your bill for any over- or under-charges from the months when they were not able to do the reading. SWSC is taking measures to make sure that bills are fair and based on actual water usage. They are committed to honest billing practices.

carry over to water use; if a respondent expects that she can get back the amount she shared with her spouse, for example, she would share more. In this sense, sharing in the game might be predictive of bargaining power over income within the home, which may be correlated with λ_i (defined as residual claimancy over the water bill, specifically). For the analysis, we use survey questions to measure which spouse is the residual claimant on the water bill, with the idea that the household may arrive at different arrangements for more or less monitorable consumption goods. Specifically, both spouses are asked whose income is used to pay the water bill and who physically pays the bill (which is done in person in this context). If the respondent's answers to those questions match, that person is labeled as the residual claimant. If they do not, then follow up questions ask about how much discretion the one making the physical payment has over savings on the bill. We define gender-specific residual claimant indicators based on the respondent's perception of his or her claim on any savings on the bill (i.e. in some couples, both spouses may believe that they are the residual claimant).

We ask respondents to compare their own water use directly to that of their spouse. We define the woman (man) as the larger water user if both members of the couple indicate that she (he) uses more water. Unlike the residual claimant variable, where individual perceptions drive the incentive to reduce water use, we are interested in identifying which member of the couple actually uses more water, so we require that spouses' answers agree to assign bigger water user status. We construct a measure of incentive alignment within the home, which equals one if the agreed upon bigger water user also perceives him or herself to be the residual claimant on the bill. We also construct an indicator for the "stereotypical" intrahousehold arrangement, in which the husband is residual claimant on the water bill and the wife is the larger water user.

We measure the information about water use available to each spouse, which is necessary but not sufficient for monitoring and enforcement to result in efficient consumption. Specifically, we ask whether the respondent looks at the water meter. We test knowledge of the quantity of consumption and the total charge on the household bill. We also ask each spouse about their top three water using activities, and construct a variable equal to one if both spouses know above the median share of each others' reported water using activities.

Finally, in addition to survey questions on water use and intrahousehold decision making, the survey collects information on demographics and socioeconomic status, as well as attitudes toward the water utility.

3.5 Sample construction and summary statistics

3.5.1 Sample construction

Our sampling takes the universe of metered household accounts as provided by the water utility (SWSC) and imposes some restrictions based first on billing data, and then based on a short screening exercise that was conducted in the field. Using the panel of billing data for metered residential customers as of February 2015 (N=9,868),[17] we eliminate households that did not have a working meter for at least 3 out of the 4 preceding months. We also excluded households that use no water (i.e. are billed for zero cu.m.) in more than half of the preceding 4 months. Households with very low variation in usage over the preceding four months were considered to have possibly tampered with the meter or have a delinquent meter reader were also excluded based on the following criteria: if the coefficient of variation in this period was less than 0.05, or if the quantity reported was identical for 3 or more months. Households with consistently low usage were also excluded as they would be least able to adjust their water consumption in response to a price shock; we drop households if their usage was on the lowest price tier (less than 6 cubic meters) for more than 2 of the preceding 4 months. Households whose median water usage in the preceding four months was above the 99th percentile were also dropped as they could also have had malfunctioning meters, or may not be as responsive to price, and may also have been significantly more difficult to survey (because they were presumably very wealthy households or firms mislabeled as residential customers by SWSC). Finally we drop households with an extremely high outstanding balance with the water utility, or households that are owed a significant amount of money by SWSC, defined as 6 times or 4 times their median bill in the preceding four months, respectively. This yields a total of 7,425 households that we target for an in-person screening.

Households were visited by a surveyor to collect data on characteristics not observed in the billing data that were also important for sampling. Specifically, we require that the water meter not be shared with other households, that the primary bill payer be married (or cohabiting) and that both spouses live at that address, and that the household was in residence for at least the 4 month period prior to April 2015. We also exclude households planning to move in the following 6 months. Our surveyors attempted to screen each household 3 times; any adult member of the household could be given the screening questionnaire. In total, 6,594 households were screened, of which 31 percent (2,051) met all our screening criteria.[18] We scheduled survey appointments with 1,817 households from our eligible sam-

[17]This number excludes households involved in a pilot of the project, who were deemed ineligible for the full study.

[18]Reasons for not screening a household include that the home was vacant or under construction, that it was occupied by a business, or that no one was home for three consecutive attempts.

ple. Of these, we completed surveys with 1,282 households. This high "attrition" rate is due largely to stopping our attempt to survey households at the end of December 2015.

Because of the restrictions we impose during the sampling stage, idiosyncratic factors that cause households to become ineligible (because, for example, of missing consumption outcomes in the months before the survey) will result in some mechanical mean reversion in water use outcomes following the survey. To improve precision, we therefore include all households that we sampled but excluded during the screening stage. These households were sampled using the same criteria as the households that were ultimately surveyed, but were screened out after the surveyors' initial visit. This adds 5,312 households to our survey sample, which are not systematically different from the surveyed or treated households in terms of pre-survey consumption patterns.[19] Appendix table A.3 shows how the sample evolves at each stage of sampling and randomization.

Households that met the screening criteria were informed about the survey. We scheduled a follow up visit with the primary bill payer and his/her spouse, emphasizing that we needed both of them to be present for the full survey. We also informed respondents they would be compensated 40 Kwacha (~ 4 USD) for participating in the survey. At the scheduled time and date, a pair of surveyors (always a woman and a man) visited the screened household for a full survey. After a few preliminary demographic questions, husbands and wives were separated and surveyed individually in different rooms of the house. Enumerators elicited water price beliefs, asked for perceptions of own and family members' water usage, and conducted the modified dictator's game. After finishing their individual questionnaires, both surveyors and respondents met back together in a common room for the last survey questions, and to receive the price information. We brought the couple back together so that the price information would be communicated to both (in the information treatment) and to avoid an awkwardness that might arise from ending the survey immediately following the game transfers with the couple separated.

3.5.2 Sample statistics

Table 1 summarizes characteristics of the sample, including elicited price beliefs and attitudes toward the water provider, and tests for balance between the incentive treatment and control group (which includes some households that received the price information and/or some the provider credibility treatments). The top panel shows statistics from the water bill, and so we report means and standard deviations for the households that were screened out of the

[19] Because we rely on the date of the survey to define the treatment timing in the panel billing data, we define a quasi-treatment date for the households that were screened but not surveyed. For households that were screened on a day that produced at least one completed survey, we use that survey date. When that strategy is not feasible, we use the average lag between screening and surveying (7 days).

survey but are included in our analysis. The middle panel shows household characteristics gathered through the survey. Around half of the sample owns their own home and the average household size is close to six. Around 16 percent employ a maid. Our attempt to explain intrahousehold decision making around water use based on spousal preferences clearly simplifies the dynamic in a home with an average of four additional members, plus – in some cases – a maid. Around half of the sample owns their own home. Levels of English language proficiency are high. Price beliefs are reasonably accurate: in over 60 percent of households, at least one of the two respondents underestimated the price. Distrust of the service provider is high: in over 40 percent of households, both spouses say that high bills are the fault of the provider, i.e. not because of high consumption.

3.5.3 Intrahousehold measures

Our preferred measure of intrahousehold efficiency comes from the respondent's incentive compatible decision of how much to share with his or her spouse in the dictator game. We summarize the measure in the bottom panel of Table 1. Husbands send a larger fraction of their endowment to their wives than wives do to their husbands. Both spouses send a smaller share of their endowment to the water NGO than to their spouse. Table 1 also reports other water use measures associated with our theoretical predictions. In around 30 percent of households, the wife says she is the residual claimant on the bill. In around percent of households, both spouses agree that she uses more water than her husband. Thus, in the typical household, the man is the residual claimant (higher λ_i) and the woman is the bigger water user (higher γ_i). In only around 2 percent of households are the incentives aligned, i.e. the residual claimant is also the bigger water user.

Of course, measures of intrahousehold efficiency may be correlated with other household characteristics that affect both water use and price sensitivity. Table 2 shows correlations between three dictator game outcomes – the share of the endowment sent by the husband, by the wife, and whether the average share sent was above median (our main measure of intrahousehold efficiency) – and individual and household characteristics. First, the share of endowment sent to the spouse is positively correlated with the share sent to the water NGO. While this may indicate that individuals who are more altruistic in general are also more altruistic toward their spouses, it may also indicate some experimenter demand effect or confusion about the game, though we observe no correlation with a measure of social desirability bias (SDB score). Second, neither of the other measures that generate predictions in our conceptual framework (residual claimant arrangements or water use patterns) are correlated with the dictator game measure, nor are variables describing knowledge of the bill. Third, our average dictator game measure is negatively correlated with household size, with age, and with home ownership, and positively correlated with employing a maid, household assets, number of rooms in the home and English language fluency. On the whole, wealthier respondents appear to share more in the dictator game, perhaps unsurprisingly. In our robustness checks, we revisit these variables to determine if they also are associated with differential responsiveness to our incentive treatment.

4 Results

4.1 Predictions

The experimental design and data collection described in the previous section generates the following empirical predictions:[20]

1. The incentive treatments decrease water consumption.

2. The magnitude of the incentive treatment effects are increasing in intrahousehold efficiency, measured by $\bar{\alpha}$.

3. The incentive treatment effect is larger if the household has aligned incentives around water use, i.e. the larger water user is also the residual claimant, or the larger water user internalizes the payoffs from conservation that accrue to the residual claimant.

4. The individual-specific incentive treatments are more effective if they (a) go to the individual who is not otherwise the residual claimant on the water bill in the household and (b) go to the individual who consumes the most water in the household.

4.2 Identification

We implement a randomized intervention, with monthly outcome data before and after the intervention. A standard difference in differences regression identifies the treatment effect:

$$y_{it} = \beta_1 treatedhh_i + \beta_2 post_{it} + \beta_3 treatedhh_i \times post_{it} + \epsilon_{it} \tag{3}$$

where $treatedhh_i$ is a binary indicator for assignment to one of the treatment groups and $post_{it}$ is a time-varying indicator that turns on for household i in the month after the survey. Note that $post_{it}$ varies across households and not just over time because the survey and treatment were rolled out over time. β_3 identifies the differential change in the outcome

[20]Note that we derive predictions in Section 2 over water use levels while our empirical results test for effects on log water use. Rewriting the model in logs generates the same predictions.

among treated household after the survey. Also note that even though a household was only eligible for the lottery for its consumption in the first full billing cycle after the survey date, we set $post_{it}$ equal to 1 as of the survey date because it is possible the intervention had immediate effects. In our preferred specification, we omit the month in which the survey occurred since it is only partially treated.[21]

To improve precision, we include time and household fixed effects in our preferred estimates. We also include screened households in the analysis to contribute to the identification of time fixed effects and the coefficient on $post_{it}$. All specifications therefore include an indicator for whether the household was in the surveyed sample, which we define as $surveyed_{it} \equiv surveyedhh_i \times post_{it}$. We also define $treat_{it} \equiv treatedhh_i \times post_{it}$, where all treated households are also in the survey sample. We estimate:

$$y_{it} = \beta_1 treat_{it} + \beta_2 surveyed_{it} + \beta_3 post_{it} + \tau_t + \eta_i + \epsilon_{it} \quad (4)$$

where τ are month-year fixed effects and η_i are household fixed effects. In the presence of household fixed effects, β_1 identifies the treatment effect of interest, and β_2 captures any independent impact of being surveyed. We allow for arbitrary within household correlation in water use over time by clustering standard errors at the household level. Because gaps in the panel are associated with meter disconnections and other meter reading issues, we add a time-varying indicator to the month after a missing observation to control for the fact that these months may record only a partial month of consumption. Our main predictions involve heterogeneity in the response to treatment by household type, so we interact $treat_{it}$, $surveyed_{it}$ and $post_{it}$ with relevant household characteristics.[22]

We use the estimates of β_1 associated with our incentive treatment to calculate short run price elasticities as follows. First, with y_{it} equal to log of monthly water quantity, we can interpret the coefficient on $treat_{it}$ in the presence of household fixed effects as ${}^{\partial ln(q)}/_{\partial p}$ or ${}^{\partial q}/_{q} \times {}^{1}/_{\partial p}$, such that multiplying by the pre-intervention average price delivers a short run elasticity. We calculate customer specific average prices, accounting for the increasing block schedule and for inflation (CPI), in each pre-intervention month and use that as the basis for our sub-group specific average marginal prices.[23]

[21] Note that because the billing cycle starts on the 20th of each month, our definition of month corresponds to the billing cycle, i.e. July runs from June 21 to July 20.

[22] Note that we do not observe the variables associated with household type of households that we do not survey, so we set these observations to zero. Thus, there is no remaining variation in the interaction of observable characteristics and the $post_{it}$ dummy, which drops out of the specification.

[23] Note that this approach to calculating elasticities does not impose assumptions about how households perceive the price change, only that households knew their pre-treatment price. We increase the likelihood of this latter assumption by including all incentive treatment households in the price information treatment. However, given that these treatments were implemented concurrently, if the price information treatment

We show the exogeneity of treatment assignment to observable household characteristics in Table 1 for the incentive treatment and Appendix tables A.1 and A.2 for the information and provider credibility treatments. We also plot average water use across treatment conditions and our preferred measure of intrahousehold efficiency in the months leading up to the survey, after removing seasonal trends (Appendix Figure A.2). Overall, we observe parallel trends by incentive treatment and our binary measure of intrahousehold efficiency. The plot of average monthly consumption by our measure of intrahousehold heterogeneity suggests that more efficient households actually consume slightly more than their less efficient counterparts. Our conceptual framework predicts the opposite, all else equal. As shown in Table 2, other covariates are correlated with our efficiency measure and may contribute to the higher average consumption among more efficient households. We regress household average pre-intervention consumption on our dictator game measure (column 1) and a vector of other household-level covariates (column 2) and show the resulting correlations in Appendix table A.4.[24] Unconditional on other observables, we see significantly higher consumption among more efficient households. Conditional on observables, the coefficient shrinks and becomes insignificant (but remains positive). This relationship highlights a key concern associated with our empirical strategy. Our main empirical predictions involve differences in how households respond to a price shock; if unobservables affect both our key sources of heterogeneity (dictator game giving and residual claimant arrangements within the home) and price sensitivity, then we may misattribute an effect of omitted variables to our measures of intrahousehold decision-making. We address this both through testing multiple theoretically motivated hypotheses, some of which vary by treatment arm, and through robustness checks in Section 9.

4.3 Average treatment effects

We begin with a figure showing the distribution of consumption across the incentive treatment and the control (i.e. households not receiving the incentive treatment, which includes some in the information and provider credibility treatments), normalized by household-specific average consumption in the incentive treatment reference window. Figure 2 shows a decrease in consumption across most of the distribution, though relatively little increase in mass in either the treatment or control below the target level of 70 percent of reference

affected price perceptions, then past usage – which we use to calculate elasticities – is unaffected.

[24] We use a subset of the variables shown in Table 2. Specifically, we focus on (1) study design based measures, including the amount sent to the NGO in the dictator game and the survey measure of social desirability bias, (2) variables that are significantly correlated with dictator game outcomes. We construct an indicator equal to one if both spouses are over 50 and include only one of our various wealth measures, given their considerable correlation.

window consumption.

Table 3 reports average treatment effects of all three experimental interventions. Column 1 reports the standard difference in difference results from estimating equation (3), without household or time fixed effects. Columns 2 and 3 add household and month-year fixed effects, respectively. Average consumption across the treatment arms are not significantly different from the pool of surveyed control households (column 1). However, surveyed households consume less, on average, than do households that were screened out of the sample (this can be seen from the coefficient on surveyed).

Our main coefficient of interest is on Incentive x Post. We observe a statistically significant 6.2 to 6.7 percent decrease in monthly consumption in response to the incentive treatment, consistent with prediction 1: the incentive treatments decrease average water use. The implied short run price elasticity is -0.275 (column 3, based on an average pre-intervention price 4.36).[25] We observe no significant average effect from the other treatments or from being surveyed (the coefficient on Surveyed x Post). Going forward, we focus on the specification shown in Column (3), which includes household and month-year fixed effects. We define the treatment variables as time varying and report results from estimating equation (4).

While we observe no average effect of the other treatments, the effects of the price information and provider credibility treatments should depend on respondent beliefs about water prices and the correspondence between water use and bills. Thus, we supplement the analysis in Table 3 with specifications that (a) interact the price information treatment with an indicator for whether the husband and wife underestimated the price, on average, and (b) interact the provider credibility treatment with trust in the water provider. Both should lead to reductions in water use. Table 4 shows that these treatments are weak, even after taking heterogeneous responses into consideration. Going forward, we incorporate these treatments into the surveyed comparison group or the incentive treatment effect, as appropriate.

4.4 Intrahousehold heterogeneity

4.4.1 Household price incentives

Table 5 shows the main test of prediction 2, for which we pool the incentive treatment arms. We estimate the effect of the lottery treatments on household water use, and allow the effects to vary with the share of endowment sent in the dictator game on average for the household

[25]Our calculated short run price elasticity of demand is slightly below the mean found in the literature reviewed by Dalhuisen et al. (2003) and in line with the short run elasticities summarized in Worthington and Hoffman (2008). In the literature, the long run elasticity is generally shown to be larger than the short run elasticity.

(column 1) or by each spouse (column 2). Consistent with prediction 2, column 1 shows a larger reduction from households that sent above the median on average, with the total effect among these more efficient households resulting in a 10.5 percent decrease in water use. The average short run price elasticity among households with below-median dictator game contributions is -0.147 (based on a pre-intervention average price of 4.3) while the total effect for above-median households implies an elasticity of -0.463 (based on a pre-intervention average price of 4.4). Appendix table A.5 shows the robustness of these results to alternate approaches to aggregating the dictator game measure.[26]

Next we turn to prediction 3, which states that the effectiveness of the price incentive should be larger in households that have resolved the incentive problem by making the larger water user the residual claimant on the bill. Perhaps surprisingly, we see that very few households (28 out of 1282) have arrived at this solution themselves, leaving us underpowered to detect heterogenous effects. That said, column 2 of Table 5 shows substantially larger (if imprecisely estimated) reductions in households where incentives around water use are aligned.

Columns 3 and 4 show the differential effect of the incentive treatments by each spouse's dictator game outcome. Column 3 includes the full sample, while column 4 restricts the sample to households that follow traditional gender roles, in which the woman is the bigger water user and the man is the residual claimant on the bill. This restriction omits 551 households in which either the woman is not the bigger water user or the man is not the residual claimant on the bill. Our empirical prediction states that the degree to which the bigger water user internalizes her spouse's payoffs are more important is she is not the residual claimant, i.e. the woman's dictator game measure matters in households that follow traditional gender roles. Column 3 shows that, on average, the husband's dictator game decision is a better predictor of household response to the price incentive. Focusing attention on households that we classify as having traditional gender roles (column 4) reveals considerably more importance of the woman's dictator game decision on aggregate water use than in the average household. The results are based largely on qualitative comparisons of magnitudes; in most cases, we lack precision to statistically distinguish individual interaction terms.

[26]It may also matter whether spouses have similar levels of altruism. We observe similar decisions across spouses: for over half of the households we study, the difference in the share of the endowment sent by the husband versus the wife is less than 0.25, and for only 15 percent of households is it more than 0.5.

4.4.2 Individual price incentives

Our conceptual framework also clarifies that the effect of the individual lottery sub-treatments should depend on the existing incentives to conserve water. If the individual incentive recipient is not otherwise the residual claimant on water savings, then the effect is relatively large. In addition, if there is a disparity in water users' ability to affect the total amount consumed by the household, then directing the incentive to the larger user will also be more effective. The results shown in Table 6 breaks the effect down by treatment arm. Recalling the patterns of water use and residual claimant status shown in the bottom panel of Table 1, prediction 4 implies that the lottery directed to the wife should have a larger effect, given that in the typical household she is the larger water user and is not the residual claimant on reductions in the water bill. The pattern of coefficients in Table 6 confirms that the most effective lottery sub-treatment is directed toward the wife. We lack the precision, however, to reject that the effect of the wife lottery is statistically different from either of the other two sub-treatments.

We examine prediction 4 further in Table 7. Column 1 interacts the individual-specific incentive treatment with an indicator for whether individual incentive was directed to the non-residual claimant in the household. Column 2 interacts the individual-specific incentive treatment with an indicator for whether the incentive recipient is the bigger water user in the couple. We see that the individual price incentive led to significant reductions in water use if directed to the non-residual claimant or the bigger water user, though the former interaction is larger and more precisely estimated. Given the considerable correlation between these variables (specifically, the women is likely to be both), we first run both interactions together in column 3. The interaction with non-residual claimant remains significant while the interaction with bigger water user becomes very small and positive. Finally, we add, as a control, whether the individual incentive was directed toward the wife, to determine if the residual claimant effect is driven entirely by gender. It is not: the interaction with residual claimant status remains marginally significant. Overall, the results presented in Table 7 suggest that the existing arrangements around who has a claim to any savings from water consumption is an important determinant of the effectiveness of the price incentive. This raises the question of why households are not able to resolve this conflict themselves.

4.4.3 Altruism versus enforcement

As discussed in Section 2.4, households may be efficient because spouses are altruistic toward one another (or face no incentive misalignment) or because they are able to monitor and enforce water use. Our primary measure of intrahousehold efficiency – giving in the dictator game – may reflect either or both of these explanations. Though we cannot cleanly disentangle whether altruism or enforcement matters more, we use measures of the accessibility of water use information in the household to investigate whether the potential for intrahousehold monitoring predicts the household response to the incentive treatment. Table 8 shows heterogeneous treatment effects based on four measures of the observability of water use in the home. Column 1 shows a binary measure of whether both spouses report looking at the water meter. Columns 2 and 3 show binary measures (verified by the enumerator) of whether both spouses can identify the consumption quantity and total charge on the bill, respectively. Column 4 shows a binary measure of whether spouses show above-median awareness of each others' water use.[27] Looking at the meter appears to decrease price sensitivity, while the other three measures of information availability are associated with greater price sensitivity (all imprecisely estimated) with the largest differential effects based on knowing the consumption quantity on the bill and knowledge of spouse's water use. Of course, looking at the meter is both self-reported and potentially of little use if the numbers are not tracked or understood, while the other three measures of information are more obviously associated with information about water use. Overall, these results suggest that efficiency may be driven, in part, by the information needed for monitoring and enforcing agreements around water use.

4.5 Robustness checks

4.5.1 Interpretation of the heterogeneity results

As shown in Tables 2 and A.4, both the dictator game measure and average pre-intervention water use are correlated with certain observable characteristics of our study households. We examine this more directly by interacting a set of covariates with the incentive treatment indicator in Table 9, first one at a time (column 1), then all at once (columns 2 and 3). While we do see some heterogeneity in price sensitivity by these other measures, our main coefficient of interest (Sent above average to spouse) decreases only modestly and remains the largest magnitude of any of the interactions in column 3.

Note that some of these covariates should not necessarily be interpreted as sources of spurious correlation. For example, households that are more altruistic in general (as measured by Sent above average to NGO) may also be more altruistic within the home. Larger households may have a harder time enforcing income-sharing agreements of any kind. Controlling for these (and other) variables may therefore eliminate variation in our intrahousehold efficiency measure that we would like to include. The robustness check shown in Table 9 is

[27]See Section 3.4 for further details on these variables.

therefore a conservative test of our main result.

These robustness checks focus on the dictator game heterogeneity. Support for our model comes equally from the predictions of differential responses to the individual price incentives, depending on the residual claimant arrangements within the home. Given that the result is not just driven by the gender of the recipient of the individual price incentive (column 4, Table 7), the effect of unobservables has to be considerably more subtle to deliver our results. Namely, unobserved characteristics of the individual who is not typically the residual claimant would have to be associated with greater price sensitivity (at the individual not the household level, since the result depends on the individual incentive treatment arms), across both genders.

4.5.2 Specification and outcomes

We test for sensitivity to our specification by varying the sample and panel length. Results are shown in the top panel of Appendix table A.6. Both the main effect of the incentive treatment and the interaction with the above-median dictator game measure are similar when we include only the surveyed sample (columns 1 and 2) or include all treated months in the analysis (columns 3 and 4). The bottom panel of Appendix table A.6 shows the effect on quantity in levels and on the total bill (in logs). Results in levels are similar to the main results in logs, and the bill total changes in accordance with the observed consumption changes.

We test whether our observed consumption responses reflect other margins of household adjustment, namely bill non-payment or meter reader evasion. Table 10 shows little effect of any of the treatments on a measure of whether the household made a payment toward their bill (columns 1 and 2) or missing meter readings (columns 3 and 4). If anything, it appears that the provider credibility treatment may have led to a slight decrease in meter reader evasion.

5 Conclusion

We generate new evidence on the importance of intrahousehold decision making for consumption externalities that extend beyond the household. We combine water billing data with experimental measures of intrahousehold efficiency via a dictator game played between spouses and with an exogenous incentive for water conservation offered to a random subset of study households. Consistent with a simple model of individual water consumption that allows for non-cooperative intrahousehold decision making, we show that households in which spouses are less likely to internalize each others' financial gains (as measured by the dictator game) are also less responsive to the increase in the cost of consumption delivered by the conservation incentive. Also consistent with our model, targeting the conservation incentive to the individual within the household who otherwise has the least incentive to conserve water leads to greater overall response. Together, these results show that households that are internally less efficient are also less price sensitive.

Our main result relies on heterogeneity in a measured characteristic across households, so is susceptible to omitted variable concerns (i.e. some other factor may be correlated both with our measure of intrahousehold efficiency and with price sensitivity). While we cannot fully rule out this type of concern, we address it in three ways. First, we test multiple predictions that rely both on measured differences across household and design based variation in the payoffs to individuals within the household. While any one of these tests may be affected by some omitted variable, they are less likely to all be driven by unmeasured heterogeneity across households. Second, we rely on experimental measures of intrahousehold efficiency, which are more likely to accurately pick up the underlying characteristic of interest. Third, we measure a number of controls that allow us to perform more or less strict analyses of the variation allowed to influence household level price sensitivity.

As more households in the developing world acquire piped water and electricity connections, overuse associated with intrahousehold inefficiencies may impose large social costs. Where price-based policy solutions are feasible, our results imply that the price needed to achieve a given level of consumption will be increasing in the degree of intrahousehold inefficiency. While higher prices will correct this intrahousehold "internality" (Allcott et al. 2014), the loss to households – in the form of money that could have been spent on other goods – will be increasing in both the magnitude of the intrahousehold friction and the marginal utility of consumption. These challenges of intrahousehold misalignment may be disproportionately large in developing countries, where more traditional arrangements between spouses are likely to persist and where utility bills constitute a substantial share of monthly income. Somewhat surprisingly, in our study sample, the most obvious intrahousehold solution appears uncommon: only 2 percent of households have made the bigger water user the residual claimant on the water bill. While pricing water at the individual level may be infeasible, other interventions, such as real-time information about consumption or technologies that lower the cost of conservation to the individual, may be able to overcome the intrahousehold inefficiency at a lower cost to consumer welfare.

References

Allcott, H. (2011). Consumers' perceptions and misperceptions of energy costs. *The American Economic Review 101*(3), 98–104.

Allcott, H., S. Mullainathan, and D. Taubinsky (2014). Energy policy with externalities and internalities. *Journal of Public Economics 112*, 72–88.

Angelucci, M. and R. Garlick (2016). Heterogeneity in the efficiency of intrahousehold resource allocation: Empirical evidence and implications for investment in children. *Working Paper*.

Ashraf, N. (2009). Spousal control and intra-household decision making: An experimental study in the philippines. *The American Economic Review 99*(4), 1245–1277.

Attanasio, O. P. and V. Lechene (2014). Efficient responses to targeted cash transfers. *Journal of Political Economy 122*(1), 178–222.

Bloch, F. and V. Rao (2002). Terror as a bargaining instrument: A case study of dowry violence in rural india. *The American Economic Review 92*(4), 1029–1043.

Bobonis, G. (2009). Is the allocation of resources within the household efficient? new evidence from a randomized experiment. *Journal of Political Economy 117*(3), 453–503.

Browning, M. and P.-A. Chiappori (1998). Efficient intra-household allocations: A general characterization and empirical tests. *Econometrica*, 1241–1278.

Chen, J. J. (2013). Migration and imperfect monitoring: Implications for intra-household allocation. *The American Economic Review 96*(2), 227–231.

Dalhuisen, J. M., R. J. Florax, H. L. De Groot, and P. Nijkamp (2003). Price and income elasticities of residential water demand: a meta-analysis. *Land economics 79*(2), 292–308.

Duflo, E. and C. Udry (2004). Intrahousehold resource allocation in cote d'ivoire: Social norms, separate accounts and consumption choices. *NBER Working Paper Series w10498*.

Hoel, J. B. (2015). Heterogeneous households: A within-subject test of asymmetric information between spouses in kenya. *Journal of Economic Behavior and Organization 118*, 123–135.

Holmstrom, B. (1982). Moral hazard in teams. *The Bell Journal of Economics*, 324–340.

Ito, K. (2014). Do consumers respond to marginal or average price? evidence from nonlinear electricity pricing. *The American Economic Review 104*(2), 537–563.

Jacobsen, M., M. Webster, and K. Vairavamoorth (2013). The future of water in african cities. resreport, The World Bank.

Jessoe, K. and D. Rapson (2014). Knowledge is (less) power: Experimental evidence from residential energy use. *The American Economic Review 104*(4), 1417–1438.

Kahn, M. E. and F. A. Wolak (2013). Using information to improve the effectiveness of nonlinear pricing: Evidence from a field experiment. *Working Paper*.

Karlan, D. S. (2005). Using experimental economics to measure social capital and predict financial decisions. *The American Economic Review 95*(5), 1688–1699.

Kebede, B., M. Tarazona, A. Munro, and A. Verschoor (2013). Intra-household efficiency: An experimental study from ethiopia. *Journal of African Economies*, ejt019.

Levinson, A. and S. Niemann (2004). Energy use by apartment tenants when landlords pay for utilities. *Resource and Energy Economics 26*(1), 51–75.

Lundberg, S. and R. A. Pollak (1994). Noncooperative bargaining models of marriage. *The American Economic Review 84*(2), 132–137.

Lundberg, S. and R. A. Pollak (2008). Noncooperative bargaining models of marriage. *The American Economic Review 84*(2), 132–137.

Mack, E. A. and S. Wrase (2017). A burgeoning crisis? a nationwide assessment of the geography of water affordability in the united states. *PloS one 12*(1).

Mani, A. (2011). Mine, your or ours? the efficiency of household investment decision: An experimental approach. *Working Paper*.

McRae, S. and R. Meeks (2016). Price perception and electricity demand with nonlinear tariffs. *Working Paper*.

Myers, E. (2015). Asymmetric information in residential rental markets: Implications for the energy efficiency gap. *Working Paper* (246R).

National Water Supply and Sanitation Council, Zambia (2015). Strategic plan: 2016-2020. Technical report, NWASCO.

NWASCO (2014). Guidelines on tariff setting. Technical report, National Water Supply and Sanitation Council.

Ostrom, E. (2006). The value-added of laboratory experiments for the study of institutions and common-pool resources. *Journal of Economic Behavior and Organization 61*, 149–163.

Udry, C. (1996). Gender, agricultural production, and the theory of the household. *Journal of Political Economy 104*, 1010–1046.

Van der Bruggen, B., K. Borghgraef, and C. Vinckier (2010). Causes of water supply problems in urbanised regions in developing countries. *Water Resources Management 24*, 1885–1902.

Velez, M. A., J. K. Stranlund, and J. J. Murphy (2009). What motivates common pool resource users? experimental evidencefrom the field. *Journal of Economic Behavior & Organization 70*(3), 485–497.

Worthington, A. C. and M. Hoffman (2008). An empirical survey of residential water demand modelling. *Journal of Economic Surveys 22*(5), 842–871.

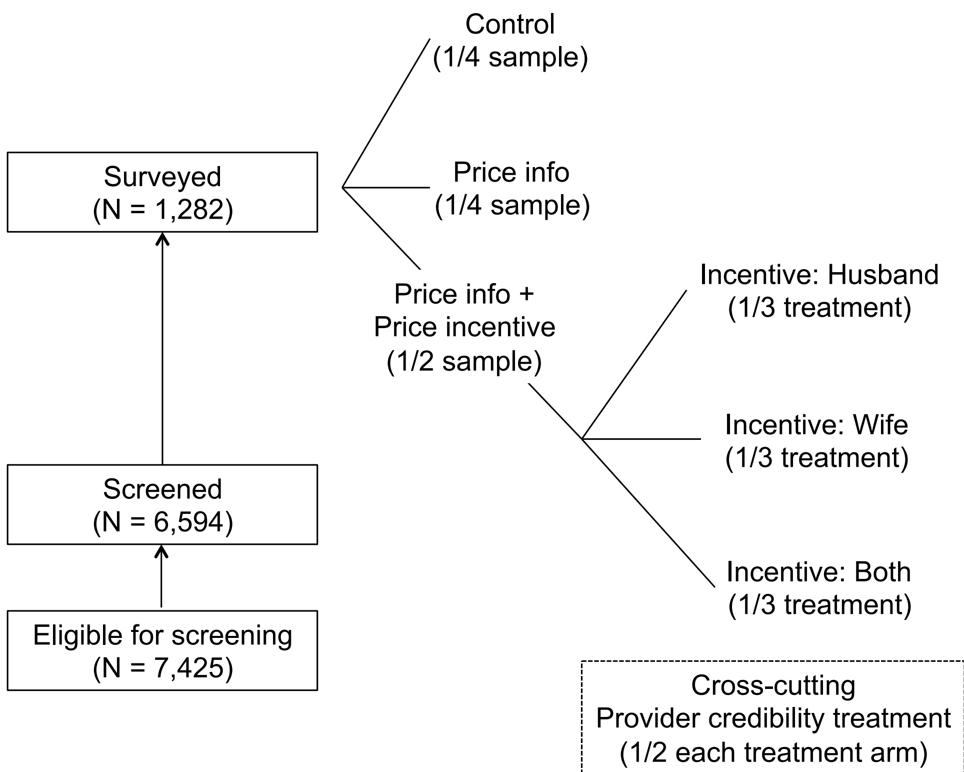

Figure 1: Experimental design

Notes: Experimental design, and sampling flow. Treatment was assigned on a rolling basis to accommodate the high rate of ineligibility that led screened households to fall out of the survey sample.

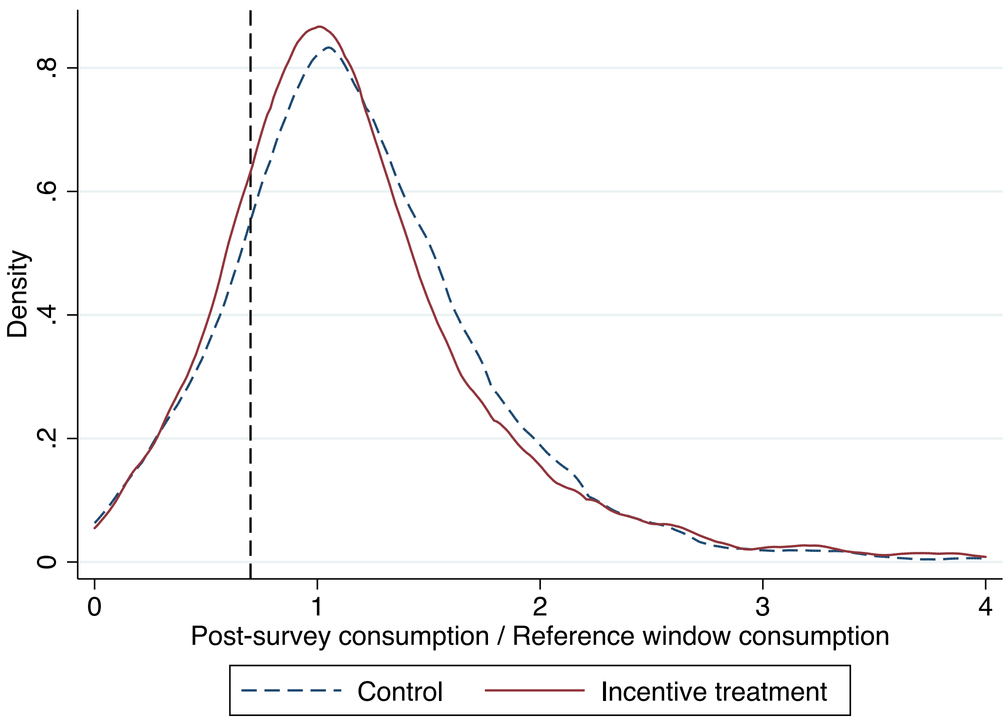

Figure 2: Consumption, relative to incentive reference month

Notes: Density plots of post-survey monthly consumption relative to the average monthly consumption in the reference months used to determine incentive lottery eligibility. The control group includes all surveyed households (i.e. screened out households are excluded) not assigned to the incentive treatment. The dashed vertical line shows the 70 percent threshold for lottery eligibility.

Table 1: Sample statistics & Balance

	Screened only (1)	No incentive (2)	Incentive (3)	P-val (2)=(3) (4)
Quantity consumed	20.940	18.995	18.247	0.239
	(14.525)	(12.097)	(10.515)	
Any payment	0.738	0.764	0.769	0.566
	(0.195)	(0.166)	(0.166)	
Missing meter reading	0.137	0.100	0.112	0.210
	(0.188)	(0.157)	(0.170)	
Total monthly bill	99.848	92.925	87.309	0.124
	(88.152)	(69.044)	(60.949)	
People in HH		5.860	5.888	0.822
		(2.286)	(2.218)	
HH has maid		0.169	0.149	0.333
		(0.375)	(0.356)	
Owns home		0.512	0.495	0.546
		(0.500)	(0.500)	
Rooms in HH		3.529	3.553	0.751
		(1.264)	(1.444)	
Both read and write English		0.768	0.772	0.873
		(0.422)	(0.420)	
Either underestimated price		0.619	0.637	0.551
		(0.486)	(0.481)	
Both blame high bill on SWSC		0.440	0.414	0.356
		(0.497)	(0.493)	
Share sent to spouse by husband		0.702	0.690	0.398
		(0.269)	(0.254)	
Share sent to spouse by wife		0.520	0.513	0.597
		(0.262)	(0.260)	
Share sent to NGO by husband		0.312	0.303	0.522
		(0.253)	(0.232)	
Share sent to NGO by wife		0.275	0.276	0.923
		(0.222)	(0.220)	
W: Residual claimant		0.307	0.316	0.749
		(0.462)	(0.465)	
W: Bigger user		0.795	0.838	0.047
		(0.404)	(0.369)	
Aligned usage incentives		0.027	0.016	0.181
		(0.163)	(0.126)	
Households	5312	664	618	

Notes: Pre-treatment means for all households (top panel), household-level statistics for surveyed households only (second panel), and measures of dictator game decisions and survey responses for surveyed households only (bottom panel). Column (1) shows averages for households screened out of the survey sample, column (2) for the survey sample that did not receive the incentive treatment, and column (3) for the survey sample that did receive the incentive treatment. Column (4) reports the p-value for a test of equal means between columns (2) and (3). The quantity consumed is measured in cubic meters per month. H and W refer to husband and wife. The share sent to the spouse is measured as a share of the respondent's endowment.

Table 2: Correlates of dictator game decisions

	Husband share sent (1)	Wife share sent (2)	Sent above median (3)
H: Share NGO	0.192***	0.079***	0.262***
	(0.030)	(0.030)	(0.057)
W: Share NGO	0.034	0.198***	0.269***
	(0.033)	(0.033)	(0.063)
H: High SDB score	0.020	0.010	0.016
	(0.015)	(0.015)	(0.028)
W: High SDB score	0.009	-0.011	-0.005
	(0.015)	(0.015)	(0.028)
W: Residual claimant	0.003	0.001	0.013
	(0.016)	(0.016)	(0.030)
W: Bigger water user	0.003	-0.006	-0.000
	(0.019)	(0.019)	(0.036)
Household size	-0.004	-0.009***	-0.015**
	(0.003)	(0.003)	(0.006)
H: Age 51+	0.001	-0.069***	-0.074**
	(0.015)	(0.015)	(0.029)
W: Age 51+	0.004	-0.061***	-0.068**
	(0.017)	(0.017)	(0.033)
HH has maid	0.019	0.068***	0.086**
	(0.020)	(0.020)	(0.038)
HH assets	0.008***	0.018***	0.030***
	(0.003)	(0.003)	(0.005)
HH owns home	-0.015	-0.034**	-0.060**
	(0.015)	(0.015)	(0.028)
HH rooms in home	0.014***	0.015***	0.034***
	(0.005)	(0.005)	(0.010)
HH english fluency	0.022	0.082***	0.112***
	(0.017)	(0.017)	(0.033)
Both know bill quantity	0.005	0.014	0.046
	(0.022)	(0.022)	(0.043)
Both know bill charge	0.010	-0.004	-0.004
	(0.016)	(0.016)	(0.030)
Either underestimated price	0.009	0.017	0.033
	(0.017)	(0.017)	(0.032)
Both blame high bill on SWSC	0.012	0.016	0.000
	(0.015)	(0.015)	(0.028)
Aligned usage incentives	-0.012	0.029	-0.045
	(0.050)	(0.050)	(0.095)

Notes: Pairwise correlation coefficients for each household characteristic and each dictator game measure. The share sent to the spouse and share sent to the NGO are measured as a fraction of the respondent's endowment. The SDB score measures social desirability bias using an adapted Crowne-Marlowe (1964) instrument. H and W refer to husband and wife.

Table 3: Average treatment effects

	log (quantity) (1)	log (quantity) (2)	log (quantity) (3)
Incentive x Post	-0.067**	-0.062**	-0.063**
	[0.030]	[0.029]	[0.029]
Information x Post	-0.019	-0.011	-0.008
	[0.034]	[0.033]	[0.033]
Credibility x Post	0.026	0.025	0.024
	[0.026]	[0.025]	[0.024]
Surveyed x Post	0.051*	0.026	0.014
	[0.029]	[0.028]	[0.027]
Incentive treatment	-0.013		
	[0.041]		
Information treatment	0.008		
	[0.047]		
Provider credibility treatment	-0.011		
	[0.033]		
Surveyed	-0.086**		
	[0.038]		
HH FE		x	x
Month-Year FE			x
Observations (HH)	6,594	6,594	6,594
Observations (HH-months)	129,899	129,899	129,899

Notes: The panel begins 20 months (billing cycles) prior to the month of the survey and ends 4 billing cycles after the survey. Since treatment was provided at the survey, we use the recorded survey date to define the treatment variables. The HH assignment indicator is consistent across the panel, while the *Post-survey* indicator switches from 0 to 1 in the first full billing cycle after the date the household was surveyed; observations for billing cycles that contain the survey date are dropped. All households in the incentive treatment also received the information treatment. Standard errors are clustered at the at the household level.

Table 4: Price beliefs and provider trust

	log (quantity) (1)	log (quantity) (2)
Info treatment	-0.012	
	[0.055]	
Info x Underestimated price	-0.020	
	[0.073]	
Provider credibility treatment		0.005
		[0.033]
Provider credibility x Distrust billing		0.046
		[0.049]
Total effect	-0.032	0.050
	[0.048]	[0.036]
HH FE	x	x
Month-year FE	x	x
Observations (HH)	6,337	6,594
Observations (HH-months)	124,826	129,899

Notes: Regressions include the time-varying survey indicator and the post indicator interacted with the heterogeneity variables. *Underestimated price* equals one if either spouse underestimated the marginal price of water. *Distrust billing* equals one if both spouses blame a high water bill on the provider. The bottom panel reports the linear combination of the treatment effect and the interaction term. Standard errors are clustered at the at the household level. Price beliefs are missing for either the husband or wife in 257 households.

Table 5: Heterogeneity in household efficiency

	log (quantity) (1)	log (quantity) (2)	log (quantity) (3)	log (quantity) (4)
Incentive treatment	-0.034	-0.066***	-0.058*	-0.093**
	[0.032]	[0.025]	[0.033]	[0.044]
Incentive x Sent above median on average	-0.071			
	[0.050]			
Incentive x Aligned incentives		-0.051		
		[0.177]		
Incentive x Husband sent above median			-0.022	-0.011
			[0.051]	[0.066]
Incentive x Wife sent above median			-0.003	-0.079
			[0.058]	[0.075]
Total effect	-0.105***	-0.117		
	[0.038]	[0.175]		
Total effect, husband			-0.079*	-0.104*
			[0.043]	[0.055]
Total effect, wife			-0.061	-0.172**
			[0.055]	[0.072]
Sample	full	full	full	gender roles
HH FE	x	x	x	x
Month-year FE	x	x	x	x
Observations (HH)	6,587	6,594	6,587	6,038
Observations (HH-months)	129,775	129,899	129,775	118,452

Notes: Regressions include the time-varying survey indicator and the post indicator interacted with the heterogeneity variables. *Shared above median* equals one if the share of the endowment transferred in the dictator game was above the median. *Aligned incentives* equals one if the residual claimant on the water bill is also the bigger water user. Column 4 restricts the sample to households in which the woman is the larger water user and the man is the residual claimant on the bill. The bottom panel reports the linear combination of the treatment effect and the interaction term. Standard errors are clustered at the at the household level. Dictator game outcomes are missing for at least one member of the couple in 7 households.

Table 6: Average treatment effects: Price incentive sub-treatments

	log (quantity)
	(1)
Couple incentive	-0.050
	[0.041]
Husband incentive	-0.043
	[0.037]
Wife incentive	-0.095**
	[0.037]
Couple = Husband (p-val)	0.887
Couple = Wife (p-val)	0.323
Husband = Wife (p-val)	0.222
HH FE	x
Month-Year FE	x
Observations (HH)	6,594
Observations (HH-months)	129,899

Notes: Regressions include the time-varying survey indicator and the post indicator. The bottom panel reports p-values for tests of equal coefficients across teh sub-treatment arms. Standard errors are clustered at the at the household level.

Table 7: Heterogeneity by lottery sub-treatments

	log (quantity) (1)	log (quantity) (2)	log (quantity) (3)	log (quantity) (4)
Individual incentive	-0.030	-0.054	-0.026	-0.024
	[0.033]	[0.033]	[0.035]	[0.036]
Incentive to non-residual claimant	-0.091**		-0.087*	-0.084*
	[0.043]		[0.046]	[0.046]
Incentive to bigger user		-0.045	-0.013	
		[0.043]	[0.046]	
Wife incentive				-0.018
				[0.046]
Total effect claimant	-0.121***		-0.113**	-0.108**
	[0.035]		[0.046]	[0.048]
Total effect user		-0.099***	-0.039	
		[0.036]	[0.048]	
HH FE	x	x	x	x
Month-year FE	x	x	x	x
Observations (HH)	6,412	6,412	6,412	6,412
Observations (HH-months)	126,136	126,136	126,136	126,136

Notes: Regressions include the time-varying survey indicator and the post indicator interacted with the heterogeneity variables. The bottom panel reports the linear combination of the treatment effect and the interaction term. Standard errors are clustered at the at the household level. The couple incentive treatment arm is excluded.

Table 8: Information and monitoring

	log (quantity) (1)	log (quantity) (2)	log (quantity) (3)	log (quantity) (4)
Incentive treatment	-0.090***	-0.055**	-0.053	-0.038
	[0.031]	[0.027]	[0.045]	[0.030]
Incentive x Look at meter	0.067			
	[0.051]			
Incentive x Know bill quantity		-0.081		
		[0.068]		
Incentive x Know bill charge			-0.020	
			[0.053]	
Incentive x Know spouse's water use				-0.080
				[0.051]
Total effect	-0.024	-0.137**	-0.074**	-0.118***
	[0.041]	[0.063]	[0.029]	[0.041]
HH FE	x	x	x	x
Month-year FE	x	x	x	x
Observations (HH)	6,594	6,594	6,594	6,594
Observations (HH-months)	129,899	129,899	129,899	129,899

Notes: Regressions include the time-varying survey indicator and the post indicator interacted with the heterogeneity variables. *Look at meter* equals one if both spouses report looking at their water meter. *Know bill quantity* and *Know bill charge* equal one if both spouses can identify the quantity and total amount owed on the bill, respectively. *Know spouse's water use* equals one if both spouses know above the median share of their spouses primary water using activities. The bottom panel reports the linear combination of the treatment effect and the interaction term. Standard errors are clustered at the at the household level.

Table 9: Robustness checks: Measurement

	log (quantity) (1)	log (quantity) (2)	log (quantity) (3)
Incentive x Sent above median	-0.071		-0.060
	(0.050)		(0.052)
Incentive x Sent above median to NGO	-0.069	-0.037	-0.034
	(0.049)	(0.050)	(0.050)
Incentive x Above median SDB score	0.034	0.024	0.030
	(0.049)	(0.050)	(0.050)
Incentive x Household size	-0.005	-0.006	-0.006
	(0.011)	(0.011)	(0.011)
Incentive x Both spouses 51+	0.067	0.029	0.026
	(0.054)	(0.058)	(0.058)
Incentive x Maid	-0.046	-0.003	-0.001
	(0.068)	(0.069)	(0.069)
Incentive x HH assets	-0.017*	-0.015	-0.013
	(0.009)	(0.010)	(0.011)
Observations (HH)	6,587	6,587	6,587
Observations (HH-months)	129,775	129,775	129,775

Notes: Column 1 show separate regressions in each cell, where each of the household level characteristics is interacted with treatment and the time varying survey variables. Columns 2 and 3 each correspond to a single regression. Standard errors are clustered at the at the household level.

Table 10: Robustness checks: Other outcomes

	Any pay (1)	Any pay (2)	Missing quant (3)	Missing quant (4)
Incentive	0.010	0.011	-0.003	-0.005
	[0.014]	[0.014]	[0.008]	[0.008]
Price info	0.007	0.006	0.000	0.002
	[0.016]	[0.016]	[0.009]	[0.009]
SWSC credibility	0.011	0.011	-0.012*	-0.012*
	[0.011]	[0.011]	[0.007]	[0.006]
Surveyed	0.004	0.002	0.011	0.010
	[0.014]	[0.014]	[0.007]	[0.007]
HH FE	x	x	x	x
Month-Year FE		x		x
Observations (HH)	6,594	6,594	6,594	6,594
Observations (HH-months)	140,431	140,431	152,971	152,971

Notes: Regressions replicate columns (2) and (3) of Table 3, with different outcomes. Columns (1) and (2) report the probability that a household made any payment during the month. Columns (3) and (4) report the probability that log(quantity) is missing, conditional on the household having received a prior non-missing meter reading.

Appendix

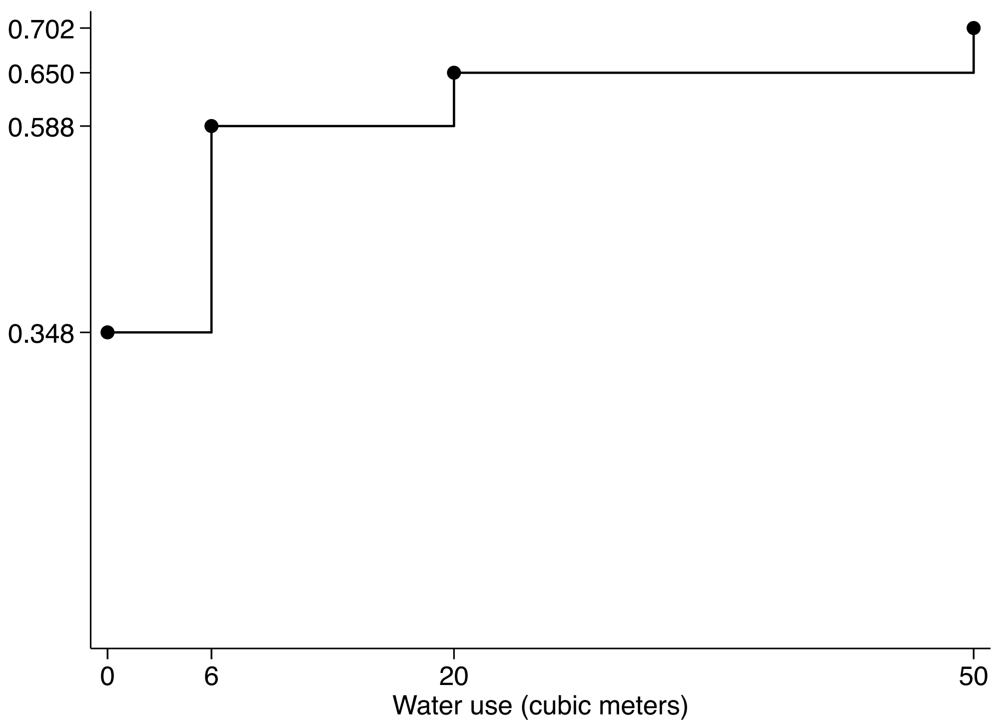

Figure A.1: 2015 tariff schedule

Notes: Increasing block tariff for residential piped water in Livingstone. The price is shown in 2015 USD per cubic meter and is increasing in cumulative consumption over the course of the month.

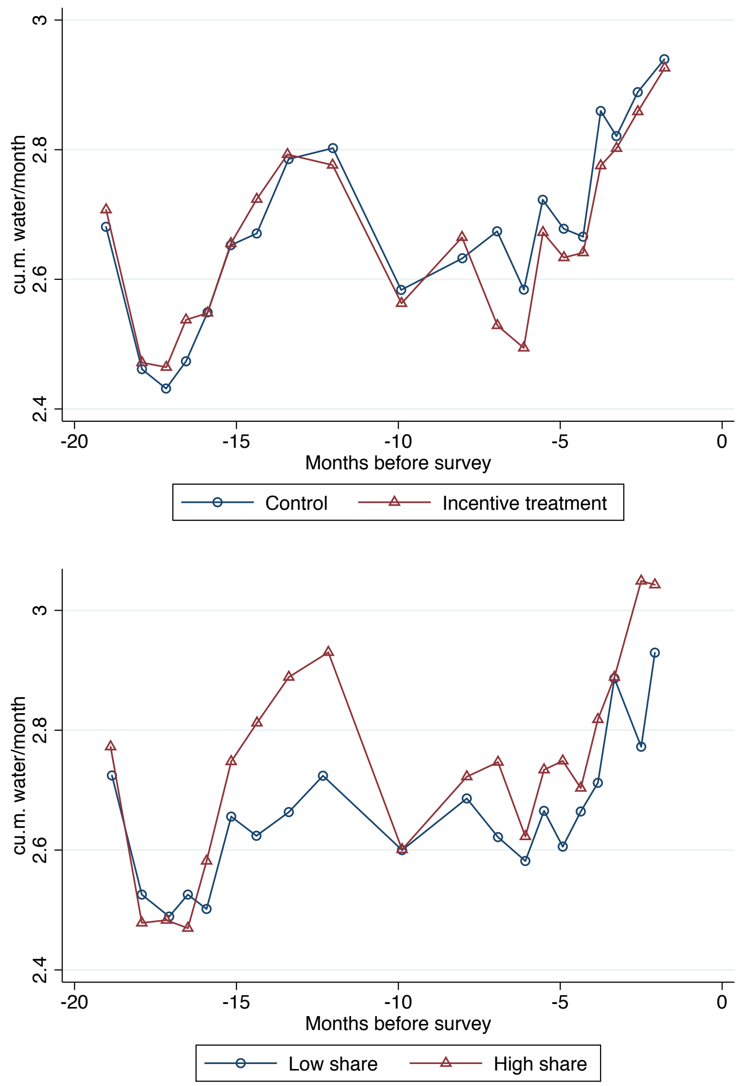

Figure A.2: Pre-treatment water use

Notes: Average monthly water consumption in the months preceding the survey. The top figure splits the surveyed sample into those receiving the incentive treatment and the control. The bottom figure splits the surveyed sample into households with a measured level of efficiency below (low share sent in the dictator game) and above the median (high share sent). The plots control for seasonal (calendar month) effects.

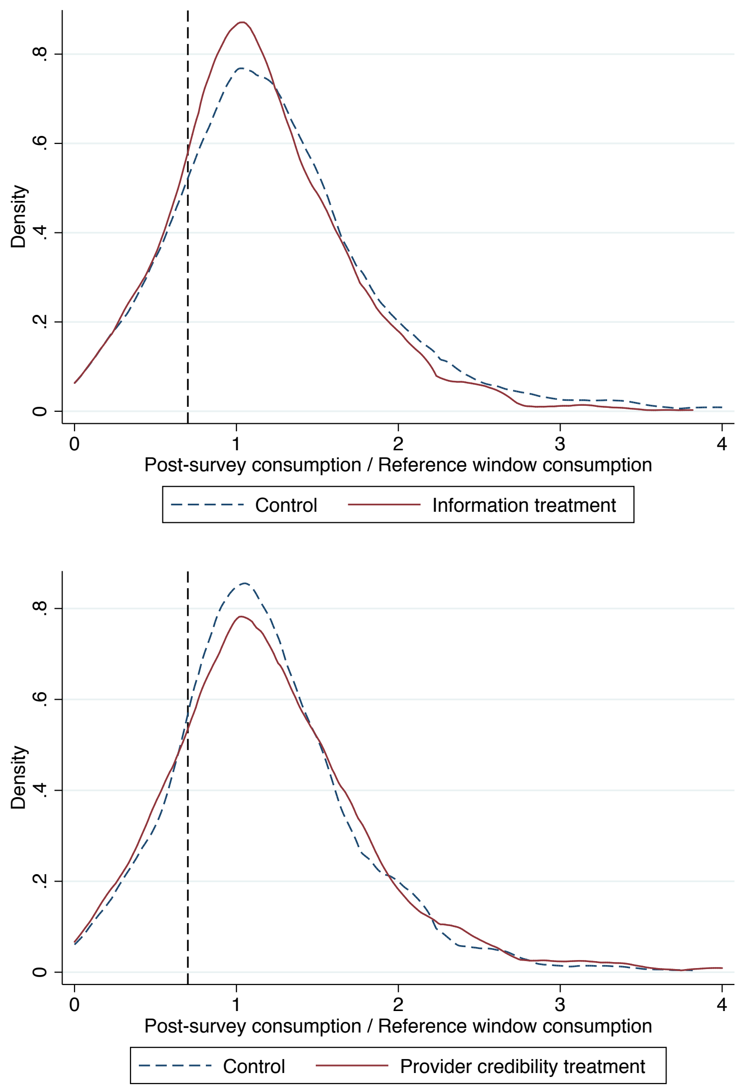

Figure A.3: Consumption, relative to incentive reference months

Notes: Density plots of post-survey monthly consumption relative to the average monthly consumption in the reference months used to determine incentive lottery eligibility. The top figure compares households with and without price information. The bottom figure compares households with and without the provider credibility treatment. Note that the incentive treatment is excluded from these plots. The dashed vertical line shows the 70 percent threshold for lottery eligibility.

Table A.1: Balance: Information treatment

	Screened only (1)	No info (2)	Info treatment (3)	P-val (2)=(3) (4)
Quantity consumed	20.940	18.927	18.531	0.584
	(14.525)	(12.255)	(11.037)	
Any payment	0.738	0.772	0.765	0.500
	(0.195)	(0.159)	(0.168)	
Missing meter reading	0.137	0.100	0.108	0.439
	(0.188)	(0.165)	(0.163)	
Total monthly bill	99.848	91.720	89.689	0.625
	(88.152)	(68.509)	(64.163)	
Households	5312	334	948	
People in HH		5.832	5.888	0.697
		(2.375)	(2.209)	
HH has maid		0.165	0.157	0.748
		(0.371)	(0.364)	
Owns home		0.491	0.508	0.584
		(0.501)	(0.500)	
Rooms in HH		3.485	3.560	0.386
		(1.249)	(1.388)	
Both read and write English		0.778	0.767	0.666
		(0.416)	(0.423)	
Either underestimated price		0.557	0.653	0.005
		(0.498)	(0.476)	
Both blame high bill on SWSC		0.428	0.427	0.976
		(0.496)	(0.495)	
Share sent to spouse by husband		0.702	0.694	0.660
		(0.273)	(0.257)	
Share sent to spouse by wife		0.526	0.513	0.458
		(0.263)	(0.260)	
Share sent to NGO by husband		0.320	0.303	0.260
		(0.251)	(0.241)	
Share sent to NGO by wife		0.285	0.272	0.349
		(0.224)	(0.220)	
Both agree that H is residual claimant		0.497	0.489	0.812
		(0.501)	(0.500)	
W: Bigger user		0.766	0.833	0.007
		(0.424)	(0.373)	
Aligned usage incentives		0.036	0.017	0.041
		(0.186)	(0.129)	
Households		334	948	

Notes: See Table 1 for a description of the table set up. Column (1) shows averages for households screened out of the survey sample, column (2) for the survey sample that did not receive the information treatment, and column (3) for the survey sample that did receive the information treatment. Column (4) reports the p-value for a test of equal means between columns (2) and (3).

Table A.2: Balance: Provider credibility treatment

	Screened only (1)	No credibility (2)	Credibility treatment (3)	P-val (2)=(3) (4)
Quantity consumed	20.940	18.884	18.379	0.426
	(14.525)	(11.830)	(10.870)	
Any payment	0.738	0.768	0.766	0.822
	(0.195)	(0.166)	(0.166)	
Missing meter reading	0.137	0.098	0.114	0.096
	(0.188)	(0.157)	(0.170)	
Total monthly bill	99.848	91.859	88.547	0.364
	(88.152)	(67.023)	(63.510)	
Households	5312	647	635	
People in HH		5.740	6.009	0.032
		(2.185)	(2.314)	
HH has maid		0.159	0.159	0.994
		(0.366)	(0.366)	
Owns home		0.491	0.517	0.370
		(0.500)	(0.500)	
Rooms in HH		3.519	3.562	0.564
		(1.362)	(1.345)	
Both read and write English		0.782	0.757	0.296
		(0.413)	(0.429)	
Either underestimated price		0.595	0.660	0.031
		(0.491)	(0.474)	
Both blame high bill on SWSC		0.427	0.428	0.949
		(0.495)	(0.495)	
Share sent to spouse by husband		0.698	0.694	0.792
		(0.259)	(0.264)	
Share sent to spouse by wife		0.520	0.514	0.671
		(0.267)	(0.254)	
Share sent to NGO by husband		0.313	0.302	0.440
		(0.254)	(0.231)	
Share sent to NGO by wife		0.279	0.272	0.568
		(0.226)	(0.217)	
Both agree that H is residual claimant		0.464	0.520	0.045
		(0.499)	(0.500)	
W: Bigger user		0.802	0.830	0.200
		(0.399)	(0.376)	
Aligned usage incentives		0.022	0.022	0.960
		(0.146)	(0.147)	
Households		647	635	

Notes: See Table 1 for a description of the table set up. Column (1) shows averages for households screened out of the survey sample, column (2) for the survey sample that did not receive the provider credibility treatment, and column (3) for the survey sample that did receive the credibility treatment. Column (4) reports the p-value for a test of equal means between columns (2) and (3).

Table A.3: Balance at each sampling stage

	Coefficient on:		
	Successfully screened (1)	Screened in (2)	Surveyed (3)
Price information treatment	0.008	-0.025**	0.011
	[0.016]	[0.012]	[0.020]
SWSC trust treatment	-0.042**	-0.001	-0.005
	[0.018]	[0.013]	[0.023]
Lottery treatment	-0.001	-0.017	-0.011
	[0.018]	[0.013]	[0.023]
Lottery to couple	0.002	-0.019*	-0.026
	[0.014]	[0.010]	[0.017]
Lottery to woman	-0.000	-0.005	0.008
	[0.014]	[0.010]	[0.017]
Lottery to man	-0.003	0.007	0.006
	[0.014]	[0.010]	[0.017]
Mean of regressor	0.888	0.311	0.608
Sample	Eligible for screening	Successfully screened	Screened in
Observations	7,425	6,594	2,051

Notes: Each cell of columns 1-3 represents one regression, with each column representing a binary regressor. The sample for regressions in each stage is conditional on surviving the previous stage. See text for further detail.

Table A.4: Consumption correlates

	Avg quantity (1)	Avg quantity (2)
Sent above median to spouse	2.190***	0.918
	(0.634)	(0.577)
Sent above median to NGO		1.074*
		(0.580)
Above median SDB score		0.516
		(0.581)
Household size		0.572***
		(0.127)
Both spouses 51+		3.362***
		(0.698)
HH has maid		6.465***
		(0.818)
HH assets (count)		1.370***
		(0.115)
Observations (HH)	1,275	1,275

Notes: Cross sectional regression of pre-intervention monthly water use on an indicator for whether the households sent above median in the dictator game (column 1) and other observables (column 2). Only surveyed households are included in the specification.

Table A.5: Robustness checks: Dictator game aggregation

	log (quantity) (1)	log (quantity) (2)	log (quantity) (3)	log (quantity) (4)
Incentive treatment	-0.034	0.006	-0.057	0.039
	[0.032]	[0.080]	[0.056]	[0.082]
Incentive x Above median	-0.071			
	[0.050]			
Incentive x Avg sent		-0.122		
		[0.130]		
Incentive x Minimum sent			-0.022	
			[0.115]	
Incentive x Max sent				-0.139
				[0.106]
HH FE	x	x	x	x
Month-Year FE	x	x	x	x
Observations (HH)	6,587	6,587	6,587	6,587
Observations (HH-months)	129,775	129,775	129,775	129,775

Notes: Regressions replicate the specification in column 1 of Table 5 showing different measures from the dictator game. Column 1 is the same as Table 5. Column 2 shows a continuous measure of the dictator game endowment share sent to spouse. Columns 3 and 4 show the minimum and maximum share of endowment sent by either spouse.

Table A.6: Robustness checks: Sample

Sample restriction:	Surveyed only		All treated months	
	(1)	(2)	(3)	(4)
Incentive treatment	-0.067***	-0.035	-0.057**	-0.026
	[0.024]	[0.032]	[0.025]	[0.032]
Incentive x Sent above median		-0.069		-0.062
		[0.049]		[0.049]
Total effect		-0.104***		-0.088**
		[0.038]		[0.038]
HH FE	x	x	x	x
Month-year FE	x	x	x	x
Observations (HH)	1,282	1,275	6,594	6,587
Observations (HH-months)	26,246	26,122	136,573	136,442

Outcome is:	quantity		log (total bill)	
	(1)	(2)	(3)	(4)
Incentive treatment	-1.241**	-0.564	-0.067***	-0.032
	[0.493]	[0.589]	[0.024]	[0.032]
Incentive x Sent above median		-1.458		-0.076
		[1.011]		[0.049]
Total effect		-2.022**		-0.108***
		[0.822]		[0.037]
HH FE	x	x	x	x
Month-year FE	x	x	x	x
Observations (HH)	6,594	6,587	6,594	6,587
Observations (HH-months)	129,899	129,775	129,899	129,775

Notes: Regressions replicate the specification in column 3 of Table 3 and column 1 of Table 5 excluding the screened-out sample.