Online Appendix to
Distortionary Fundraising for Energy Efficiency Subsidies:
Implications for Efficient and Equitable Program Design

Chris Bruegge*

This Version: November 9, 2017

*The main text is available online at http://stanford.edu/~cbruegge/jmp.pdf

Department of Economics, Stanford University, 579 Serra Mall, Stanford, CA, 94305.
Website: http://stanford.edu/~cbruegge/ E-mail: cbruegge@stanford.edu
A1 Legal Background for E.E. Programs

By law in many states, money to support energy efficiency (E.E.) subsidies is collected through an increase in the marginal price of electricity. Several noteworthy examples include

- California Public Utility Code, §381, §381.5 (Electricity), §890, §892 (Gas)
- Massachusetts Code, TITLE II: Section 19.
- Vermont Statutes, §209

In some states, an increase to the volumetric price of electricity is not mandated by law, but utility companies still recover costs for energy efficiency programs in this manner. For example, Rocky Mountain Power, which operates in Idaho, Utah, and Wyoming, charges a percentage-based increase for the variable component (energy and transmission) of customer bills in order to fund it’s efficiency programs. Note that this is not a feature of utilities in liberal, coastal states. Even in Wyoming, ranked 50th by the ACEEE on the state energy efficiency scorecard, allows RMP to charge a 1% surcharge on energy consumption in order to fund energy efficiency programs.

A2 Back of the Envelope Savings Decomposition

To corroborate the evidence I present in the paper on the importance of the fundraising distortion, this section performs a back of the envelope savings decomposition to determine the fraction of reductions in energy use due to the appliance subsidy and the energy price change respectively. The calculation in this section is intended to represent the cumulative effect of all energy efficiency programs, while in the paper I focus on two particular programs. First, I will describe the inputs needed to estimate energy saved by the Energy Star appliances purchased because of the program, and second I will sketch the parameters that are relevant to compute energy saved by the higher electricity prices that are associated with the program. The goal of this section is to convince the reader that the importance of the energy price change is independent of modeling assumptions I make later in the paper. As I discuss the inputs to this decomposition, I’ll also point out how the energy price change might bias existing estimates of the relevant parameters.

Savings from Energy Star Appliance Adoption

To estimate how much energy is saved by more energy efficient appliances adopted because of all energy efficiency programs in this utility territory, a natural starting place is to multiply the number of participants by the average energy

---

1 The utility’s filing with the Idaho Public Utilities Comission to change to a percentage-based rate can be found here.
2 There are 51 spots since Washington DC is included, and North Dakota was last place in 2016. More information on the ACEEE Energy Efficiency Scorecard can be found here.
saved per participant. Two problems arise with this method, first many participants would have purchased an Energy Star appliance even without the rebate, and second, only the average savings rate should be computed for this selected group of marginal participants, not for the average participant. Note that the change in the energy price can affect the number of marginal participants as well as the composition of marginal participants.

1. Inframarginal Participation: Since only the marginal customers change their behavior in response to the policy, energy reductions because of inframarginal participants should not be counted towards savings caused by the program. Houde and Aldy (2014) estimate that more than 90% of participants in a suite of federal appliance subsidies were inframarginal (revenue for this program was from federal stimulus money, not higher energy prices), and Boomhower and Davis (2014) find significant inframarginal participation in a program in Mexico. It’s possible that these marginal households made their Energy Star purchase both because of the subsidy and because of the generally higher energy prices which makes energy efficiency more attractive, although only the former effect is captured by existing studies.

2. Selection: Households that purchase Energy Star appliances are different than non-purchasers in many ways. Furthermore, those whose Energy Star appliance purchase is marginal to the subsidy might differ from those who would have purchased an Energy Star appliance even without the subsidy. For example, participants might have a strong incentive to invest in energy efficient appliances because they have a large demand for energy use. Thus selection into program participation could potentially bias estimates of energy savings because a comparison of participants and non-participants conflates both appliance ownership and underlying preferences for energy consumption.

After adjusting for inframarginal participation and selection, a more refined representation of total savings from the appliance subsidy is given by

\[
\Delta kWh_{\text{Month}} = \left( \frac{\text{Savings per Marginal Participant}}{\text{Marginal Participants}} \right) \times \left( \text{Number of Marginal Participants} \right)
\] (A2.1)

Total savings caused by the program is simply given by multiplying the number of households whose purchase was caused by the subsidy with the amount of energy saved by these households as a result of their Energy Star appliance purchase. Several experimental evaluations such as Fowlie et al. (2015) measure marginal participation and savings with respect to an experimental price change. Although their strategy measures the effect on the households marginal to the

---

3 Most program evaluators now make a net-to-gross correction, but in practice there isn’t a great way to estimate this.

4 Marginal households in a given month are those who own an Energy Star appliance, but who would own a conventional appliance in that month if the subsidy program didn’t exist. If some marginal households simply shifted their purchase forward in time, then the expression below gives an upper bound on the savings per month.
experiment, the households marginal to the experimental incentives are not necessarily the same as those marginal to the actual program incentives. Experiments move the appliance price incentive by a different amount than the actual program, and furthermore existing experimental evaluations haven’t manipulated the energy price.

If the participants in a typical program reduce their energy use by 5% per month and 10% of participants are marginal to the subsidy incentive, one can convert Equation A2.1 into percent savings per month to show that on average across all program participants, there is a causal $0.05 \cdot 0.10 = 0.5\%$ reduction in energy use per month. In my utility, less than 2% of households participate in an energy efficiency subsidy per year ($< 0.17\%$ per month), so on average across all customers in the entire utility service territory there is a $0.005 \cdot 0.0017 = 0.0085\%$ reduction in energy consumption due to the more efficient appliances purchased because subsidy programs that would not have been purchased otherwise. In Table I I make the same computation for variety of different parameter inputs taken from estimates in the literature.

### Savings from Energy Price Changes

Energy savings caused by energy price changes can be computed by multiplying the price elasticity of demand for energy by the price change and the baseline energy consumption. The price increases for all consumers regardless of whether they claim an appliance rebate, and furthermore the price change affects consumption from all appliances, not just the appliance category subsidized through the program.

$$\text{Price Channel: } \frac{\Delta \text{kWh}}{\text{Month}} = (\% \text{ Price Change}) \cdot (\text{Price Elasticity}) \cdot (\text{Baseline Consumption}) \cdot (\text{Total Households}) \quad (A2.2)$$

Using an estimate of $-0.16$ for the price elasticity of demand for electricity and the 3.3% price increase caused by energy efficiency subsidies in my utility territory, I obtain that on average households save $0.16 \cdot 0.033 = 0.53\%$ per month as a result of the higher electricity price. Notice that I haven’t multiplied by the baseline consumption or the total number of households so that this number is comparable to savings from the appliance channel.

---

5The regulator breaks down the marginal price of electricity into components such as generation, transmission, distribution, and public purpose charges, and on average 3.3% of the marginal price is used to fund energy efficiency programs (37% of the public purpose category that accounts for 9% of the marginal price).

6It’s natural to ask what the counterfactual world would look like with no energy efficiency subsidy programs. There is a large literature in public finance on the flypaper effect and the fungibility of funds, but in this context I think the counterfactual is clearer. The utility has no incentive to offer public purpose programs in the absence of regulation, and the regulator requires it to spend a given budget on these programs each year. Since it passes these costs on to customers, the price increases by the exact amount needed to pay for the programs. The only dimension along which the utility could change the counterfactual is on how it passes charges on to different customer classes in terms of actual incidence, not statutory incidence.

7Costs to the utility for subsidizing energy efficiency are passed on to customers each year. At the beginning of the year, projected needs are assessed and incorporated into rates, and at the end of each year there is a true-up adjustment to rates the following year based on over or under-funding. Public purpose programs are included in the Energy Resource Recovery Account (ERRA) proceedings, and recuperated the same way generation costs are recovered CPUC (2016).
Using the estimates above, 98% = 100 \cdot 0.53/(0.53+0.0085) of the reductions in energy consumption occur through the higher energy prices that result from the policy. Table 1 shows sensitivity to a range values for these parameter, and regardless of the inputs, the energy price change is more than half of the total effect of the program.

### Table 1: Back of the Envelope

<table>
<thead>
<tr>
<th>% Savings / Household · Month</th>
<th>Appliance Channel</th>
<th>Energy Price Channel</th>
<th>Energy Price % of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conservative</td>
<td>0.125</td>
<td>0.5</td>
<td>79.4%</td>
</tr>
<tr>
<td>Midrange</td>
<td>0.013</td>
<td>1.0</td>
<td>98.7%</td>
</tr>
<tr>
<td>High</td>
<td>0.001</td>
<td>2.5</td>
<td>&gt;99.9%</td>
</tr>
</tbody>
</table>

Notes: Baseline consumption is roughly 700kWh / month per household. The conservative estimate uses an own price elasticity of -.08, a savings rate of 10% per participant from Fowlie et al (2017), and 50% marginal participants. The savings rate of 10% for the weatherization is likely to be substantially higher than the savings rate for the average program, which is the correct number to include here. Midrange estimates are -.16, 5%, and 10%, and high estimates are -.39, 1%, and 3% for the price elasticity, savings per household, and fraction of marginal participants respectively. My estimates later in the paper apply to the washing machine and the refrigerator program, not the average program.

### A3 Consumer Utility Maximization

As described in the text, the household solves the following utility maximization problem:

\[
\max_{(j,s)} U(j, s_{ijt}) = n_{i0} + \text{FE}_{ij}(\theta_i) + \sigma \varepsilon_{ij} \left[ \sum_t \delta^t \left( n_{it} + \frac{1}{2\beta_i} (s_{ijt} - \alpha_i - \nu_{ijt})^2 \right) \right]
\]

subject to

\[
\begin{align*}
I_{i0} & \geq n_{i0} + p_{ij}^A, \\
I_{it} & \geq n_{it} + p_{ij}^S \cdot s_{ijt}, \quad t \in \{1, \ldots, 120\}.
\end{align*}
\]

The perceived price should enter its budget constraints in Equation A3.1 even if this isn’t actually the price it faces. This distinction only arises because electricity rates follow an increasing block price, meaning that the marginal price is an increasing function of the amount of electricity used in a month. I will formulate the household optimization problem separately for three different perceptions of the price schedule: (1) the household could perceive the correct marginal cost but not the prices on other steps of the electricity tariff, (2) the household could respond to the average price (perhaps the average price it faced in the previous month), or (3) the household could respond
to the full nonlinear price schedule. Ito (2017) demonstrates that household perceive the marginal price schedule in cases such as simple schedules with a fixed charge and a flat marginal price. Ito (2014) provides empirical evidence for optimization with respect to several notions of average price. Finally, work such as Reiss and White (2005) allows households to optimize with respect to the full price schedule.

A3.1 Optimization w.r.t. Marginal Price or Average Price

The theoretical treatment of optimization with respect to the marginal price and the average price in the previous month are equivalent, so I will address both cases in this section. If the household perceives the marginal price (or average price from the previous month), then

\[ P_{jt}^S = p_j^S \cdot s_t, \]

where \( p_j^S \) is the marginal price it faces in month \( t \) (average price it faces in \( t - 1 \)).

Since Note that with quasilinear utility, the household chooses the optimal level of energy service consumption and spends leftover income on the numeraire good. This means that even though the household might spend more ex-post on energy than it had planned, I don’t have a situation where the combination of energy service consumption and numeraire consumption are infeasible given actual (as opposed to perceived) prices.

Writing out the Lagrangian for the optimization problem above (conditional on \( j \)), we obtain

\[
\mathcal{L} = n_0 + F \mathcal{E}_j(\theta) + \sigma \epsilon_j + \mathbb{E}_\nu \left[ \sum_t \delta^t \left( n_t + \frac{1}{2\beta^S} (s_{jt} - \alpha - \nu_{jt})^2 \right) \right] + \lambda_0 (I_0 - n_0 - p_j^A) + \sum_t \lambda_t (I_t - n_t - p_j^S \cdot s_t)
\]

Taking the first order condition with respect to \( n_t \), we obtain \( \lambda_t = \delta^t \). Substituting this into the first order condition with respect to \( s_t \), we get

\[
\delta^t \cdot \frac{1}{\beta^S} (s_{jt} - \alpha - \nu_{jt}) = \lambda_t p_j^S
\]

\[
= \delta^t p_j^S
\]

and it follows that the optimal level of energy service consumption is given by

\[
s_{jt}^* = \alpha + \beta^S p_j^S + \nu_{jt}
\]

(A3.2)

I will discuss the welfare implications of household myopia at the end of this section after I solve the optimal level of energy services with the full nonlinear price schedule.

Conditional indirect utility and the optimal appliance choice Substituting \( s_{jt}^* \) into the budget constraints, we obtain \( n_{jt}^* = I_t - p_j^S \cdot s_{jt}^*, t > 0 \) and \( n_{j0}^* = I_0 - p_j^A \). This implies that the indirect utility is given by the expression
\[ V_j = I_0 - p_A^j + FE_{j}(\theta) + \sigma \epsilon_j + \mathbb{E}_\nu \left[ \sum_t \delta^t \left( I_t - p_S^j \left( \frac{1}{2} \beta S p_S^j + \alpha + \nu_{jt} \right) \right) \right] \]
\[ = I_0 - p_A^j + FE_{j}(\theta) + \sigma \epsilon_j + \sum_t \delta^t \left[ I_t - p_S^j \left( \frac{1}{2} \beta S p_S^j + \alpha \right) \right] \quad (A3.3) \]

Since \( \mathbb{E}_\nu(\nu) = 0 \). The utility maximizing appliance choice \( j \) will just be \( \text{argmax}(V_j) \).

### Conditional expenditure function

Since we have already solved the utility maximization problem, the most straightforward way to recover the expenditure function is to use the identity

\[ \bar{U} = V_j(p, e_j(p, \bar{U})) \quad (A3.4) \]

where \( \bar{U} \) is an arbitrary level of utility, \( V_j(p, I) \) is the conditional expected indirect utility and is a function of prices \( p = (p^A, p^S) \) and income \( I \), and \( e_j(p, U) \) is the conditional expected expenditure function, which is a function of prices \( p \) and utility \( U \). Setting \( A3.3 \) equal to \( \bar{U} \) and rearranging\(^8\), we obtain

\[ e_j(p, \bar{U}) = p_A^j - FE_{j}(\theta) - \sigma \epsilon_j + \sum_t \delta^t \left( p_S^j \left( \frac{1}{2} \beta S p_S^j + \alpha \right) \right) + \bar{U} \quad (A3.5) \]

and the unconditional expected expenditure function is given by the lower envelope \( \min_j e_j(p, U) \). Notice that the expenditure function is simply the negative of the conditional indirect utility function, where \( \bar{U} = \sum_{t=0}^T \delta^t I_t \).

### A3.2 Optimization w.r.t the full nonlinear price schedule

If the household perceives the actual price schedule, then I need to introduce another choice variable into the problem which describes the tier of the price schedule that the household chooses. Let the variable \( k_t \) denote the step of the marginal price schedule in period \( t \) and \( g(k_t) \) be the lowest level of consumption on the \( k \)th tier. Note that \( g(1) = 0 \) and \( g(\text{max}(k) + 1) = \infty \), which simply says that the lowest tier is bounded below by zero and the highest tier has no upper bound.

\(^8\)Where I have substituted \( e_j(p, \bar{U}) \) for \( \sum_{t=0}^T \delta^t I_t \) as implied by \( A3.4 \) So expenditure is given in present discounted value terms.
The household optimization problem is
\[
\max_{(j,k,s)} U(j, s_t)
\]
subject to
\[
I_0 \geq n_0 + p_j^A, \\
I_t \geq n_t + P_{jt}^S(s_t), \ t \in \{1, \ldots, 120\} \\
s_t \geq \underline{s}(k_t), \ t \in \{1, \ldots, 120\} \\
\underline{s}(k_t + 1) \geq s_t, \ t \in \{1, \ldots, 120\}.
\]
where I’ve defined the upper bound on consumption so that the domain of \(s_t\) is compact. Writing out the Lagrangian for this optimization problem (conditional on \(j\) and \(k\)), we obtain
\[
\mathcal{L} = n_0 + FE_j(\theta) + \sigma e_j + \mathbb{E}_\nu \left[ \sum_t \delta^t \left( n_t + \frac{1}{2\beta S}(s_{jt}^* - \alpha - \nu_{jt})^2 \right) \right] \\
+ \lambda_0(I_0 - n_0 - p_j^A) + \sum_t \lambda_t(I_t - n_t - P_{jt}^S(\underline{s}(k_t))) + \sum_t \mu_t(s_t - \underline{s}(k_t)) + \sum_t \bar{\mu}_t(\underline{s}(k_t + 1) - s_t)
\]
where \(\mu_t\) is the dual variable corresponding to the consumption lower bound constraint and \(\bar{\mu}_t\) is the dual variable corresponding to the consumption upper bound constraint. The solution is similar to the previous marginal / average price optimization, and we can show again that \(\lambda_t = \delta^t\) and the income constraint binds in every period. The Karush-Kuhn-Tucker conditions state that the following is necessary for utility maximization:
\[
\delta^t \left( \frac{1}{2\beta S}(s_{jt}^* - \alpha - \nu_{jt}) - p_j^S(\underline{s}(k_t)) \right) + \bar{\mu}_t = \mu_t \geq 0 \perp s_t^* \geq \underline{s}(k_t) \\
\mu_t - \delta^t \left( \frac{1}{2\beta S}(s_{jt}^* - \alpha - \nu_{jt}) - p_j^S(\underline{s}(k_t)) \right) = \bar{\mu}_t \geq 0 \perp \underline{s}(k_t + 1) \geq s_t^*
\]
where \(\perp\) is the standard notation used in complementarity problems that says that at least one of the expressions on either side must hold with equality.

**Case 1: \(s_t^* = \underline{s}(k_t)\):** If the agent locates at the lowest level on consumption on step \(k_t\) then it follows that \(\bar{\mu}_t = 0\) and
\[
s_{jt}^* = \beta^t p_j^S(\underline{s}(k_t)) + \alpha + \nu_{jt} + \mu_t \tag{A3.6}
\]
If consumption conditional on choosing step \(k_t - 1\) is the upper bound of the previous step, then locating at the kink is utility maximizing. If however, the customer would locate at the interior or the lower bound of step \(k_t - 1\) conditional on locating on step \(k_t - 1\), then step \(k_t\) will not maximize utility.
**Case 2:** \( s(k_t) < s^*_t < \bar{s}(k_t + 1) \): If the agent chooses a point on the interior of step \( k \), then it follows from the KKT conditions that \( \mu_t = \bar{\mu}_t = 0 \) and

\[
s^*_j = \beta^s p^s(s(k_t)) + \alpha + \nu_{jt}
\]

(A3.7)

In this case, step \( k_t \) will be the utility-maximizing step in month \( t \).

**Case 3:** \( s^*_t = \bar{s}(k_t + 1) \): If the upper bound constraint on consumption binds, this implies that

\[
s^*_j = \beta^s p^s(s(k_t)) + \alpha + \nu_{jt} - \bar{\mu}_t
\]

(A3.8)

In this case, the agent’s utility on step \( k_t + 1 \) will be greater than or equal to utility on step \( k_t \) (using an argument analogous to the one from Case 1 about the kink point). Thus we can ignore this case and without loss of generality just use \( s^*_j \) from the next highest step of the price schedule.

**Graphically determining which step of the price schedule** To see graphically how the agent chooses a step of the price schedule, we can plot both sides of the above inequalities as a function of \( s \) (where \( p^s(s(k_t)) = p^s(s_t) \) if \( s \in [s(k_t), \bar{s}(k_t + 1)] \)). Note that since the agent knows \( \beta^S \), \( \alpha \), and \( \nu_{jt} \) it follows that there is a single level of energy service consumption \( s_{jt} \) that satisfies the optimality conditions. See stylized illustration below where the marginal price schedule has three steps, and remember that \( \beta^S < 0 \).

**Figure 1:** Unique solution for \( s^*_j \) with nonlinear price schedule

\[
LHS = s_{jt}
\]

\[
RHS = \beta^s p^s(s^*_j) + \alpha + \nu_{jt}
\]
The regions of the plot where the intersection occurs at a vertical portion of $\beta^s p^s(g(k_t)) + \alpha + \nu_{jt}$ correspond to the case where the optimal energy services for this agent occur at a jump in the marginal price schedule (i.e. a kink in the budget constraint) where $\mu_{jt} \geq 0$.

**Indirect utility under the nonlinear price schedule**  
Substituting the optimal level of energy service and numeraire consumption from above into the direct utility function, we obtain

$$V_j = I_0 - p_j^A + FE_j(\theta) + \sigma \epsilon_j + \mathbb{E}_\nu \left[ \sum_t \delta^t (I_t - P^S_{jt}(s^*_jt) + \frac{1}{2} \beta^S (p^S(s^*_jt))^2) \right]$$  \hspace{1cm} (A3.9)

if $s^*_jt$ is in the interior of a marginal price step, and

$$V_j = I_0 - p_j^A + FE_j(\theta) + \sigma \epsilon_j + \mathbb{E}_\nu \left[ \sum_t \delta^t (I_t - P^S_{jt}(s^*_jt) + \frac{1}{2} \beta^S (p^S(s^*_jt) + \mu_{jt})^2) \right]$$  \hspace{1cm} (A3.10)

if $s^*_jt$ is at the kink at bottom of the $k$th step of the price schedule.

**Roy’s Identity**  
I have already derived the optimal amount of energy service consumption by maximizing the direct utility, but just for completeness we can arrive at the same answer by applying Roy’s Identity to the indirect utility function. This approach requires that we take the derivative of $V_j$ with respect to the intercept of the nonlinear price schedule. Notice that Leibniz’s Rule implies that

$$\frac{\partial P^S_{jt}(s^*_jt)}{\partial p^S_{jt}(0)} = \int_0^{\beta^S (p^S_{jt}(s^*_jt) + \mu_{jt}) + \alpha + \nu_{jt}} p^S_{jt}(z) dz$$

$$= (p^S_{jt}(s^*_jt) + \mu_{jt}) \beta^S + \int_0^{\beta^S (p^S_{jt}(s^*_jt) + \mu_{jt}) + \alpha + \nu_{jt}} 1 dz$$

$$= 2 \beta^S (p^S_{jt}(s^*_jt) + \mu_{jt}) + \alpha + \nu_{jt}$$

where $p^S_{jt}(0)$ is the intercept of the price schedule (i.e. the marginal price at $s_{jt} = 0$) and $\mu_{jt} = 0$ if the utility maximizing choice of energy service consumption occurs at the interior of a segment of the marginal price schedule. Thus

$$s^*_jt = -\frac{\partial V_j/\partial p^s(0)}{\partial V_j/\partial I_t}$$

$$= \beta^s p^s(g(k_t)) + \alpha + \nu_{jt} + \mu_{jt}$$

as we have shown above.
Welfare Implications of Household Myopia  When households don’t optimize with respect to the true price schedule but their experience utility corresponds to the true price schedule, then they consume too much $s_t$. Taking the difference between conditional indirect utility from the full nonlinear optimization in Equations A3.9 and A3.10 and the linear price schedule optimization in Equation A3.3, we get

$$
\sum_t \delta^t \left( p^S \alpha - \mathbb{E}_\nu \left[ P^S_j(s^*_j) \right] + \frac{1}{2} \alpha^S \left( (p^S)^2 - \mathbb{E}_\nu \left[ P^S_j(s^*_j)^2 \right] \right) \right) 
$$

(A3.11)

if the agent makes the same discrete durable purchase, where $p^S$ is the perceived average or marginal price and $P^S_j(s^*_j)$ and $p^S_j(s^*_j)$ are actual total expenditure and marginal price respectively at $s^*$. Notice that the convexity of the price schedule means that $\mathbb{E}_\nu(\nu)$ does not drop out of this expression. Furthermore, it’s possible that the agent makes the “incorrect” durable purchase because of the price schedule mis-perception.

A3.3 Alternative Derivation for Energy Services

Energy service consumption (e.g. loads of laundry or refrigeration) enters the household utility, but in my primary billing data I observe energy consumption measured in kilowatt hours (kWh). Converting the monthly electricity consumption data from my primary dataset into energy service consumption requires two steps: First, I need to specify an energy service production function for each discrete choice that converts kWh into energy services. Second, I need to determine how much energy was used by the modeled appliance. Let the production function for energy services be given by

$$
s_{ijt} = \gamma_j \cdot kWh^*_j + \sum_u \gamma_u \cdot kWh^*_u
$$

where $j$ is the modeled appliance, $u$ is an element of the set of all unmodeled appliances, $\gamma_k$ is a linear production function that takes energy and turns it into energy services, and $kWh^*_k$ is the unobserved number of $kWh$ of electricity used by appliance $k$ in month $t$. I observe $kWh_{ijt} = kWh^*_j + \sum_u kWh^*_u$ in the data. To determine the production function $\gamma_j$ for each modeled appliance, I use estimates from the U.S. Department of Energy. The DOE estimates that Energy Star washers use 25% less energy per load than conventional washers, so $\gamma_A = \gamma_B = \frac{1}{1-0.25} \gamma_C = 1.33 \gamma_C$. The outside option D includes households that purchased an Energy Star or a conventional

---

9The linearity assumption is very intuitive for many energy-consuming durables, and is supported by physical models of energy use. For example, washing machines produce clean clothes (an energy service), the number of clean clothing items is a function of the number of loads of laundry, and the energy used for each load of laundry is constant. Refrigerators use roughly the same amount of energy for each hour of refrigeration (given a temperature setting), and light bulbs use a constant amount of electricity for each hour of lighting (given desired brightness). Models of space heating and cooling also indicate that energy consumption is a linear function of the temperature difference between the interior and the exterior of a space.
washer more than a year ago, so I use the weighted average of the market share of these two purchase options from the RECS data for appliances older than two years to compute $\gamma_D$.

Since energy services are a composite of clean clothes, refrigeration, lighting, heating and cooling services, etc., it’s helpful to normalize $s$ so we can interpret its units. For all appliances that don’t appear in the model, normalize $\gamma_u = 1$, and define $\sum_u \gamma_u \cdot kW h^*_u \equiv kW h^*_u$. As a final normalization, I let one kWh of electricity consumed by choice $j = C$ produce one unit of energy services, so $\gamma_C \cdot kW h^*_C \equiv kW h^*_C$. Using these normalizations, it follows that

$$s_{iCt} = kW h^*_C + kW h^*_U \equiv kW h_{iCt}$$

This simply means that total energy services in households that purchased a new non-Energy star appliance is equal to their total electricity consumption, $kW h_{iCt}$.

To compute energy service consumption from overall energy consumption $kW h_{ijt}$ and modeled appliance choice, define the variable

$$\sigma_{ijt} \equiv \frac{kW h^*_ijt}{kW h^*_ijt + kW h^*_U}$$

to be the average share of total energy consumption used by the modeled appliance, and notice that the denominator is equal to observed consumption, $kW h_{ijt}$.

It follows that

$$s_{ijt} = \gamma_j \cdot kW h^*_ijt + kW h^*_U$$

$$= \left[1 + \sigma_{ijt}(\gamma_j - 1)\right] kW h_{ijt}$$

where again $kW h_{ijt}$ is just total energy consumption. The expression above now just requires an estimate of $\sigma_{ijt}$ to compute $s_{ijt}$. I estimate $\sigma_{ijt}$ using monthly plug-level data from Pecan Street and the same procedure I presented for the purchase probability imputation. In addition to using income, home size, and home age to discretize the Pecan Street data, I also use mean monthly temperature. This allows the share of energy consumed by washing machines and refrigerators to decrease during hot months where the air conditioning is running. It follows that the price of

---

10 Although there have been a number of economists who have suggested that engineering estimates such as the ones I use here overstate actual savings (Fowlie et al. (2015), Davis et al. (2014)), these concerns shouldn’t apply to my particular setting because I explicitly model the behavioral responses that would create a divergence from the engineering calculations and the experimentally measured savings in previous economic studies. It is easy to plug various appliances into a watt meter and measure how much energy is consumed for a load of laundry, an hour of refrigeration at a given temperature, etc. This is the calculation upon which I base my estimates of $\gamma_j$.

11 Also note that I assume that modeled appliance choice doesn’t affect consumption by other appliances $kW h^*_U$.
energy services is
\[ p^{S}_{ijt} = \frac{p^{kWh}_{it}}{1 + \sigma_{ijt}(\gamma_j - 1)} \]

The price of electricity, \( p^{kWh}_{it} \), varies over time and across 10 different regions in the utility territory. I collected these prices for my sample period from the utility.\(^\text{12}\)

A4 Discrete Choice Derivations from Indirect Utility

A household’s discrete and continuous choices are deterministic since the household knows its idiosyncratic preferences \( \epsilon_{ij} \).\(^\text{13}\) However, since these idiosyncratic tastes are unobserved to the econometrician, I need assume a distribution for these terms. This will allows me to integrate over the distribution and determine a probability that a particular household makes a given decision. Recall from the text that the CDF of \( \epsilon_iB, \epsilon_iC, \) and \( \epsilon_iD \) is given by

\[ F(\epsilon_iA, \epsilon_iB, \epsilon_iC) = \exp\left( -\left( \exp(-\epsilon_iA/\rho) + \exp(-\epsilon_iB/\rho) \right)^\rho - \exp(-\epsilon_iC) \right) \]

and let the random parameter \( \theta_g \) have the distribution \( f_{\theta_g}(\cdot) \), where \( \theta_g \in (-\infty, 0] \). I will let \( f_{\theta_g} \) be an exponential distribution on the negative real numbers in my primary results, and I also test robustness to a uniform distribution and a Rayleigh distribution. It follows that the probability that a household makes each of the discrete choices is given by

\[ Pr_{ij} = Pr(V_{ij}(\bar{\epsilon}, \theta) > V_{ij'}(\bar{\epsilon}, \theta), j \neq j') = \int_R f(\theta, \epsilon_B, \epsilon_C, \epsilon_D)d\epsilon_Dd\epsilon_Cd\epsilon_Bd\theta \quad (A4.1) \]

where \( R \) is the region in which the above inequalities are satisfied. I can analytically evaluate this multidimensional integral over \( \epsilon \) and simplify this expression to depend on an integral over \( \theta \). The integration over \( \theta \) does not have a closed form and I will compute this numerically in the model estimation.

By way of notation, I will drop the subscripts \( i \) and \( g \) everywhere and let \( \mu_j = V_j - \epsilon_j \) (i.e. \( \mu_j \) is the portion of the indirect utility not determined by the idiosyncratic preferences in \( \epsilon \)). Thus \( \mu_j \)

\(^{12}\)The 10 different baseline regions are defined by a customer’s zipcode. This utility offers several different pricing plans to its electricity customers, including time of use pricing and critical peak pricing. Despite offering several plans, only 6.8% of customers were enrolled in one of these plans in 2015, up from less than 5% in 2013. This information is available from the EIA’s form 861, which collects information on US utilities. The form is available from https://www.eia.gov/electricity/data/eia861/ There is also a special rate for low-income households. Since the households in my dataset are mostly owner-occupied units, the fraction of low-income customers is less than the full utility sample.

\(^{13}\)The variable \( \nu \) could represent how many children the family has, which is unobserved in the data but clearly a know factor to the family that would affect their demand for energy services.
is a function of $\nu_j$ and $\mu_A$ is a function of both $\nu_A$ and $\theta$. I will also suppress this dependence for notational simplicity below.

**A4.1** $Pr_{iD}$

The probability than individual $i$ chooses discrete choice $D$ is given by

$$Pr_{iD} = Pr(\mu_D + \epsilon_D > \max(\mu_C + \epsilon_C, \mu_B + \epsilon_B, \mu_B + \epsilon_B + \text{Subsidy} + \theta))$$

$$= \int \int \int_{R} f(\theta, \epsilon_B, \epsilon_C, \epsilon_D) d\theta d\epsilon_B d\epsilon_C d\epsilon_D$$ (A4.2)

where $f(\cdot)$ is the joint density of $(\theta, \epsilon_B, \epsilon_C, \epsilon_D)$ and the region $R = \{(\theta, \epsilon_B, \epsilon_C, \epsilon_D)\}$ such that all of the following are satisfied:

1. $\epsilon_B < \mu_D + \epsilon_D - \mu_B - \text{Subsidy} - \theta$,
2. $\epsilon_B < \mu_D + \epsilon_D - \mu_B$, and
3. $\epsilon_C < \mu_D + \epsilon_D - \mu_C$

The first condition (1) corresponds to the region where $D$ is preferred to $A$, (2) corresponds to the region where $D$ is preferred to $B$, and (3) corresponds to the region where $D$ is preferred to $C$.

For notational convenience, define $\bar{a}(\theta, \epsilon_D)$, $\bar{b}(\epsilon_D)$, and $\bar{c}(\epsilon_D)$ to be the right hand side of the above inequalities, i.e. $\bar{a}(\theta, \epsilon_D) = \mu_D + \epsilon_D - \mu_B - \text{Subsidy} - \theta$. Also define

$$A(\theta, \epsilon_C, \epsilon_D) = \int_{\epsilon_B = -\infty}^{\bar{a}(\theta, \epsilon_D)} f(\theta, \epsilon_B, \epsilon_C, \epsilon_D) d\epsilon_B$$

$$= f_\theta(\theta) f_{D}(\epsilon_D) \int_{\epsilon_B = -\infty}^{\bar{a}(\theta, \epsilon_D)} f_{\epsilon_B, \epsilon_C}(\epsilon_B, \epsilon_C) d\epsilon_B$$

$$= f_\theta(\theta) f_{D}(\epsilon_D) \frac{\partial F_{\epsilon_B, \epsilon_C}(\epsilon_B, \epsilon_C)}{\partial \epsilon_C} \bigg|_{\epsilon_B = \bar{a}(\theta, \epsilon_D)}$$

$$= f_\theta(\theta) f_{D}(\epsilon_D) \cdot \exp\left(-\left(\bar{a}(\theta, \epsilon_D)/\rho + e^{-\epsilon_C/\rho}\right)^\rho\right)\left(e^{-\bar{a}(\theta, \epsilon_D)/\rho} + e^{-\epsilon_C/\rho}\right)^{\rho-1} e^{-\epsilon_C/\rho}$$

where the second equality follows because $\theta \perp \epsilon_D \perp (\epsilon_B, \epsilon_C)$, the second equality follows from the fundamental theorem of calculus, and the final equality is a result of the functional form of $F(\epsilon_B, \epsilon_C, \epsilon_D)$. Define $B(\theta, \epsilon_C, \epsilon_D)$ analogously where $\bar{a}(\theta, \epsilon_D)$ is simply replaced with $\bar{b}(\epsilon_D)$. Now notice that the integral in Equation [A4.2] can be decomposed into two parts, so that
\[ Pr_{iD} = \iiint_{R_A} A(\theta, \epsilon_C, \epsilon_D)d\theta d\epsilon_C d\epsilon_D + \iiint_{R_B} B(\theta, \epsilon_C, \epsilon_D)d\theta d\epsilon_C d\epsilon_D \]

where \( R_A = \{ (\theta, \epsilon_C, \epsilon_D) : \theta + \text{Subsidy} > 0, \epsilon_C < \mu_D - \mu_C + \epsilon_D \} \) and \( R_B = \{ (\theta, \epsilon_C, \epsilon_D) : \theta + \text{Subsidy} \leq 0, \epsilon_C < \mu_D - \mu_C + \epsilon_D \} \). I will evaluate the integral on the left, but the procedure is very similar for the second integral.

\[
\int_{\theta \in R_A} f_\theta(\theta) \int_{\epsilon_D} f_{\epsilon_D}(\epsilon_D) e^{z_1(\theta, \epsilon_D)} d\epsilon_D d\theta
\]

where in the first line I have used the change of variable

\[
z_1 = -(e^{-(\mu_D + \epsilon_D - \mu_B - \text{Subsidy} - \theta)/\rho} + e^{-\epsilon_C/\rho})^\rho
\]

and in the third line I have used the change of variable

\[
z_2 = -e^{\epsilon_D} K(\theta)
\]

Using the same steps to integrate \( B(\theta, \epsilon_C, \epsilon_D) \) over the region \( R_B \), we obtain

\[
Pr_{iD} = \int_{\theta = -\infty}^{\text{Subsidy}} \frac{f_\theta(\theta)}{1 + K(\theta)} d\theta + \frac{1}{1 + K} \int_{-\infty}^{0} f_\theta(\theta) d\theta
\]

where \( K = (e^{-(\mu_D - \mu_B)/\rho} + e^{-(\mu_D - \mu_C)/\rho})^\rho \).

### A4.2 Other Probabilities

The derivation for the other probabilities is similar and results in the following expressions:
\[ Pr_{iA} = \int_{-\text{Subsidy}}^{0} f_\theta(\theta) \cdot \frac{(1 + \exp(-(\mu_A - \mu_C)/\rho))^{\rho-1}}{\exp(-\mu_A) + (1 + \exp(-(\mu_A - \mu_C)/\rho))^{\rho}} d\theta \quad (A4.3) \]

\[ Pr_{iB} = \int_{-\infty}^{-\text{Subsidy}} f_\theta(\theta) \cdot \frac{(1 + \exp(-(\mu_B - \mu_C)/\rho))^{\rho-1}}{\exp(-\mu_B) + (1 + \exp(-(\mu_B - \mu_C)/\rho))^{\rho}} d\theta \]

\[ Pr_{iC} = \int_{-\text{Subsidy}}^{0} f_\theta(\theta) \cdot \frac{(1 + \exp(-(\mu_C - \mu_A)/\rho))^{\rho-1}}{\exp(-\mu_C) + (1 + \exp(-(\mu_C - \mu_A)/\rho))^{\rho}} d\theta + \int_{-\infty}^{-\text{Subsidy}} f_\theta(\theta) \cdot \frac{(1 + \exp(-(\mu_C - \mu_B)/\rho))^{\rho-1}}{\exp(-\mu_C) + (1 + \exp(-(\mu_C - \mu_B)/\rho))^{\rho}} d\theta \]

\[ Pr_{iD} = \int_{-\text{Subsidy}}^{0} f_\theta(\theta) \cdot \frac{1}{1 + (\exp(\mu_A/\rho) + \exp(\mu_C/\rho))^{\rho}} d\theta + \int_{-\infty}^{-\text{Subsidy}} f_\theta(\theta) \cdot \frac{1}{1 + (\exp(\mu_B/\rho) + \exp(\mu_C/\rho))^{\rho}} d\theta \]

### A5 Moment Conditions

To derive the moment conditions I use in my estimation, I make the following assumption:

\[ Z \perp \epsilon \sim \text{Generalized Extreme Value} \quad (A5.1) \]

\[ \text{Subsidy} \perp \theta \sim \text{Rayleigh, Lognormal, or Exponential} \quad (A5.2) \]

\[ \epsilon \perp \theta \quad (A5.3) \]

\[ \mathbb{E}(\nu|\epsilon) = 0 \quad (A5.4) \]

\[ \mathbb{E}(\Delta \nu \Delta p^{s,iV}|\epsilon) = 0 \quad (A5.5) \]

Where Z are all the observables that enter the utility function. This implies that

\[ \mathbb{E}\left\{ \left( I_{ij}(Z, \epsilon) - Pr_{ij}(Z) \right) f(Z) \right\} = \mathbb{E}_Z \left\{ \mathbb{E}_{\epsilon|Z} \left[ \left( I_{ij}(Z, \epsilon) - Pr_{ij}(Z) \right) f(Z) \right] \right\} = \mathbb{E}_Z \left\{ \mathbb{E}_{\epsilon|Z}[I_{ij}(Z, \epsilon)] - Pr_{ij}(Z) \right\} f(Z) \]

\[ = \mathbb{E}_Z \left\{ \left( Pr_{ij}(Z) - Pr_{ij}(Z) \right) f(Z) \right\} = 0 \]

where \( f(Z) \) is an arbitrary function of Z. The first equality follows from the law of iterated expectations, the second equality follows because Z can be treated as a constant inside of the
square brackets, and the third line follows from the independence of \( Z \) and \( \epsilon \) and the generalized extreme value distribution of \( \epsilon \).

### A5.1 10 Moment Conditions in Nonlinear Estimation

I use four functions of \( Z \) as instruments: (1-3) \( I_j = 1, j \in \{A, B, C\} \) and (4) \( \mu_j(Z_i) = -p^A_{ij} + \xi_{ij}(\theta_i) + V^*_ij \), the portion of the indirect utility that doesn’t contain \( \epsilon \).

\[
E[(\mathbb{I}_{iA} - P_{rA})] = 0 \tag{A5.6}
\]

\[
E\left((\mathbb{I}_{ij} - P_{rij})\right) = 0, \quad j \in \{B, C\} \tag{A5.7}
\]

\[
E\left((\mathbb{I}_{iA} - P_{rA})\mu_{ij}\right) = 0 \tag{A5.8}
\]

\[
E\left((\mathbb{I}_{ij} - P_{rij})\mu_{ij}\right) = 0, \quad j \in \{B, C\} \tag{A5.9}
\]

Finally, I impose the following moment condition implied by the continuous utilization decision:

\[
E[(\Delta s_{ijt} - \beta^S \Delta p^S_{ijt} - \beta^w \Delta w_{it}) \cdot (\mathbb{I}_{it}^{\text{tier}} \cdot \Delta p^S_{ijt})] = 0 \quad \forall j \tag{A5.10}
\]

where \( \Delta p^{S,IV} = (\mathbb{I}_{it}^{\text{tier}} \cdot \Delta p^S_{ijt}) \) which is equal to 0 if the household changed tiers and \( \Delta p^S_{ijt} \) if the household stayed on the same tier. Thus I’m just picking up the within household-by-tier variation in electricity prices.

### A5.2 “Definitional” Moment Conditions

I estimate a separate intercept parameter \( \alpha_i \) for each household, but it would be impractical to impose these moment conditions in the nonlinear estimation. Consequently I use the condition that \( E(\nu_{ijt}) = 0 \quad \forall i, j \) to define the parameter \( \alpha_{ij} = s_{ijt} - \beta_g p^S_{ijt} - 1/2 \Delta \nu_{ijt} \) where the \( 1/2 \Delta \nu_{ijt} \) term is used because I used two time periods to compute \( \Delta \nu_{ijt} \). I then compute \( P_{rij} \) four times (one for each \( \alpha_i \)) and weight the four values by \( I_{ij} \) observed in the data.

To further reduce the number of parameters computed in nonlinear estimation, I embed the following moment conditions (which hold by definition) in the nonlinear estimation:

\[
E[(\Delta s_{ijt} - \beta^S \Delta p^S_{ijt} - \beta^w \Delta w_{it}) \cdot (\mathbb{I}_{it}^{\text{tier}} \cdot \Delta w_{it})] = 0 \quad \forall j \tag{A5.11}
\]

This is easily implemented since for any value of \( \beta^S \) I can run least squares to compute \( \beta^w \) (i.e. \( \beta^w \) is defined by the projection of \( \Delta s_{ijt} - \beta^S \Delta p^S_{ijt} \) onto the subspace of \( \mathbb{R}^N \) spanned by the vector \( \Delta w_{it} \)).
A5.3 5 Parameters per Income Quintile

I estimate 5 parameters in my GMM objective function, and I run the estimation separately for each income group:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta^S_g$</td>
<td>Determines price elasticity of Energy Service Consumption</td>
</tr>
<tr>
<td>$\tau_g$</td>
<td>Utility of owning a new appliance</td>
</tr>
<tr>
<td>$\kappa_g$</td>
<td>Non-energy utility of owning an Energy Star appliance</td>
</tr>
<tr>
<td>$\lambda_g$</td>
<td>Parameter governing disutility of rebate form hassle costs</td>
</tr>
<tr>
<td>$\sigma_{\epsilon,g}$</td>
<td>Variance of the logit error</td>
</tr>
</tbody>
</table>

A6 Simulation Details

Since I don’t observe realizations of the idiosyncratic taste parameters $\epsilon_{ij}$ or rebate cost parameter $\theta_i$, I simulate from the distributions of these random variables to compute purchase probabilities in my estimation procedure and to compute expected consumer surplus in my policy evaluations.

Recall that $\epsilon$ is i.i.d across households, with joint distribution given by

$$F(\epsilon_i^A, \epsilon_i^B, \epsilon_i^C) = \exp\left(-\left(\exp(-\epsilon_i^A/\rho) + \exp(-\epsilon_i^B/\rho)\right)^\rho - \exp(-\epsilon_i^C)\right)$$  \hspace{1cm} (A6.1)

Notice that $\epsilon_C$ is independent of ($\epsilon_A, \epsilon_B$), so I begin by drawing from the joint distribution of ($\epsilon_A, \epsilon_B$) using a “Normal to Anything” (NORTA) transformation. First, draw from a multivariate normal distribution with mean 0, variance 1, and covariance $1 - \rho^2$. Denote this random vector ($n_A, n_B$).

I then transform ($n_A, n_B$) into a correlated uniform random vector ($u_A, u_B$) by transforming each component of $n_j$ to a uniform random variable using a standard normal CDF. Notice that the correlation structure will be preserved in ($u_A, u_B$). At this stage, I also draw from a $U(0,1)$ distribution to obtain $u_C$, which is independent of ($u_A, u_B$). Finally, using the inverse of the (marginal) CDFs of $\epsilon_A, \epsilon_B,$ and $\epsilon_C$ I can transform the vector ($u_A, u_B, u_C$) into ($\epsilon_A, \epsilon_B, \epsilon_C$) that has joint distribution [A6.1]. This last step relies on the inverse CDF method, which can transform a $U(0,1)$ variable into a random variable with an arbitrary distribution. Notice that $F_{\epsilon_j}^{-1}(u_j) = -ln(-ln(u_j))$, so it follows that the random variable $\epsilon_j = F_{\epsilon_j}^{-1}(u_j)$ will have distribution $F_{\epsilon_j}^{-1}(\cdot)$ if $u_j$ is distributed $U(0,1)$. The correct correlation structure is generated because of the initial correlation between ($n_A, n_B$) in the first step.

The cost variable $\theta$ is independent of $\epsilon$. To draw from the distribution of $\theta$, I begin by using Matlab’s built-in Rayleigh random number generator to generate realizations of a Rayleigh(1) random variable. The variable $\theta = \lambda_g \cdot r_{\theta}$, where $r_{\theta}$ denotes a draw from the Rayleigh(1) distribution. I use the same 100 draws from the joint distribution of ($\epsilon_A, \epsilon_B, \epsilon_C, \theta$) for all households.
A7 Producer Surplus

In the policy evaluation and design in the paper, I relax the assumption that the appliance market is competitive and consider the programs’ effect on total welfare if producers earn rents on Energy Star appliances. In this section, I develop a simple model of the supply side of the appliance market that would rationalize the markups I consider in the paper.

Consider a monopolist Energy Star appliance seller in market $m$. Consistent with notation in the paper, households in market $m$ are indexed by $i$ and have conditional indirect utility for discrete choice $j$ given by $V_{ij}$. For expositional ease in this section, consider the case where the consumer choice set consists of options $j \in \{A, B, C\}$ and where the error terms in the conditional indirect utility are independent. This allows me to use simple, closed form logit formulas to illustrate how producers set markups, but these assumptions can be relaxed without affecting the conclusions.

Notice that the consumer chooses option $j$ with probability given by

$$Pr_{ij} = \frac{\exp(V_{ij}/\sigma_{gi})}{\sum_j \exp(V_{ij}/\sigma_{gi})}$$

The Energy Star appliance producer maximizes expected profits, which are equal to price less marginal cost for each unit times the expected number of sales over all consumers $i$ in the market:

$$\max_{p_A} \sum_i Pr_{iA} (p_A - MC)$$

where the Energy Star appliance is produced, shipped, and sold in this market at constant marginal cost $MC$. The first order condition results in a variant of the standard Lerner equation:

$$p_A - MC = \frac{\sum_i Pr_{iA}}{\sum_i 1/\sigma_{gi} Pr_{iA}(1 - Pr_{iA})}$$

If $Pr_{iA}$ and $\sigma_{gi}$ are the same for everybody, and as in the data if 10% of households purchase an Energy Star appliance, then the profit per appliance is given by

$$p_A^* - MC = 1.1 \sigma$$

Since my estimates of $\sigma$ are close to $100$, a markup over marginal cost of $125$ seems very reasonable.

---

14 This assumption is just for expositional ease. I have also worked out the case of multiple appliance suppliers competing in a horizontally differentiated product space.
A8  Optimal Policy Design

The policy space depends on the constraints placed on the welfare-maximizing planner. I consider several progressively more flexible constraints that are illustrative of the tradeoffs made in the policy design problem. First, I set a fixed program budget and allow the policy-maker to choose the best way to distribute subsidy money across households in different parts of the income distribution. This choice affects both equity and efficiency, as I will show in the next section that restricting program eligibility and increasing the subsidy amount reduces the share of inframarginal households. Second, I allow the planner to set the subsidy rates, and energy tariffs (both fixed and marginal prices) subject to a balanced budget constraint.\footnote{I am working on two additional counterfactuals, one in which I also allow the planner to tax new inefficient appliances \((j = C)\) instead of subsidizing efficient appliances, and a second in which I consider the benefits of a technology that decreases the distribution of \(\theta\) in a first order stochastic dominance sense.} \footnote{If the planner separately sets subsidy amounts and fixed charges for each income quintile (and each income quintile is self-funding), then I actually only need to solve a univariate optimization problem in \(S_{gi}\). Optimal program size is zero if the externality reductions are less than the private welfare cost \(\theta\) for all (not just marginal) participants.}

Planner Chooses \(S_{gi}\) Given Fixed Program Budget  Let the subsidy program budget be fixed at \(B\), and allow the policy-maker to set a different subsidy amount \(S_{gi}\) for households in quintile \(g_i\) of the zipcode income distribution (i.e. the subsidy varies by zipcode). The planner’s optimization problem is given by

\[
\max_{S_{gi}} \sum_i w_{gi} V_i - \phi \sum_i \sum_j Pr_{ij} kWh_{ij} \\
\text{subject to} \sum_i Pr_{iA} S_{gi} = B, \\
S_{gi} \geq 0 \ \forall g_i
\]

The Karush-Kuhn-Tucker conditions for this problem with respect to \(S_{gi}\) are given by

\[
\sum_i w_{gi} \frac{\partial V_i}{\partial S_{gi}} - \phi \sum_i \sum_j \frac{\partial Pr_{ij}}{\partial S_{gi}} kWh_{ij} = \lambda^B \left( \sum_i Pr_{iA} \cdot I_{gi} + \sum_i \frac{\partial Pr_{iA}}{\partial S_{gi}} S_{gi} \right) \\
\sum_i Pr_{iAS_{gi}} = B
\]  

(\ref{A8.1}) (\ref{A8.2})

where \(\lambda^B\) is the Lagrange multiplier on the budget constraint and \(\lambda^{S_{gi}}\) is a dual variable cor-
responding to the constraint. We also have the complementarity conditions necessary for an optimal solution that $S_g \lambda^{S_g} = 0$ for all $g$. At an interior optimum, this is a system of six nonlinear equations and six unknowns (five subsidy amounts $S_{g_i}, g_i = 1, \ldots, 5$ and the Lagrange multiplier $\lambda$).

To understand the above expression, imagine increasing the subsidy in the $g^{th}$ income quintile $S_g$ by $\$1. This increases social welfare by $w_g$ times the marginal utility to households in the $g^{th}$ quintile of the zipcode income distribution of an extra dollar of subsidy. For households who are inframarginal to the $\$1$ change in the subsidy, the change in expected indirect utility is given by $S_{g_i} + \theta$, where $\theta < 0$. For marginal households, the change in expected indirect utility is given by $p_i^A - p_{ij} + \theta + \frac{V_i^A - V_{ij}^A}{\beta_{g_i} \sum_t \delta(p_{iAt}^A - p_{ijt}^A)}$ (A8.3)

where $j \neq B$. Notice that in both cases, the expected indirect utility decreases in the magnitude of $\theta$, creating an incentive for the planner to target customers that have a low participation cost. Furthermore, the difference in consumption utility for inframarginal households $V_i^A - V_{ij}^A$ increases in $\beta_{g_i}$, creating an incentive for the planner to target customers that have a high value of energy service consumption.

The second term in Equation A8.1, which represents the decrease in environmental externalities, contains two separate forces. First, the planner wants to maximize the change in the probability of purchasing an efficient appliance, i.e. maximize the probability of being marginal to the subsidy. At the low levels of marginal participation observed in the data, the probability that household $i$ is marginal to the subsidy increases in the subsidy amount. Thus with a fixed budget, there is an incentive for the planner to offer a large incentive to a small number of households to increase the proportion of marginal program participants. This intuition is depicted graphically in Figures ?? and ??, The second force that affects the reduction in environmental damages is savings conditional on purchase of an efficient appliance. These savings decrease with a household’s elasticity of energy service demand because of the rebound effect. For example, an elastic household will use more energy services after installing an efficient appliance because the price of utilization has decreased, while an inelastic household will consume the same amount of services and have strictly lower energy consumption. This effect counteracts the incentive of the planner to target elastic households that we saw in the first term. The right hand side of the expression gives the additional dollars of program budget needed to pay the higher subsidy to quintile $g$, and this money is no longer available to provide to other quintiles. Thus the optimal policy will equate the marginal utility gains and reduced externality damages in each quintile to the marginal increase in program budget needed to offer such a subsidy (times the factor $\lambda$).

To review, the optimal subsidy for each income quintile $g$ is (1) decreasing in the cost of program

---

\(^{17}\)In the appendix, I demonstrate that the unconstrained problem where $S_{g_i}$ is allowed to be less than zero has the same solution, so the dual variables $\lambda^{S_{g_i}}$ which enter A8.1 additively are all zero.
participation $\text{abs}(\theta)$, (2) increasing in private value of energy service consumption for marginal households (linear in $\beta_g$), (3) increasing in the number of marginal households (i.e. convex in $S$), (4) increasing in average energy savings conditional on begin marginal (also determined by $\beta_g$), and (5) depends on the total number of participants.

**Planner Chooses** $S_g$, $x$, and $F$  In this case, the planner is allowed to choose the subsidy amount for each zipcode income quintile, the energy tax $x \geq 0$, and a fixed payment added to the monthly bill $F \geq 0$, subject to the constrain that the program be revenue neutral, i.e. 
\[ \sum_i \left( F + x \cdot \sum_j P_{ij}kWh_{ij} \right) = \sum_i P_{i\Delta}S_g \] and that the electricity tax and fixed bill charge are non-negative.$^{18}$

The Karush-Kuhn-Tucker conditions with respect to $S_g$ are still given by $A8.1$, and there are several additional equations that pin down $x$, $F$, and the corresponding dual variables. Intuitively, if the marginal price of electricity is already above the social cost, then $x = 0$ and we can exclusively use a fixed charge to recover the program funds.$^{19}$ The optimal program size is then pinned down by the first order condition:

\[ \frac{1}{N} \sum_i w_{gi} = \lambda^B \]

where again $\lambda^B$ is the Lagrange multiplier on the budget constraint.

Now I have a system of seven equations and seven unknowns, which is a very computationally difficult problem. To illustrate the applicability of this framework, I evaluate the benefits of two alternative policies that raise subsidy money more efficiently or distribute it more progressively. First, I will show that increasing the generosity of the subsidy in tandem with restrictions on subsidy eligibility can be both equity and efficiency improving.$^{20}$ Second, I measure the effect of a program funded through a fixed charge to customers’ bills. While this means that substantially less energy is saved (the energy price incentive that saved 99% is shut down in this case), such a program could be justified by a failure in the appliance market that results in suboptimal investment in efficient durables.

$^{18}$A more general optimal policy problem would allow $x < 0$ and $F_g$ to depend on the zipcode income.

$^{19}$In the appendix, I consider the more general case where the retail price of electricity is below the social cost of electricity.

$^{20}$Where equity improvements are defined by more progressive income transfers.
References


