Determinants of the Willingness to Pay for Electricity in Developing Countries: Evidence from a Household Survey from India*

Zeyu Wang† Frank A. Wolak‡

June 16, 2022

Abstract

This paper estimates a model of the household-level demand for electricity services such as lighting, heating and cooling, home appliances, and business use in the Indian state of Rajasthan using a combination of household-level survey data and administrative data. This model incorporates customer-level demographic characteristics, billing cycle-level weather variables, and the fact that households are subject to electricity outages and face increasing block price schedules for their electricity consumption. We estimate two versions of the model that differ in how the relationship between electricity use and consumption of each electricity service is modeled. The first model uses a shape-constrained kernel regression and the second model uses a customer-level constant elasticity of electricity consumption with respect to energy service model. Both energy service demand models produce estimates of the response of each of the above four categories of energy services to changes in the price of each energy service. Both versions of the model also produce estimates of the marginal willingness to pay for an additional hour of each of the four categories of energy services. The mean marginal willingness to pay across customers for an additional hour an energy service is the smallest for lighting and the largest for home appliance services.

*We would like to thank Jin Chen, Kevin Behan and Evan Magnusson for outstanding research assistance.

†Department of Economics, Stanford University, Stanford, CA 94305, zeyuw@stanford.edu

‡Program on Energy and Sustainable Development and Department of Economics, Stanford University, Stanford, CA 94305-6072, wolak@zia.stanford.edu

1 Introduction

Getting customers to pay for the electricity they consume is major challenge to the long-term financial viability of electricity supply industries in many developing countries. It may be cheaper for the customer to have an informal connection to the grid, make financial arrangements with a meter reader to reduce their electricity bill, or simply not pay their bill. These informal sources of electricity have the advantage of being lower cost, but they are also likely to be less reliable and less able to provide a sizeable supply of electricity. Consequently, customers with these kinds of connections are less likely to make investments in capital goods that use a substantial amount of high quality electricity. This logic suggests that customers with a sizeable demand for high quality electricity because of their electricity-consuming capital goods holdings are more likely to obtain a formal connection and pay their electricity bill.

A customer's monthly electricity consumption depends on the demand for individual electricity services such as lighting, heating and cooling, household appliance use, and business use. A customer's willingness to pay for an additional hour of each of these services is likely to differ for a variety of reasons. For example, during the nighttime hours, a customer might have a significantly lower willingness to pay for an additional hour of heating or cooling services than his willingness to pay for an additional hour of lighting services. A customer that operates a business out of their dwelling might also be likely to be willing to pay substantially more for an additional hour of business electricity services that other electricity services.

An understanding of the willingness to pay of customers for additional hours of different kinds of energy services can help firms and regulators design pricing and other policy mechanisms to encourage customers to obtain formal connections that they pay for. For example, if a customer has the highest willingness to pay for an additional hour business services, then policies that support investments in electricity consuming goods for business use should encourage the customer to obtain a formal connection and pay her bill. The goal of this paper is to provide customer-level information on the willingness to pay for four electricity services in the Indian state of Rajathan: (1) lighting, (2) heating and cooling, (3) household appliance use, and (4) business use.

To do this, we specify a model of the household-level demand for electricity services and electricity demand using a combination of household-level survey data, conducted in two districts of Rajasthan, and administrative data composed of household-level billing cycle consumption, electricity bills, and tariffs. This model incorporates customer-level demographic characteristics, billing cycle-level weather variables, and accounts for the fact that households are subject to electricity outages and increasing block price schedules for their electricity consumption.

Our model of the household-level electricity demand embodies the fact that electricity consumption is the result of a household's demand for hours of service from all electricity-consuming capital goods that it owns. Households do not directly consume electricity, but instead they purchase capital goods that when combined with electricity provide the services they desire such as lighting, which combines electricity use with an electricity-consuming lighting device. The amount of electricity that an hour of lighting service requires depends on the number of lighting devices a household owns, the specific capital goods used to produce light, and when the capital good is used. Combining household-level survey data on electricity appliance holdings and the monthly hours of use of each electricity-consuming capital good with the household's billing cycle electricity consumption and total bill allows us to recover an estimate of the household's demand for these individual electricity services.

We employ a two-step estimation procedure that first recovers what we call a household-level "electricity consumption function" which characterizes the relationship between the household's billing cycle level electricity consumption in kilowatt-hour (kWh) and total hours of use of each of the four electricity services during that billing cycle. The estimated "electricity consumption function", together with the hours of electricity services consumed and information on the customer's monthly electricity bill are used to estimate the household's billing cycle-level demand for each electricity service. The demand for each electricity service is derived from an underlying model of utility-maximizing behavior subject to a nonlinear budget constraint that arises because the household faces an increasing block price schedule for their electricity consumption. Moreover, there is a nonlinear relationship between a household's demand for each electricity service and its total electricity consumption for the billing cycle.

Our electricity service demand model recovers customer-level own-price and income elasticities for different electricity services, as well as cross-price demand elasticities between the different electricity services. The model also provides customer-level marginal willingness pay values for an additional hour of each energy services. These measures can provide valuable input into the design of policies that increase the likelihood that customers pay their bills. For example, if customers have a high marginal willingness to pay for a certain electricity service, then providing subsidies to purchase the capital goods necessary for the household to consume these services is likely to increase their willingness to pay for a reliable and sizeable supply of electricity.

Our electricity consumption function estimates implies significant heterogeneity across customers in the predicted increase in electricity consumption as a result of a marginal increase in each energy service demand. There are two main reasons for this result. First, there are differences in the quantity of energy consuming capital goods that produce each energy service that each customer owns. For example, depending on what kind or how many televisions, washing machines, or microwave ovens a household owns, the electricity consumption per hour of use of each energy service can vary across customers.

Second, how intensively the various appliances within each energy service category are used can vary depending on weather conditions. For example, when it is extremely hot outside a household might stay inside and watch more television and wash clothes rather than use other appliances in that energy service category. A given appliance can consume more electricity depending on weather conditions. For example, one hour of use of an air conditioner when it is extremely hot outside consumes more electricity than it does during milder days.

Results from our energy services demand system finds significant differences in the own-price elasticity of demand across the four energy services and significant differences across customers in these own price elasticities for the same energy service. There is also considerable heterogeneity in the income elasticity of demand across these four services with the highest income elasticity for domestic end-uses and lowest for heating and cooling or lighting, depending on the version of the model.

Finally, we demonstrate how our model can be used to recover estimates of household willingness to pay for each electricity service. Both models find that the mean (across billing cycles and customers) willingness to pay for an additional hour of an energy service is highest for appliances and then business uses. The lowest mean willingness to pay is for lighting. These results suggests a number of policies to increase the willingness of customers to pay for the electricity they consume that we discuss in Section VIII.

The remainder of the paper proceeds as follows. Section II sets the context for this research by describing the current challenges facing developing country power sectors and the Indian power sector in particular and summarizes previous research that is related to our research either topically or methodologically. Section III describes the datasets used in our analysis and presents descriptive statistics from these data to provide context for our analysis. Section IV describes the two econometric models of energy services demand. Section V describes the estimation procedure and results. Section VI estimates the willingness to pay for the four categories of electricity services for households in Rajasthan. We conclude and discuss the policy implications of our results in Section 7.

2 Background

This section first describes the revenue shortfall relative to production cost problem that is common to virtually all developing country electricity supply industries. We then summarize dire financial conditions facing most electricity retailers in India. We then discuss the use of increasing block tariffs for residential electricity consumers in India and how they impact modeling the electricity demand for electricity and other infrastructure services.

2.1 Two energy worlds

Burgess et al. (2020) introduce the distinction between two energy worlds. In the first, citizens enjoy universal access to electricity 24 hours per day. In the second many citizens are not connected to the transmission and distribution network and those that are have an unreliable supply that many customers don't fully pay for. The authors argue that a major reason for the persistence of the extreme version of the second world is that the governments in these countries treat electricity as a right whether or not the customer pays for it. This leads to significant revenue shortfalls relative to the cost of the electricity delivered to consumers.

These revenue shortfalls can occur both because customers with formal connections do not pay their bill or customers with informal connections consume without being billed. There have been a number of studies exploring the determinants of this behavior. Yurtseven (2015) considers the case of Turkey and finds a number of predictors of this behavior such as whether a region is rural, low income, or has higher electricity prices. Depuru et al. (2010) present across-country evidence that supports these factors as important predictors of revenue shortfalls. Although there are many reasons that customers do not pay their bills or establish informal connections, the final outcome is that customers appear to value an unreliable, but low cost, supply of electricity over a possibility more reliable supply that they pay for.

One reason for customers to pay for the electricity in first energy world is because they have a high marginal willingness to pay for an additional hour of at least one electricity service. For example, customer with large refrigerator filled with food that will spoil unless it is kept cold has a high marginal willingness to pay for an additional hour of appliance services. Customers with a minimal amount of electricity consuming capital, such a few lights and a fan, are significantly less likely to be willing to pay for a reliable supply of the electricity than customers that own expensive appliances that consume significantly more electricity per hour of use and are more likely to be permanently damaged by an unreliable supply of electricity.

By estimating the customer-level demand for electricity services, we can recover estimates of the marginal willingness to pay for an additional hour of each electricity service by each customer in our sample. This information identifies individual customers that are likely to be willing to pay for reliable supply of electricity. Distribution utilities can focus their grid upgrade efforts on ensuring a reliable supply of electricity in regions with customers that have a high marginal willingness to pay for at least one electricity service.

2.2 Indian electricity supply industry experience

Electricity for household consumption accounts for approximately a quarter of all electricity sales in India. A sizable portion of these sales are made by publicly-owned electricity distribution companies at prices significantly lower than the average revenue required to recover costs, which puts much of India squarely in the second energy world. Burgess et al. (2020) provides significantly more detail on the Indian experience.

Electricity prices in Rajasthan, like much of rest of India, follow an increasing block tariff (IBT) structure for energy and a fixed charge. As shown in Figure 1, the marginal price is the same for all consumption on a block or range of monthly consumption but increases for higher amounts of consumption. Rajasthan's tariff schedule also has a lower first-block energy charge for households possessing a below poverty line (BPL) card issued by the government. BPL households are required to pay the usual energy and fixed charges if their consumption exceeds more than 50 kWh during a billing cycle.

Using a household survey, we use appliance level ownership and usage data to estimate a demand model for energy services. In estimating this model, we account for consumption-specific charges under the IBT price schedule and variation in weather conditions within a household's billing cycle. The model also incorporates heterogeneity in residential demand due to observable socio-economic characteristics-–such as, income, family sizes and other observable factors that differ across customers. In addition to household survey data, the analysis also uses an administrative billing data set containing four years of metered consumption data for every household in two districts of Rajasthan – Jaipur and Alwar – served by a local electricity utility called JVVNL.

In the Indian context, two data sources that have been previously used to study electricity demand are the National Sample Survey Office's (NSSO) consumption expenditure rounds and panels of the India Human Development Survey. Both sources are known to miss the top of the income distributions and report electricity consumption based on self-reported figures. These surveys are not designed to accurately measure residential electricity consumption and therefore do not capture the full distribution of electricity consumed in a region. Figure 2 compares the consumption distribution from the administrative billing data of JVVNL consumers from 2015-16 to NSSO's consumption expenditure survey of 2011-12 for districts served by JVVNL. The figure highlights the fact that, assuming distribution neutrality between these sets of years, NSSO consumption data is particularly imprecise in capturing the top ends of the consumption distribution.

Our data on a household's billing cycle-level electricity consumption is obtained from the household's actual electricity bill. We also use data on appliance holding and hours of use of each appliance for the inventory of energy-consuming appliances owned by the household from our survey data. Our model of the demand for energy services also depends on observable socioeconomic characteristics of the household from our survey and weather conditions during the household's billing cycle. This allows us to recover the price-responsiveness of demand into four categories of energy services demand as well as the overall responsiveness of electricity demand to the price of electricity.

2.3 Customer demand under IBT schedules

IBTs are not unique to India and the estimation of residential electricity demand under IBT schedules has been studied extensively in industrialized country settings. Reiss and White (2005) use annual data from a sample of California households to estimate electricity demand under IBT pricing. The model is then used to analyze the effect of tariff changes on changes in consumption and the share total monthly expenditure a household spends on electricity.

McRae (2015) conducts a similar exercise in a developing country setting using billing cycle level data, rather than annual data. He uses appliance ownership data from the Colombian census paired with the customer's utility billing data to estimate his model of customer-level billing cycle demand under nonlinear pricing. McRae uses the demand estimation under non-linear pricing econometric modeling framework developed by Hanemann (1984) to recover the parameters of household-level preference functions.

More recently, Wang and Wolak (2022) introduce a new model customer-level demand under IBTs that accounts for the fact that customers cannot precisely control their consumption during the billing cycle. Instead, customers make at the beginning of the billing cycle estimate of what marginal price they will face at the end of the model. Conditional on this expected utility maximizing marginal price choice, they consume for the entire billing. The authors apply this modeling framework to two samples of water utility customers facing IBTs.

3 Data Used in Analysis

The data used in this paper are from three different sources: (i) administrative data on billing cycle-level electricity consumption and bills issued; (ii) detailed household demographic characteristics and appliance ownership and use from a survey designed and implemented in 2017; and, (iii) daily temperature and precipitation data. Each of these data sources are described in below with additional details in Appendix B.

3.1 Electricity consumption and prices data

We collected administrative cycle-level billing data directly from JVVNL and exclude all non-residential consumer bills from it. The dataset has a unique connection identifier which is used to match households across survey and administrative datasets. The dataset provides the total amount electrical energy consumed during the billing cycle, the calendar dates on which the meter was read, the total energy charges and fixed charges, electricity duties, subsidies, and other charges included in the bill. It does not record the consumption tier and the corresponding marginal price of electricity.

We back out the consumption and marginal prices by calculating the per-unit energy price from the total energy cost provided in the billing dataset and comparing it to the per-unit charges across consumption tiers as prescribed in the tariff schedule. Appendix B contains an example of this exercise.

JVVNL generally produces bills at a bimonthly frequency, depending on the schedule of a roving meter reader. Meter readers walk a different route over each day of a bimonthly period to visit residences and compile the total kWhs consumed since their last visit. This results in different sets of customers having different bi-monthly billing cycles based on the days on which the meter reader can reach the customers residence. To guard against possible billing errors, we exclude observations for which the billing cycle is more than 180 days.

If the meter reader is unable to establish contact with the household or if the meter appears to be malfunctioning at the time of reading, the consumption figures for the household are not captured in the billing dataset. In such cases, JVVNL makes an imputation based on the average consumption of the household over the past six months. We drop these observations from our sample, resulting in exclusion of an additional 1,250 observations (13% of the data).

To provide an overview of prices and consumption variables in the dataset, Table 1 shows the summary statistics for all households over 2014-2017 period. The average monthly per capital consumption of electricity in the sample is 39.7 kWhs – more than double the per capital residential usage of 15 kWh in Rajasthan in 2012 according to Prayas (2016). The increase in the per capital electricity consumption reflects robust economic growth and rapid rates of electrification in the region.

The marginal prices on the increasing block tariff for our sample of customers was revised twice during our sample period. Figure 3 shows the distribution of average daily household consumption by the month, with red labels in the horizontal axis indicating the periods in which these tariff changes occurred. There is clear seasonality in the consumption levels– with peak and troughs during the summer and winter months respectively. We refer the reader to Appendix B for additional validation checks conducted on the dataset.

We also have outage data for our sample period for the distribution network served by each household in our sample. We merge this information with household location data to obtain household-specific outage data.

3.2 Household characteristics

The household characteristics we include in our model are taken from a survey of households carried out in two districts of Rajasthan. The survey, conducted as part of World Bank electricity lending program in the state, was administered by a local team of fieldworkers with extensive experience working and residing in these areas. The goal of the survey was to collect household demographics, socio-economic data, appliance ownership and use, and unique household level billing codes to match survey data to the billing database. The survey enumerated approximately 2,000 households.

Our data on a household's billing cycle-level electricity consumption is obtained from the household's actual electricity bill. We also use data on appliance holding and hours of use of each appliance for the inventory of energy-consuming appliances owned by the household from our survey data. Our model of the demand for energy services also depends on observable socioeconomic characteristics of the household from our survey and weather conditions during the household's billing cycle. This allows us to recover the price-responsiveness of the demand for the four categories of energy services as well as the overall responsiveness of electricity demand to the price of electricity.

Table 2 shows that the consumption distribution of the sampled households, weighted by the sampling weights, is able to replicate the population consumption distribution from the billing cycle-level dataset for all customers in the two districts of Rajasthan for our sample period. These matching patterns allay concerns related to sample selection in the overall distribution, except from the 95th percentile onwards of the consumption distribution–where our sample appears to represent a lower fraction of households than in utility's customer database.

To measure household-level income, the survey contained several modules pertaining to farming, raising livestock, self-employment, casual labor activities, salaries from jobs and remittance earnings for each adult member of the household. Total household income was calculated as the sum of incomes across each of these modules and household members. The survey had a non-response rate of approximately 9% on income questions, so we exclude these households from our base sample. To keep the survey tractable and short, we did not capture household consumption expenditure or assets and liabilities information. As a result, we are unable to calculate disposable income separately from total household income. Finally, we exclude household-bills for which the annualized electricity bill amount was greater than 75% of annual reported household income. This restriction is imposed to account for the fact that households also need to pay for food and other essentials besides electricity throughout the year.

After excluding observations based on the income criteria and after matching household survey data to the administrative data, we are left with a observation count of 7,615 billing-cycle level observations comprising of 805 unique customers. Ninety-six percent of these observations are from consecutive cycles, implying that a majority of households do not sort in and out of the panel. Table 3 summarizes socioeconomic characteristics of the sampled households and shows the share of households by various employment categories.

The survey also collected information on ownership and intensity of use of electrical appliances, allowing us to estimate the demand for residential energy services. We categorize each of the 24 appliances covered in the survey to one of the four energy service categories (as shown in Appendix A). The energy demand for each of the four energy services is calculated by summing up the hours of energy services demand for all electricity consuming capital goods in that category. We further restrict our sample to household-bills that have positive energy demand for all four services, which leaves us with 3977 billing-cycle observations and 611 unique customers. In Figure 4, we plot the histograms of the service demand for the four categories in hours of use per day.

3.3 Temperature and precipitation data

The daily temperature and rainfall data are available at a 1 degree latitude by 1 degree longitude and 0.25 degree by 0.25 degree gridded resolution from Srivastava et al. (2009) and Pai et al. (2014), respectively. We match the locations of the villages and census enumeration blocks to these grids to obtain household-level measures of temperature and rainfall. Figure 7 shows deviations of daily temperature and rainfall in an area from its three-year period average (2014-2017). The precipitation curve shows that there is little variation in rainfall across time and regions. The low average precipitation levels in the region also explains the high ownership rates of evaporative air coolers as observed in the survey data. The temperature plot highlights the substantial heterogeneity in climatic conditions across villages and enumeration blocks even within the same month. This heterogeneity in temperatures across regions is exploited in later sections to estimate household-level demand.

4 Econometric Model

The development of our econometric model of the billing cycle-level household demand for electricity services proceeds in three steps. First, we present our model of the customer-level outage distribution. Second, we present the household-level electricity consumption function relating the household's energy services demands to its electricity consumption. Third, using results from the previous two steps, we specify a household-level utility function for energy services and estimate it using the first-order conditions for expected utility maximizing choices of the electricity services.

4.1 Outage Factor

The outage factor is defined as the fraction of time during the billing cycle that the electricity is unavailable to the household. For households included in our dataset, the average outage factor is 0.3. This can have significant impact on both household's electricity consumption function and demand model. Model how households take into account random outages in making their electricity service choices, we first need to model the distribution of the outage factor, since it is a random variable.

We plot the histogram of the outage factor in Figure 8. From the histogram, we can see that most of the outage factors fall into two clusters, the range of 0 to 0.2 or the range of 0.6 to 0.8. Therefore, we model the outage factor as a mixture of two beta distributions, with the probability p of having distribution $Beta(a_1, b_1)$ and probability $1-p$ of having distribution $Beta(a_2, b_2)$. We choose beta distribution mainly due to its flexibility and the fact that it has support on the interval (0,1).

Further, different household may have different outage factor distributions. For example, electricity grid might be more reliable in high-income neighborhood, or less reliable in rainy days. Therefore, We allow the probabilities p and parameters of the two beta distributions (a_1, b_1, a_2, b_2) to vary based on household characteristics and weather information. Specifically, let $\beta_p, \beta_{a1}, \beta_{b1}, \beta_{a2}, \beta_{b2}$ be the parameters that we want to estimate. Let X_i be the household-level variables for household i, including both household characteristics and weather information. Then the outage factor distribution for household i is the mixed distribution with probability $exp(\beta_p X_i)/(1 + exp(\beta_p x_i))$ of having distribution $Beta(exp(\beta_{a1} X_i), exp(\beta_{b1} X_i))$ and probability $1/(1 + exp(\beta_p x_i))$ of having distribution $Beta(exp(\beta_{a2} X_i), exp(\beta_{b2} X_i))$. In addition, consistent with our earlier discussion, the reliability of electricity grid can be quite different in urban versus rural areas. Therefore, we estimate the outage factor distribution separately for urban and rural areas.

4.2 Electricity Consumption Function

Household-level electricity consumption depends on the hours of various categories of electricity services demand by the household–namely heating and cooling, lighting, household appliances for domestic end-uses, and appliances for business end-uses. However, the household pays for electricity consumption that results from their demand for these services. Therefore, we need to characterize the relationship between the household's electricity consumption and household's hours of use of each of four electricity services.

Let s_i equal the household's demand for energy service i in hours of use, $\mathbf{s} = (s_1, s_2, ..., s_N)'$ equal the vector of energy services demand for the N services that the household consumes. Let $S_i = ln(s_i)$ and $S = (S_1, S_2, ..., S_N)'$. Let e equal the household's electricity consumption in kilowatt-hours (KWh) and $E = ln(e) = f[S] + \epsilon$ where $f[S]$ is the household's electricity consumption function that converts the vector of the logarithms of individual energy services into the logarithm of electricity use, and ϵ is a random variable that is unobserved by the household and the researcher that captures the technological uncertainty in amount of electricity (e) consumed by a given vector of the logarithm of energy services S. We use two approaches to characterize the function $f(\cdot)$. The first takes a nonparametric approach that does not account for observable customer-level heterogeneity. The second specifies a parametric model that imposes restrictions not rejected by our nonparametric approach that accounts for significant observable heterogeneity across customers in the electricity consumption function.

4.2.1 Constrained Kernel Regression

Our nonparametric model for the electricity consumption function imposes two restrictions consistent with the physics governing electricity service use and electricity consumption. First, we expect that an increase in any electricity consuming service should increase electricity consumption. Therefore, we impose the constraint that the partial derivative of $f[S]$ with respect to any element of S is positive.

Second, we expect that the electricity consumption function should at least exhibit constant return to scale. For example, if we double the demand for all the categories of service, we expect the electricity consumption to at least double as well. In terms of the logarithm, this means that the sum of the partial derivatives of $f[S]$ with respect to the four elements of S should sum up to at least 1. As we discuss below, the validity of both of these sets of restrictions can be tested empirically using our dataset.

4.2.2 Parametric Functional Form with Customer-Level Heterogeneity

Our preferred alternative is to specify an parametric functional form for the electricity consumption function that allows for differences in this function across households and billing cycles based on household demographic characteristics and weather information. Let w_{jt} be the household-level information for household j for billing cycle t. Let s_{ijt} equal the household j's demand for energy service i in hours of use during billing cycle t. Let $S_{ijt} = ln(s_{ijt})$ and $S_{jt} = (S_{1jt}, S_{2jt}, ..., S_{Njt})'$. Let e_{jt} equal the household j's electricity consumption during billing cycle t in kilowatt-hours (kWh) and $E_{jt} = ln(e_{jt}) = f_{jt}[S_{jt}] + \epsilon_{jt}$.

We assume that

$$f_{jt}[S_{jt}] = \sum_{i=1}^{N} \alpha_{ijt} S_{ijt} + \delta_j$$

where

$$\alpha_{1jt} = \frac{1}{1 + \sum_{k=2}^{N} exp(w'_{jt}\gamma_k)}$$

$$\alpha_{ijt} = \frac{exp(w'_{jt}\gamma_i)}{1 + \sum_{k=2}^{N} exp(w'_{jt}\gamma_k)} \text{ for } i = 2, 3, \ldots, N$$

We assume the electricity consumption function to be a household fixed effect plus a linear combination of service demand, allowing the coefficient to vary based on household-level variable and weather information for that billing cycle. Note that the functional form guarantees that the partial derivatives of the logarithm of electricity consumption with each respect to the logarithm of each service demand are non-negative and sum up to 1.

4.3 Model of Electricity Services Demand

Our model of the demand for energy services makes use of data from the household survey of appliance use to account for the well-known fact that electricity is a derived demand. The model also accounts for the fact that customers outages in the distribution grid which can prevent them from consuming electricity during the billing cycle. Our model implies that a household's realized demand for electricity is derived from its demand for each energy service provided by each electricity-consuming capital good owned such as a light bulb, fan, or air conditioner. Moreover, the amount of electricity consumed to provide a fixed quantity of electricity services, say an hour of computer use, is uncertain because of factors such as the background temperature and intensity of use of the appliance as well as outages. We now describe the household demand model incorporates both the outage factor distribution and the electricity consumption function to recover the household's demand for the four energy services.

Let s_i equal the household's demand for energy service i in hours of use, $s = (s_1, s_2, ..., s_N)'$ equal the vector of energy services demand for the N services that the household consumes. Let $S_i = ln(s_i)$ and $S = (S_1, S_2, ..., S_N)'$. Let e equal the household's electricity consumption in kilowatt-hours (KWh) and $E = ln(e) = f[S] + \epsilon$ where $f[S]$ is the household's electricity consumption function which is described previously, and ϵ is a random variable that is unobserved by the household and the researcher.

To account for impact of distribution network outages, S is assumed to decomposed into a planned vector of energy services used, $\mathbf{S}$, and a random outage fraction, AF, which is a random variable defined on the interval (0,1) and equal to the fraction of hours in the billing cycle that electricity is available to the household. The distribution is described previously. These variables are assumed to satisfy the equation: $S = \mathbf{S} + \iota * ln(AF)$, where $\iota = (1, 1, 1, 1)'$, so that all services consumed are reduced by the same proportion, AF, as a result of outages within the billing cycle.

Let $p(e)$ equal the potentially nonlinear price schedule that the household faces, where $p(e)$ is the marginal price paid at electricity consumption level e. Let $T(e^*) = \int_0^{e^*} p(e)de$ equal household's total bill under nonlinear price schedule $p(e)$ for consumption level, e^*. Note the $T(e)$ includes any fixed charge that must be paid regardless of the household's monthly consumption. Suppose the household consumes a composite "outside" good besides electricity that equals the difference between the households's total monthly income and its electricity bill. Let x_{it} equal this "outside" good expenditure for household i during billing cycle t. The distribution x_{it} across households and billing cycles is presented in Figure 9).

The household is assumed to have the preference function, $U(S, X, A, W)$ where $S = \mathbf{S} + \iota *$

ln(AF) and $X = ln(x)$. This utility function depends on the vector of realized energy services consumed by the household, S, its demand for x and observable characteristics of the household, A, and a vector of weather variables, W. Note that because of the outages during the billing cycle the household's desired vector of energy services, **S**, is not equal the actual energy services is consumes, S. The household's budget constraint is equal to $T(e)+p_x x \leq M$, where p_x is the price of x, M is the household's income, and $e = exp(f[\mathbf{S}+\iota*\ln(AF)]+\epsilon)$. In terms of the $(S', X)'$, the budget constraint becomes $M = T\{exp(f[\mathbf{S}+\iota*ln(AF)]+\epsilon)\}+p_x exp(X)$.

Each billing cycle the household is assumed to choose their consumption of each energy service to maximize expected utility (where the expectation is taken with respect to the technological uncertainty ϵ and the outage uncertainty, AF). The problem takes the form:

$$\max_{\mathbf{s},x} E_{\epsilon,AF}[U(\mathbf{S} + \iota * ln(AF), X, A, W) \mid (A, W)] \tag{1}$$
$$\text{subject to } T\{exp(f[\mathbf{S} + \iota * ln(AF)] + \epsilon)\} + p_x exp(X) \leq M$$

where $E_{\epsilon,AF}[(\cdot)|(A, W)]$ implies taking the expectation with respect to the distribution of ϵ and AF conditional on the values of (A, W). Using the budget constraint to solve for the demand for X given the demand for **S** yields:

$$X = ln[M - T\{exp[f(\mathbf{S} + \iota * ln(AF)) + \epsilon]\}] - ln(p_x) \tag{2}$$

Substituting into the household's utility function yields the equivalent problem to (5.2.1):

$$\max_{\mathbf{S}} E_{\epsilon,AF}[U(\mathbf{S}+\iota*ln(AF), ln[M-T\{exp[f(\mathbf{S}+\iota*ln(AF))+\epsilon]\}]-ln(p_x), A, W)|(A, W)] \tag{3}$$

which has the following first-order conditions:

$$\frac{\partial E_{\epsilon,AF}[U(\mathbf{S} + \iota * ln(AF), ln[M - T\{exp[f(\mathbf{S} + \iota * \ln(AF)) + \epsilon)\}] - ln(p_x), A, W)|(A, W)]}{\partial S_j} = 0$$
$$\text{for } j = 1, 2, ..., N \tag{4}$$

Switching the order of integration and differentiation yields:

$$\frac{E_{\epsilon,AF}[\partial U(\mathbf{S} + \iota * ln(AF), ln[M - T\{exp[f(\mathbf{S} + \iota * ln(AF)) + \epsilon]\} - ln(p_x), A, W)|(A, W)]}{\partial S_j} = 0 \tag{5}$$

This implies:

$$E_{\epsilon,AF}\{\frac{\partial U(\mathbf{S}+\iota * ln(AF), ln[M-T(e)]-ln(p_x), A, W)}{\partial S_j}$$
$$-\frac{\partial U(\mathbf{S}+\iota * ln(AF), ln[M-T(e)]-ln(p_x), A, W)}{\partial X}\frac{p[e]e}{M-T(e)}\frac{\partial f}{\partial S_j}|(A,W)\} = 0 \quad (6)$$
$$\text{for } j = 1, 2, ..., N.$$

Note that $ln(e) = f(\mathbf{S}+\iota * AF) + \epsilon$, where $\mathbf{S}$ is observed by the econometrician. Also note $\frac{\partial f(\mathbf{S} + \iota * AF)}{\partial S_j}$ depends on the realization of AF. Therefore, the first-order conditons for the expected utility-maximizing choices of the $N = 4$ energy services become:

$$E_{\epsilon,AF}\{\frac{\partial U(\mathbf{S}+\iota * ln(AF), ln[M-T(e)]-ln(p_x), A, W)}{\partial S_j}$$
$$-\frac{\partial U(\mathbf{S}+\iota * ln(AF), ln[M-T(e)]-ln(p_x), A, W)}{\partial X}\frac{p[e]e}{M-T(e)}\frac{\partial f(\mathbf{S}+\iota * AF)}{\partial S_j}\} = 0 \quad (7)$$
$$\text{for } \quad j = 1, 2, ..., N$$

5 Estimation Procedure and Results

In this section, we describe how we estimate the econometric models described in the previous section and present the estimation results.

5.1 Outage factor distribution

We estimate the parameters of the mixed distribution by maximizing the overall likelihood. We start with the specification that includes both household demographic variables and billing cycle-level weather information. We then test whether the parameters for household demographic variables are jointly equal to 0. For urban area model, we cannot reject the null hypothesis that household demographic variables predict differences in the outage factor. Thus we only include weather information for the rural area, but include both weather information and household demographics for the urban area. The included variables and the parameter estimates for both models are presented in Table 4.

5.2 Electricity consumption function

In this subsection, we present our estimates of our two approaches to estimate the electricity consumption function. For both models we plot the histogram of the partial derivatives of

E with respect to the elements of S.

5.2.1 Constrained Kernel Regression

To perform the constrained kernel regression, we follow the procedure outlined in Du et al. (2013). The key idea is to re-weight data points that make up the kernel regression so that the estimated kernel model satisfies all the constraints. Let $\{Y_i, X_i\}_{i=1}^M$ denote sample data we have on electricity consumption and energy service demand. We are trying to estimate the electricity consumption function $f(\cdot)$, while $Y_i = f(X_i) + \epsilon_i$. It is not possible to impose the constraints on all points in the function support. Thus we have to choose a subset of points in the function support to impose the constraints. Let $\{X_i^c\}_{i=1}^K$ denote such subset of points.

Consider a generalized kernel regression with Nadaraya-Watson estimator in the following form

$$\hat{f}(x) = \sum_{i=1}^{M} p_i A_i(x) Y_i$$

where $A_i(x) = MK_h(X_i, x)/\sum_{j=1}^M K_h(X_j, x)$. For an unrestricted kernel regression, we take $p_i = 1/M, i = 1, \ldots, M$. Since we have certain constraints on the derivative of $f(\cdot)$ described in section 4.2.1, we need to select p_i to satisfy those constraints. Let p_u be the M-vector of uniform weights and let p be the vector of weighs to be selected. We choose p to minimize the distance from p to p_u but still satisfy our constraints. Here we allow for both positive and negative weights while retaining $\sum_i p_i = 1$. We define distance as $D(p) = (p_u - p)'(p_u - p)$.

For constraints, let $\hat{f}^w(X)$ denote the derivative of $\hat{f}(\cdot)$ with respect to the wth service. The two constraints are then:

$$\hat{f}^w(X) > 0 \text{ for } w = 1, 2, 3, 4$$

$$\sum_{w=1}^{4} \hat{f}^w(X) = 1$$

In practice, it is very hard to satisfy the equality constraint. Thus we relaxed the constraint on the sum of the partial derivatives to lie in the range between 0.98 and 1.02. As a robustness check, we tried other ranges and got similar results.

To sum up, we are solving the following optimization problem:

$$
\begin{aligned}
\min_p (p_u - p)'(p_u - p), \text{ subject to} \\
\sum_i p_i = 1 \\
\hat{f}^w(X_i^C) > 0 \text{ for } w = 1, 2, 3, 4, i = 1 \dots K \\
0.98 \leq \sum_{j=1}^{4} \hat{f}^j(X_i^C) \leq 1 \text{ for } i = 1 \dots K
\end{aligned} \tag{8}
$$

Now we need to decide what data we use to estimate the kernel regression, namely, $\{Y_i, X_i\}_{i=1}^M$. It is very computationally intensive to perform the constrained kernel regression on bill-level data (3977 data points), which means optimizing the objective function over 3977 parameters under thousands of constraints. In addition, bill-level data are more volatile and susceptible to measurement error that they create more extreme data points, which cause extra difficulty for the kernel estimation and constraints. Thus for computing reason, we aggregate the household-bill level data into household level data (611 household) to decrease the size of the optimization problem. For each household, we calculate its daily average of electricity consumption and service demand over the entire period considered as its observation.

We also need to decide which points we impose the constraints on, namely, $\{X_i^c\}_{i=1}^K$. Following Du et al. (2013), we impose the constraints on all the points that we use to estimate the model, namely $\{X_i\}_{i=1}^M$. In addition, we impose the constraints on a grid with 7 points on each dimension, with the 7 points evenly placed, and covered the main range of support for each dimension (a total of $7^4 = 2401$ grid points).

The estimated partial derivatives of $\hat{f}(\cdot)$ with respect to the elements of energy service are presented in Figure 10. One may notice that some gradients fall out of the range between 0 and 1, which seems to violate the constraints we impose. This is because we impose the constraints on the household-level data points, while the gradients presented are on the bill-level data points. Out of the 3977 bill-level data points, more than 3400 data points satisfy both type of constraints. For the points that do not satisfy both constraints, they are usually not far off. To be consistent and comparable with the parametric model, we still keep these data points in the sample.

Lastly, we want to check whether the constraints we imposed are valid. Following Du et al. (2013), we performed a hypothesis test on whether the constraints are correct. We use a bootstrap approach. We first calculate the residual $\hat{\epsilon}_i$ using the model and parameter estimates discussed above, and then generate resamples for Y_i via *iid* residual resampling. These resamples are generated under the null hypothesis that these constraints are correct. By recomputing $D(p^*)$ for the bootstrap sample , we can generate the null distribution of $D(\hat{p})$. We compute the empirical P value, P_B, as the proportion of the bootstrap resamples $D(p^*)$ that exceed $D(\hat{p})$. In our case, we repeat the process 100 times, and get a empirical P-value of 0.96, thus failing to reject the null hypothesis.

5.2.2 Parametric model with customer-level heterogeneity

Our parametric model with customer-level heterogeneity takes the following form

$$E_{jt} = \sum_{i=1}^{N} \alpha_{ijt} S_{ijt} + \delta_j + \eta_{jt}$$

where

$$\alpha_{1jt} = \frac{1}{1 + \sum_{k=2}^{N} exp(w'_{jt}\gamma_k)}$$

$$\alpha_{ijt} = \frac{exp(w'_{jt}\gamma_i)}{1 + \sum_{k=2}^{N} exp(w'_{jt}\gamma_k)} \text{ for } i = 2, 3, \ldots, N$$

and the η_{jt} are a sequence of independent identically distributed mean zero random variables with finite support. We fit the parameters of the model by nonlinear least squares. The estimates for γ are presented in Table 5. Note that we treat heating and cooling as the "first" service, so that there are no γ for heating and cooling. Similarly, we plot the distribution of the derivative of E_{it} with with respect to the four elements of S_{it} across customers and billing cycles in Figure 11.

5.3 Model of electricity services demand

To use the four sets of moment conditions in equation (4.3) to estimate the parameters of the household's preference function. In order to facilitate the estimation process and simulation of counterfactual solutions, we assume the following functional form for $U(S, X, A, W)$. Let $Z = (S', X)'$. Define

$$U(Z) = a(A, W)'(Z - \bar{Z}) + \frac{1}{2}(Z - \bar{Z})'\Gamma(Z - \bar{Z})$$

where $a(A,W) = (exp(V_1'\beta_1), ..., exp(V_{N+1}'\beta_{N+1}))'$ is a $(N+1)\times 1$ vector of positive functions of elements of A and W for that household and billing cycle and $\bar{Z}$ is the sample mean of the vector Z. The vectors V_1 to V_{N+1} are composed of elements of A and W expressed in deviations from the mean, which implies the sample mean of the non-constant elements of V_i are zero for $i = 1, 2, ..., N+1$. Note the the constant term in V_{N+1} must be normalized to equal 1. This sets the scale of the cardinal utility function $U(Z)$. The elements of $a(A,W)$ will allow the marginal utility of consuming each electricity service to differ across households and billing cycles. The matrix Γ is assumed to be symmetric and negative definite, which implies the following Cholesky factorization: LDL', where L is a lower triangular matrix with 1's along the diagonal and D is a diagonal matrix with $d_{kk} \leq 0$ for all k.

In terms of this notation, the gradient vector of $U(Z)$ is:

$$\frac{\partial U(Z,A,W)}{\partial Z} = a(A,W) + \Gamma(Z - \bar{Z})$$

Note that the first four elements of this vector are $\frac{\partial U(Z,A,W)}{\partial S_j}$ for $j = 1$ to N and the last element is $\frac{\partial U(Z,A,W)}{\partial X}$.

Let the index i denote households, t denote billing cycles, and j denote energy services. Let

$$\ell_{jit}(\theta, e_{it}, AF_{it}) = H_{it}[\frac{\partial U(\mathbf{S} + \iota * ln(AF_{it}), ln[M - T(e_{it})] - ln(p_x), A, W)}{\partial S_{jit}}$$
$$-\frac{\partial U(\mathbf{S} + \iota * ln(AF_{it}), ln[M - T(e_{it})] - ln(p_x), A, W)}{\partial X_{it}} \frac{p(e_{it})e_{it}}{M - T(e - it)} \frac{\partial f(\mathbf{S} + \iota * AF)}{\partial S_j}] \quad (9)$$

where $H_{it} = (1, A_i', W_{it}', AW_{it}')'$ and θ is the vector of parameters to be estimated: β_i for $i = 1, 2, ...(N+1)$ and the elements of L and D, where $\Gamma = LDL'$. The vector AW_{it} is composed of functions of elements of the A_i and W_{it} sufficient to identify all of the parameters of model.

Let $f_{AF}(.|i, t, e_{it})$ equal the estimated density of $AF_{i,t}$ given the value of e_{it}, household i's electricity consumption in billing cycle t. Draw K values from this density and substitute it into the above equation and compute:

$$\ell_{jit}(\theta, e_{it}) = \frac{1}{K}\sum_{k=1}^{K} H_{it}[\frac{\partial U(\mathbf{S} + \iota * ln(AF_{it}^{k}), ln[M - T(e_{it})] - ln(p_x), A, W)}{\partial S_{jit}}$$
$$- \frac{\partial U(\mathbf{S} + \iota * ln(AF_{it}^{k}), ln[M - T(e_{it})] - ln(p_x), A, W)}{\partial X_{it}} \frac{p(e_{it})e_{it}}{M - T(e_{it})} \frac{\partial f(\mathbf{S} + \iota * ln(AF_{it}^{k}))}{\partial S_j}] \tag{10}$$

Stacking the $\ell_{jit}(\theta, e_{it})$ into a vector for each energy service yields:

$$\ell_{it}(\theta, e_{it}) = (\ell_{1it}(\theta, e_{it})', \ell_{2it}(\theta, e_{it})', ..., \ell_{Nit}(\theta, e_{it})')'.$$

Solving for the values of θ that minimizes:

$$\min_{\theta} L(\theta)'\Sigma L(\theta) \tag{11}$$

where

$$L(\theta) = \frac{1}{I}\sum_{i=1}^{I} \frac{1}{T(i)} \sum_{t=1}^{T(i)} \ell_{it}(\theta, e_{it}) \tag{12}$$

and $T(i)$ is the number of billing cycles available for household i, and Σ is a positive definite weighting matrix of the same dimension as the number of rows of $\ell_{it}(\theta, e_{it})$, will yield a consistent estimate of θ. Starting with $\Sigma = I$ the identity matrix is the most straightforward.

Given a consistent estimate of θ, say $\hat{\theta}$, compute

$$V(\hat{\theta}) = \frac{1}{I}\sum_{i=1}^{I}[\frac{1}{T(i)}\sum_{t=1}^{T(i)} \ell_{it}(\hat{\theta})][\frac{1}{T(i)}\sum_{t=1}^{T(i)} \ell_{it}(\hat{\theta})]'$$

Solving for the value of θ that minimizes:

$$\min_{\theta} L(\theta)'V(\hat{\theta})^{-1}L(\theta) \tag{13}$$

yields consistent and asymptotically efficient estimates of θ. Call this value $\tilde{\theta}$.

Estimated variance of these estimates can be computed as:

$$\frac{1}{I}[D(\tilde{\theta})'V(\hat{\theta})^{-1}D(\tilde{\theta})]^{-1}$$

where

$$D(\theta) \;=\; \frac{1}{I}\sum_{i=1}^{I}\frac{1}{T(i)}\sum_{t=1}^{T(i)}\frac{\partial \ell_{it}(\theta)}{\partial \theta}$$

When estimating the service demand model with constrained kernel regression as the electricity consumption function, some of the parameter estimates were very imprecisely estimated because this model does not allow for customer-level heterogeneity in the electricity consumption function. Therefore, when performing the optimization problem that computes the estimates of the elements of θ, we add a lasso regularization term to zero out some of the parameters. These parameters are shown as zero with NA standard error when we present the estimation results for this model. This problem did not arise for our preferred specification which allowed for customer-level and weather-dependent heterogeneity in the electricity consumption function.

The parameter estimates for the service demand model when using constrained kernel regression for the electricity consumption function are presented in Table 6 and 7. The parameter estimates for the service demand model using the electricity consumption function with customer and billing cycle weather heterogeneity are presented in Table 8 and 9. We also report the Hansen (1982) test for the validity of the over-identification restrictions for each model. For both models we do not find any evidence against the null hypothesis of the validity of these over-identifying restrictions.

6 Implications of Model

6.1 Service demand elasticities

Our demand models can be used to calculate the price elasticity of demand and income elasticity of demand for each energy service. Let p_E be the marginal price of electricity for the household, and the marginal price for energy service i as $ps_i = \frac{\partial exp(f(S_i))}{\partial exp(S_i)} p_E = \frac{exp(f(S_i))}{s_i}\frac{\partial f(S_i)}{\partial S_i} p_E$, which is the product of the marginal price of electricity and increase electricity consumption associated with one hour increase in the use of energy service i.

For the purpose of calculating elasticity with respect to service demand, we reformulate the household's the utility maximization problem as a standard household choice problem subject to a linear budget constraint using the above marginal prices of each energy service.

The solution to this problem yields conventional demand functions that depend on the values of ps_i for $i = 1, 2, .., N$ and the price of the outside good:

$$\max_Z U(Z) \text{ subject to } \sum_{i=1}^{5} exp(P_i)exp(Z_i) \leq exp(M) \tag{14}$$

where $Z = (ln(s_1), ln(s_2), ln(s_3), ln(s_4), ln(x))$, $P = (ln(ps_1), ln(ps_2), ..., ln(p_x))$ and $M = ln(m)$. Let λ be the Lagrange multiplier. Taking the first order condition yields

$$\frac{\partial U}{\partial Z_i} - \lambda exp(P_i)exp(Z_i) = 0, \forall i \tag{15}$$

$$-\sum_{i=1}^{5} exp(P_i)exp(Z_i) + exp(M) = 0 \tag{16}$$

Note that Z and λ are functions of P and M. Now we take derivative of Equation 15 for Z_i with respect to P_i, P_j for $j \neq i$ and M, we get the following equations:

$$\sum_{k=1}^{5} \frac{\partial^2 U}{\partial Z_i \partial Z_k}\frac{\partial Z_k}{\partial P_i} - \frac{\partial \lambda}{\partial P_i} exp(P_i)exp(Z_i) - \frac{\partial Z_i}{\partial P_i}\lambda exp(P_i)exp(Z_i) = \lambda exp(P_i)exp(Z_i) \tag{17}$$

$$\sum_{k=1}^{5} \frac{\partial^2 U}{\partial Z_i \partial Z_k}\frac{\partial Z_k}{\partial P_j} - \frac{\partial \lambda}{\partial P_j} exp(P_i)exp(Z_i) - \frac{\partial Z_i}{\partial P_j}\lambda exp(P_i)exp(Z_i) = 0 \text{ for } j \neq i \tag{18}$$

$$\sum_{k=1}^{5} \frac{\partial^2 U}{\partial Z_i \partial Z_k}\frac{\partial Z_k}{\partial M} - \frac{\partial \lambda}{\partial M} exp(P_i)exp(Z_i) - \frac{\partial Z_i}{\partial M}\lambda exp(P_i)exp(Z_i) = 0 \tag{19}$$

Similarly, we can take derivative of Equation 16 with respect to P_j for any j and M, we get the following:

$$-\sum_{i=1}^{5} exp(P_i)exp(Z_i)\frac{\partial Z_i}{\partial P_j} = exp(P_j)exp(Z_j) \text{ for any } j \tag{20}$$

$$-\sum_{i=1}^{5} exp(P_i)exp(Z_i)\frac{\partial Z_i}{\partial M} = -exp(M) \tag{21}$$

Denote $\frac{\partial^2 U}{\partial Z_i \partial Z_j}$ as U_{ij}, and $exp(P_i)exp(Z_i)$ as PZ_i. We can aggregate all these equations into a matrix format as:

$$A \cdot \begin{bmatrix} \frac{\partial Z_1}{\partial P_1} & \frac{\partial Z_1}{\partial P_2} & \dots & \frac{\partial Z_1}{\partial P_5} & \frac{\partial Z_1}{\partial M} \\ \frac{\partial Z_2}{\partial P_1} & \frac{\partial Z_2}{\partial P_2} & \dots & \frac{\partial Z_2}{\partial P_5} & \frac{\partial Z_2}{\partial M} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\partial Z_5}{\partial P_1} & \frac{\partial Z_5}{\partial P_2} & \dots & \frac{\partial Z_5}{\partial P_5} & \frac{\partial Z_5}{\partial M} \\ \frac{\partial \lambda}{\partial P_1} & \frac{\partial \lambda}{\partial P_2} & \dots & \frac{\partial \lambda}{\partial P_5} & \frac{\partial \lambda}{\partial M} \end{bmatrix} = B \tag{22}$$

where

$$A = \begin{bmatrix} U_{11} - \lambda PZ_1 & U_{12} & \dots & U_{15} & -PZ_1 \\ U_{21} & U_{22} - \lambda PZ_2 & \dots & U_{25} & -PZ_2 \\ \vdots & & \ddots & & \vdots \\ U_{51} & U_{52} & \dots & U_{55} - \lambda PZ_5 & -PZ_5 \\ -PZ_1 & -PZ_2 & \dots & -PZ_5 & 0 \end{bmatrix}$$

and

$$B = \begin{bmatrix} \lambda PZ_1 & 0 & \dots & 0 & 0 \\ 0 & \lambda PZ_2 & \dots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & \lambda PZ_5 & 0 \\ PZ_1 & PZ_2 & \dots & PZ_5 & -exp(M) \end{bmatrix}$$

Note that $\frac{\partial Z_i}{\partial P_i}$ is own-price elasticity, $\frac{\partial Z_i}{\partial P_j}$ is cross-price elasticity, and $\frac{\partial Z_i}{\partial M}$ is the income elasticity of demand. λ can be calculated as $\frac{\partial U}{\partial Z_5}/PZ_5$, according to Equation 15 for $i = 5$. Solving for the system of equations yield our elasticity estimates. With the price elasticities with respect to service demand, and the electricity consumption function, we can also calculate the overall price elasticity with respect to electricity consumption as following.

$$\begin{aligned} \epsilon_E &= \frac{\partial f(S)}{\partial \ln(p_E)} = \sum_{i=1}^{4} \frac{\partial f}{\partial S_i} \frac{\partial S_i}{\partial \ln(p_E)} = \sum_{i=1}^{4} \frac{\partial f}{\partial S_i} \Big[\sum_{j=1}^{4} \frac{\partial S_i}{\partial P_j} \frac{\partial P_j}{\partial \ln(p_E)} \Big] \\ &= \sum_{i=1}^{4} \frac{\partial f}{\partial S_i} \Big[\sum_{j=1}^{4} \frac{\partial S_i}{\partial P_j} \frac{\partial \ln(ps_j)}{\partial \ln(p_E)} \Big] = \sum_{i=1}^{4} \frac{\partial f}{\partial S_i} \Big[\sum_{j=1}^{4} \frac{\partial S_i}{\partial P_j} \Big] = \sum_{i=1}^{4} \frac{\partial f}{\partial S_i} \Big[\sum_{j=1}^{4} \frac{\partial Z_i}{\partial P_j} \Big] \end{aligned}$$

We present these billing-cycle level price elasticities for each service, overall price elasticity for electricity consumption, and the income elasticity for the service demand model using constrained kernel regression for the electricity consumption function in Figure 12, 13 and 14. We present the same results for the service demand model when using parametric model with household and weather heterogeneity in Figure 16, 17 and 18. Note that the own-price elasticities for the two versions of the model are quite close, except for business-end use.

The median and mean overall price elasticity for both versions of our energy service demand model is inelastic. The median and mean income elasticities imply that electricity is a normal (positive elasticity) good and a luxury (elasticity greater than one) good. The income elasticities of the kernel model show a two-peak pattern. The peak on the left is mainly urban households while the peak on the right is mainly rural households. This can relate to the outage factor distribution since we model it differently for urban and rural areas. In our customer and weather heterogeneity electricity consumption model does not have the two peaks. One potential reason is that the outage factor is less important in this electricity consumption function because it does not change the derivative of the electricity consumption function, while it does change the value of the derivative in the shape-constrained kernel model.

6.2 Willingness to pay

A major payoff from modeling the demand for electricity services is the ability to calculate a household's willingness to pay (WTP) for an additional hour of each electricity service at their current level of electricity consumption. We utilize the outside good to estimate a rupee per hour marginal willingness to pay for each electricity service at the customer's current demand for these services. The marginal willingness to pay in rupees for an additional hour of electricity service i during billing cycle $t$is:

$$\begin{aligned} WTP_{it} &= \frac{\partial U}{\partial s_{it}}/(\frac{\partial U}{\partial x_{it}}/p_x) \\ &= \frac{\partial U}{\partial S_{it}}\frac{\partial S_{it}}{\partial s_{it}}/(\frac{\partial U}{\partial X_{it}}\frac{\partial X_{it}}{\partial x_{it}}/p_x) \\ &= \frac{\partial U}{\partial Z_{it}}\frac{1}{s_{it}}/(\frac{\partial U}{\partial Z_{5t}}\frac{1}{x_{it}p_x}). \end{aligned} \tag{23}$$

Distributions of the marginal willingness-to-pay across billing cycles and households for each energy service for the two versions of energy services demand model are plotted in Figures 15 and 19.

A number of conclusions emerge from these figures. First, the mean and median marginal willingness to pay for an additional hour of appliance and business energy services are significantly higher than those for lighting and heat and cooling services for both models. Second, there is considerable heterogeneity in marginal willingness to pay for given energy service across customers and even within billing cycles for the same customer.

Tables 10 and 11 illustrate an additional feature of these marginal willingness to pay distributions. For each customer we computed the sample mean across billing cycles of WTP_{it}. We then computed the fraction of customers that had the highest billing cycle-level mean for energy service i. For both models, business services had the highest mean marginal willingness to pay for almost 50 percent of the customers. However, each of the other three energy services had non-trivial fractions of customers with the highest mean marginal willingness to pay. For the shape-constrained kernel regression model, approximately 34 percent had the highest mean marginal willingness to pay for heating and cooling services. For the customer demographic and weather heterogeneity electricity consumption function model, 23 percent of customers had the highest mean marginal willingness to pay for lighting. The remaining of Tables 10 and 11 given the percentages of energy services with the second highest, third highest, and fourth highest mean marginal willingness to pay.

A surprising result from these tables is that for both models the appliance energy service has, by far, the largest fraction of customers with the fourth highest mean marginal willingness to pay. However, for both models it has the, by far, the highest mean marginal willingness to pay. This results suggests an important channel for increasing the willingness of households to pay for a reliable supply of electricity. Consistent with the results in McRae (2015), households with significant appliance holdings are likely to value a reliable supply of electricity to be willing to pay for it.

Tables 12 and 13 explore the second moment properties the distribution of the billing cycle-level means of customer-level marginal willingness to pay for each energy service for the shape-constrained kernel regression model. Table 12 presents the across-customer covariance matrix of mean marginal willingness to pay for the four energy services. Consistent the results in Table 10, the variance of mean marginal willingness to pay for appliance services is by far the largest, followed by the mean marginal willing to pay for business services. Table 13 presents the correlation matrix associated with covariance matrix in Table 12. There is surprising little correlation across the four services in the marginal willingness to pay.

Tables 14 and 15 present the information in 12 and 13 for the customer and weather heterogeneity model. These two tables confirm the two conclusions from the results for the shape-constrained kernel regression model.

Tables 16 and 17 present the scatter plot of every billing cycle's service demand, marginal utility, and marginal willingness to pay for both models. To maintain a meaningful scale, we only include observations with service demand ≥ 0.1 for all four service categories. To have a better view of the 3D plots, we present each scatter plot from 2 angles. Note that the vertical axis is always marginal willingness to pay. Tables 18 and 19 presents the same plot but only include observations with service demand ≥ 1. These plots show that customers with high marginal willingness to pay usually have low level of service consumption currently. Thus policies that support household-level investments in electricity consuming capital goods should target population with low service consumption. They have higher marginal willingness to pay for services and are currently less likely to have a formal grid connection. Once they have the capital goods, they have a much higher chance to obtain formal connection and pay for more stable electricity service.

Results of our household-level marginal willingness pay computation provide strong empirical evidence in favor of the view that if the customer owns electricity consuming capital goods that yield high value energy services, customers are more likely to be willing to pay for a reliable supply of electricity.

7 Conclusion

Our analysis of the customer-level demand for energy services reveals substantial within and across customer heterogeneity in the marginal willingness to pay for different energy services. Although basic lightening and heat and cooling services are often cited as reasons for increasing electrification in developing countries, our empirical results suggest that modern household appliances and business uses of electricity are the major drivers of a high willingness to pay for electricity.

Consequently, policies which support household-level investments in electricity consuming capital goods that provide appliance services and business services combined with policies to increase the reliability of the supply of electricity for those households could significantly increase the fraction of households that pay for the electricity they consume.

References

Burgess, Robin, Michael Greenstone, Nicholas Ryan, and Anant Sudarshan, "The consequences of treating electricity as a right," *Journal of Economic Perspectives*, 2020, *34* (1), 145–69.

Depuru, Soma Shekara Sreenadh Reddy, Lingfeng Wang, Vijay Devabhaktuni, and Nikhil Gudi, "Measures and setbacks for controlling electricity theft," in "North American Power Symposium 2010" IEEE 2010, pp. 1–8.

Du, Pang, Christopher F. Parmeter, and Jeffrey S. Racine, "Nonparametric kernel regression with multiple predictors and multiple shape constraints," *Statistica Sinica*, 2013, pp. 1347–1371.

Hanemann, W. Michael., "Discrete/continuous models of consumer demand," *Econometrica: Journal of the Econometric Society*, 1984, pp. 541–561.

Hansen, Lars Peter, "Large sample properties of generalized method of moments estimators," *Econometrica: Journal of the econometric society*, 1982, pp. 1029–1054.

McRae, Shaun, "Infrastructure quality and the subsidy trap," *American Economic Review*, 2015, *105* (1), 35–66.

Pai, D.S., M. Rajeevan, O.P. Sreejith, B. Mukhopadhyay, and N.S. Satbha, "Development of a new high spatial resolution (0.25× 0.25) long period (1901-2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region," *Mausam*, 2014, *65* (1), 1–18.

Prayas, A. C., "Residential electricity consumption in India: what do we know," *Prayas (Energy Group)*, 2016.

Reiss, Peter C. and Matthew W. White, "Household electricity demand, revisited," *The Review of Economic Studies*, 2005, *72* (3), 853–883.

Srivastava, A. K., M. Rajeevan, and S. R. Kshirsagar, "Development of a high resolution daily gridded temperature data set (1969–2005) for the Indian region," *Atmospheric Science Letters*, 2009, *10* (4), 249–254.

Wang, Zeyu and Frank A. Wolak, "Designing nonlinear price schedules for urban water utilities to balance revenue and conservation goals," *Stanford University, Program on Energy and Sustainable Development, Working Paper*, 2022.

Yurtseven, Çağlar, "The causes of electricity theft: An econometric analysis of the case of Turkey," *Utilities Policy*, 2015, *37*, 70–78.

A Appendix for Category of Energy Services

Category of energy services	Appliances	Standard Appliance Wattage
Heating or Cooling	Air cooler	200 W
	Air conditioner	2000 W
	Room heater	2000 W
	Warm air blower	2000 W
	Ceiling fan	80 W
	Table fan	80 W
	Immersion rod	1000 W
	Geyser	2000 W
Lighting	Cfl/leds	20 W
	Bulbs	100 W
	Tube lights	40 W
Domestic end-use appliances	Television	200 W
	Refrigerator	60 W
	Water purifier	60 W
	Microwave	800 W
	Electric iron	1000 W
	Sewing machine	100 W
	Water pump	740 W
	Washing machine	700 W
	Others: Flour grinder, juicer, milk churner, mixer	200 W
Business end-use appliances	Cell phone charging	6 W
	Desktop computer	200 W
	Laptop computer	65 W

Notes: Standard wattage information of common household appliances from Bureau of Energy Efficiency standards for 2012-13 and online load calculators provided by Tamil Nadu Generation and Distribution Corporation (`https://www.tangedco.gov.in/load_calculato.html`) and Paschim Gujarat Vij Company Limited (`http://www.pgvcl.com/consumer/CONSUMER/calculate_n.php`)

B Data Appendix

(1) An example illustrating the backing out applicable energy price of electricity based on observed consumption

Consider for example, a billing cycle for a customer starting on 1st June to 1st August 2015, with a total bi-monthly consumption of 310 kWhs and total energy cost of Rs. 1,497. Given 310 kWhs consumed over two months, we expect this consumer to fall in the fourth consumption tier (corresponding to 150 to 300 kWhs of monthly consumption tier in the tariff schedule). Based on the prevailing electricity prices at the time, the total energy cost for the customer for this cycle is calculated as: Rs. 3.50 per kWh for first the 100 kWhs of consumption (i.e., the first tier of the tariff schedule) + Rs. 5.45 per kWh for the next 200 kWhs of consumption (i.e., the second tier) + Rs. 5.7 for the remaining 10 kWhs of consumption (i.e., the third tier). The total energy cost based on this calculation is the same as that appears in the dataset (i.e., Rs. 1,497). For such bills, we attribute the Rs. 5.7 per unit figure (i.e., the energy charge of the consumption tier on which the household lies on) as the marginal price for this customer-billing cycle. Additional duties, surcharges and cess, levied on per-unit of consumption are then added to this marginal price to arrive at the gross final marginal price faced by the customer in this cycle.

Comparing the per-kWh energy cost derived from the billing dataset to per-unit prices prescribed in tariff schedule is straightforward, except for the months during which tariff revisions occur. If new tariff schedules are issued mid-cycle, JVVNL prorates the total consumption by the number of days of the billing cycle that falls under each of the tariff schedules. For instance, tariff revisions occurred on 1st September 2016. Consider a billing cycle, starting 1st August to 1st October, 2016. To calculate the energy charges, first, the total consumption over the billing cycle is divided into two, weighted by the fraction of days in the cycle that falls under each tariff schedule (1st Aug – 1st Sept = 31 days and 1st Sept – 1st October = 30 days). The total energy charge is then calculated by applying the energy charges prescribed under each tariff schedule and on basis of the prorated consumption. The marginal price of consumption for each half of consumption is then calculated using the same formulas as noted in the text above. In such cases of mid-cycle tariff revisions, we have split the cycle into two, starting 1st Aug to 1st Sept and 1st Sept to 1st October, prorating consumption, fixed costs, total electricity prices, etc. for each of the two cycles.

(2) Additional validation checks conducted on the administrative billing data

We validate if the household consumption, in general, reacts inversely to changes in prices. Figure 5 illustrates the changes in the density of consumption and energy prices for periods between the first and second price revisions (September-2016 to March-2017 and February-2015 to September-2016 respectively). The figure indicates a sharper increase in prices for consumers at top-most tiers. This higher increase in prices also appears to be correlated with a leftward movement along the consumption distribution. Between September 2016 and March 2017, the fraction of bills in the 0-50 kWhs above poverty line (APL) category increased by more than 10 percentage points while price increased by Rs. 0.4. Price increase for greater than 500 kWhs of consumption category was about double that amount and was associated with fall of about 3 percentage points in the fraction of bills. More generally, the share of household-bills in the top three consumption tiers (for which prices have risen the most) has fallen, while it has increased in the bottom two tiers (for which prices rose moderately).

The shift in the density of consumption could also be due to selection, wherein, large numbers of newly connected households with low initial levels of consumption could have added mass to the left of the distribution during September 2016 to March 2017. However, we do not find evidence of selection in the data–only 1 percent of sample comprises of consumers-bills that were newly added to the dataset during this period, insignificant enough to have moved the distribution so sharply to the left. To be sure, we exclude these households from the dataset to find the leftward shift in consumption to persist. We interpret this result to be the first indication that consumers in our sample show a negative response to rising prices.

The consumption data for a given household in our sample also appears to be stable over consecutive billing periods. Figure 6 compares the consumption tier of a household in the previous cycle (in the horizontal axes) to the consumption tier in the current cycle (in boxes). Households below the poverty line have low demand for energy services and therefore may not transition to higher consumption tiers. The opposite is true for some consumers in the higher tiers. A large proportion of both groups therefore are observed to reside within their own tier over consecutive billing cycles. For others, transitions to one-tier above or below their current consumption tier is more likely. Transitions to more than two tiers away over consecutive cycles, reassuringly, appears to be rare in the data.

Tables

Table 1: The distribution of various charges in the dataset

	Mean	10th percentile	90th percentile
monthly consumption	166.3 kWh	24.7 kWh	374.5 kWh
monthly consumption per capita	39.7 kWh	4.4 kWh	87.6 kWh
energy charges	₹1729	₹181	₹4086
fixed charges	₹348	₹180	₹480
electricity duty	₹138	₹20	₹310
tariff subsidy	₹-15.1	₹-66.3	₹0

Notes: Total number of observations in the sample are 7,615. The distribution was weighted by sampling probabilities.

Table 2: Comparing the consumption distribution of sampled households to billing dataset

Percentiles	Admin billing data (all HH)	Admin billing data (surveyed HH)
1%	1	1
5%	15	14
10%	30	28
25%	62	58
50%	115	119
75%	240	263
90%	452	436
95%	645	562
99%	1283	762
Mean Value	208	189

Notes: This table compares the distribution of consumption for all households in the two district of Rajasthan in the administrative data to the consumption distribution obtained from the survey using sampling probabilities. The sample period for both datasets is restricted to January and February 2017.

Table 3: Socioeconomic profile of the households

Variable	Proportion of households
Proportion of general caste	29%
Pukka Wall	97.1%
Pukka Roof	99%
Pukka Floor	44.4%
Tap water connection	63.4%
Top Income Source: Farming	15.8%
Top Income Source: Livestock	3.4%
Top Income Source: Own Business	19.5%
Top Income Source: Casual Labour	23.2%
Top Income Source: Salaried Work	36.2%
Top Income Source: Remittances	2.8%
Living in a tented house	5.5%
Ownership of Below Poverty Line Card	18.2%

Table 4: Parameter Estimate for Outage Factor Distribution

Parameter	Rural	S.E. for Rural	Urban	S.E. for Urban
$\beta_p^{\text{constant}}$	-1.144	0.083	-3.285	0.124
$\beta_p^{\text{temperature}}$			0.028	0.043
$\beta_p^{\text{precipitation}}$			0.077	0.106
$\beta_p^{\text{electricity consumption}}$			-0.093	0.046
$\beta_p^{\text{household size}}$			0.019	0.034
β_p^{age}			0.006	0.010
$\beta_p^{\text{number of rooms}}$			0.164	0.090
$\beta_p^{\text{year of schooling}}$			0.014	0.020
$\beta_{a_1}^{\text{constant}}$	2.449	0.067	1.559	0.371
$\beta_{a_1}^{\text{temperature}}$	0.085	0.039	0.416	0.024
$\beta_{a_1}^{\text{precipitation}}$	-0.836	0.625	4.647	0.160
$\beta_{a_1}^{\text{electricity consumption}}$	0.007	0.070	0.526	0.132
$\beta_{a_1}^{\text{household size}}$			0.052	0.079
$\beta_{a_1}^{\text{age}}$			0.041	0.013
$\beta_{a_1}^{\text{number of rooms}}$			0.030	0.103
$\beta_{a_1}^{\text{year of schooling}}$			0.035	0.038
$\beta_{b_1}^{\text{constant}}$	1.084	0.046	0.327	0.279
$\beta_{b_1}^{\text{temperature}}$	0.070	0.051	0.247	0.017
$\beta_{b_1}^{\text{precipitation}}$	-0.706	0.653	4.221	0.154
$\beta_{b_1}^{\text{electricity consumption}}$	-0.018	0.053	0.105	0.138
$\beta_{b_1}^{\text{household size}}$			0.070	0.075
$\beta_{b_1}^{\text{age}}$			0.042	0.013
$\beta_{b_1}^{\text{number of rooms}}$			0.058	0.100
$\beta_{b_1}^{\text{year of schooling}}$			0.068	0.032
$\beta_{a_2}^{\text{constant}}$	4.757	0.068	3.809	0.069
$\beta_{a_2}^{\text{temperature}}$	-0.035	0.020	-0.057	0.019
$\beta_{a_2}^{\text{precipitation}}$	-0.156	0.307	0.161	0.137
$\beta_{a_2}^{\text{electricity consumption}}$	-0.010	0.045	-0.013	0.013
$\beta_{a_2}^{\text{household size}}$			0.015	0.024
$\beta_{a_2}^{\text{age}}$			0.020	0.004
$\beta_{a_2}^{\text{number of rooms}}$			-0.040	0.058
$\beta_{a_2}^{\text{year of schooling}}$			0.023	0.009
$\beta_{b_2}^{\text{constant}}$	5.630	0.069	0.130	0.065
$\beta_{b_2}^{\text{temperature}}$	-0.032	0.019	-0.044	0.016
$\beta_{b_2}^{\text{precipitation}}$	-0.186	0.300	0.054	0.057
$\beta_{b_2}^{\text{electricity consumption}}$	-0.012	0.046	-0.003	0.011
$\beta_{b_2}^{\text{household size}}$			-0.007	0.022
$\beta_{b_2}^{\text{age}}$			0.009	0.004
$\beta_{b_2}^{\text{number of rooms}}$			-0.005	0.041
$\beta_{b_2}^{\text{year of schooling}}$			0.026	0.007

Table 5: Parameter Estimate for electricity consumption function Using Specified Functional Form

Parameter	Estimate	Standard Error
$\gamma^{\text{constant}}_{\text{lighting}}$	1.457	0.075
$\gamma^{\text{income}}_{\text{lighting}}$	0.239	0.030
$\gamma^{\text{household size}}_{\text{lighting}}$	-0.093	0.014
$\gamma^{\text{temperature}}_{\text{lighting}}$	0.003	0.003
$\gamma^{\text{urban}}_{\text{lighting}}$	0.120	0.016
$\gamma^{\text{constant}}_{\text{domestic}}$	-1.202	0.119
$\gamma^{\text{income}}_{\text{domestic}}$	-2.344	0.207
$\gamma^{\text{household size}}_{\text{domestic}}$	-2.511	0.231
$\gamma^{\text{temperature}}_{\text{domestic}}$	-1.585	0.175
$\gamma^{\text{urban}}_{\text{domestic}}$	-0.774	0.090
$\gamma^{\text{constant}}_{\text{business}}$	-0.798	0.115
$\gamma^{\text{income}}_{\text{business}}$	-5.322	0.342
$\gamma^{\text{household size}}_{\text{business}}$	-0.715	0.143
$\gamma^{\text{temperature}}_{\text{business}}$	-2.427	0.280
$\gamma^{\text{urban}}_{\text{business}}$	-4.375	0.432

Table 6: Coefficients for the demand model, with constrained kernel regression for the electricity consumption function

	Dependent variables				
	hc	l	a	b	x
Intercept	-3.8608	-3.3067	-7.8208	-4.4773	0
	(0.1514)	(0.1483)	(0.6653)	(0.1611)	(NA)
log(daily income)	-0.0596	0	0	-0.1029	0.8376
	(0.1873)	(NA)	(NA)	(0.2115)	(0.1936)
log(household size)	0	0	0	-0.1149	0
	(NA)	(NA)	(NA)	(0.1132)	(NA)
log(mean temp)	0	1.2537	0	1.1809	-0.8223
	(NA)	(0.4006)	(NA)	(0.4745)	(0.2859)
urban	0	0.0896	0	0	-1.3631
	(NA)	(0.8874)	(NA)	(NA)	(0.8584)
urban x log(daily income)	0.1989	0.1449	-0.0369	0.1501	0.2253
	(0.1148)	(0.1198)	(0.4245)	(0.1023)	(0.1571)

Notes: J Statistic = 50.0116, J test p-value = 0.1113, #Obs = 611, ObjValue = 0.08185

Table 7: Symmetric Gamma Matrix, with constrained kernel regression for the electricity consumption function

	hc	l	a	b	x
hc	-0.005452				
	(0.0015)				
l	0.002343	-0.003084			
	(0.0022)	(0.0044)			
a	-0.000101	0.000423	-0.000484		
	(2e-04)	(0.0012)	(9e-04)		
b	0.000641	0.001214	0.000103	-0.001563	
	(9e-04)	(0.0019)	(8e-04)	(0.0021)	
x	-0.000363	-0.000722	0.000163	0.000572	-0.000579
	(0.0028)	(0.0029)	(2e-04)	(0.0017)	(0.0148)

Table 8: Coefficients for the demand model, with specified model for the electricity consumption function

	Dependent variables				
	hc	l	a	b	x
Intercept	-4.7089	-3.2772	-5.5559	-5.6471	0
	(0.2754)	(0.5625)	(0.6427)	(0.4787)	(NA)
log(daily income)	0.489	1.1948	0.3917	-0.474	1.3351
	(0.7742)	(0.9696)	(1.7923)	(0.4711)	(0.5092)
log(household size)	0.2265	0.0624	-1.6921	0.1018	-0.0067
	(0.412)	(0.4457)	(0.5221)	(0.6482)	(0.3855)
log(mean temp)	2.209	2.3838	1.5948	0.1252	0.1882
	(0.7674)	(0.7622)	(1.0671)	(1.2622)	(0.5416)
urban	3.8203	6.6864	5.6927	3.5463	1.4348
	(4.6854)	(7.6309)	(11.326)	(2.0431)	(3.1821)
urban x log(daily income)	-0.3496	-0.7713	-0.8143	-0.6731	-0.1735
	(0.7578)	(0.9717)	(1.67)	(0.5282)	(0.5428)

Note: J Statistic = 36.9708, J test p-value = 0.1195, #Obs = 611, ObjValue = 0.06051

Table 9: Symmetric Gamma Matrix, with specified model for the electricity consumption function

	hc	l	a	b	x
hc	-0.000937				
	(8e-04)				
l	5.3e-05	-0.010303			
	(0.0019)	(0.0094)			
a	-2.9e-05	0.003391	-0.003847		
	(0.0011)	(0.0056)	(0.0048)		
b	0.000294	0.004079	0.001856	-0.005556	
	(0.0014)	(0.0085)	(0.0044)	(0.008)	
x	0.001168	0.001667	-0.000733	-0.000811	-0.001864
	(0.0012)	(0.0051)	(0.0018)	(0.0012)	(0.0204)

Table 10: Rank Willingness to Pay for All Services, constrained kernel regression

	Heat and Cool	Light	Appliances	Business
Percent of customers with highest mean willingness to pay across billing cycles	34.70%	8.84%	6.55%	49.92%
Percent of customers with 2nd highest mean willingness to pay across billing cycles	30.93%	35.52%	2.29%	31.26%
Percent of customers with 3rd highest mean willingness to pay across billing cycles	30.11%	48.61%	4.91%	16.37%
Percent of customers with 4th highest mean willingness to pay across billing cycles	4.26%	7.04%	86.25%	2.45%

Table 11: Rank Willingness to Pay for All Services, specified functional form

	Heat and Cool	Light	Appliances	Business
Percent of customers with highest mean willingness to pay across billing cycles	13.09%	23.08%	15.88%	47.95%
Percent of customers with 2nd highest mean willingness to pay across billing cycles	30.44%	37.15%	19.31%	13.09%
Percent of customers with 3rd highest mean willingness to pay across billing cycles	42.39%	27.33%	19.31%	10.97%
Percent of customers with 4th highest mean willingness to pay across billing cycles	14.08%	12.44%	45.50%	27.99%

Table 12: Covariance Matrix for Willingness to Pay for Different Services, constrained kernel regression

	Heat and Cool	Light	Appliances	Business
Heat and Cool	426.93	-0.85	-346.97	-5.09
Light	-0.85	3.33	-14.99	18.62
Appliances	-346.97	-14.99	1547285.03	-597.43
Business	-5.09	18.62	-597.43	3342.77

Table 13: Correlation Matrix for Willingness to Pay for Different Services, constrained kernel regression

	Heat and Cool	Light	Appliances	Business
Heat and Cool	1.000	-0.022	-0.013	-0.004
Light	-0.022	1.000	-0.007	0.177
Appliances	-0.013	-0.007	1.000	-0.008
Business	-0.004	0.177	-0.008	1.000

Table 14: Covariance Matrix for Willingness to Pay for Different Services, parametric functional form

	Heat and Cool	Light	Appliances	Business
Heat and Cool	28.24	0.80	-339.69	-5.94
Light	0.80	6.35	-1132.77	-37.43
Appliances	-339.69	-1132.77	61754780.35	-26325.58
Business	-5.94	-37.43	-26325.58	29672.59

Table 15: Correlation Matrix for Willingness to Pay for Different Services, parametric functional form

	Heat and Cool	Light	Appliances	Business
Heat and Cool	1.000	0.060	-0.008	-0.006
Light	0.060	1.000	-0.057	-0.086
Appliances	-0.008	-0.057	1.000	-0.019
Business	-0.006	-0.086	-0.019	1.000

Table 16: 3D Scatter Plot of Marginal Utility, Service Demand, and Marginal Willingness to Pay, constrained kernel regression, filtered on service demand≥ 0.1 for all service categories

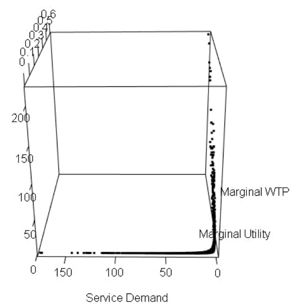
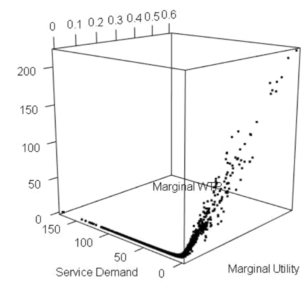
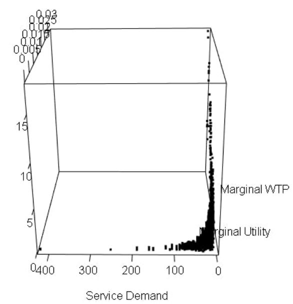
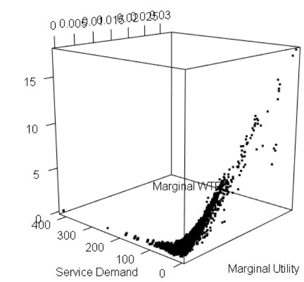
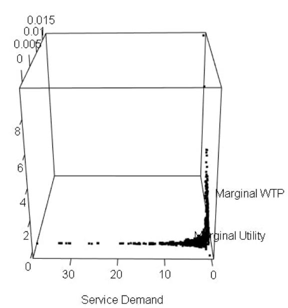
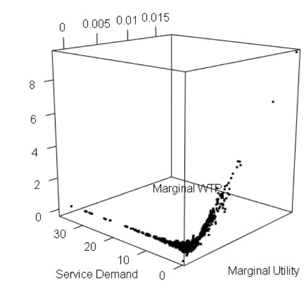
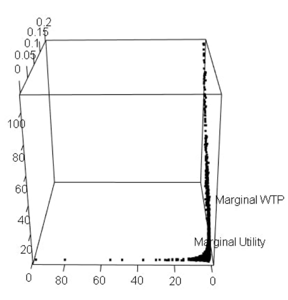
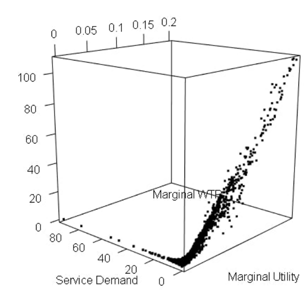
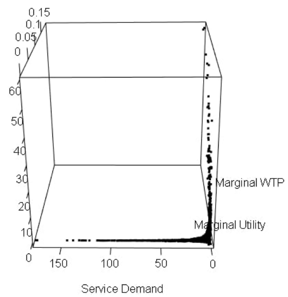
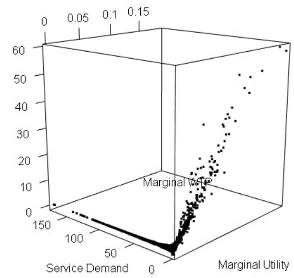
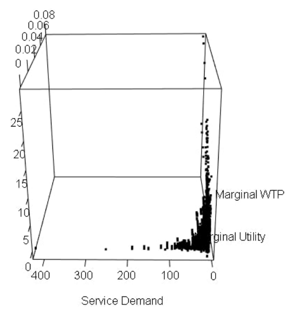
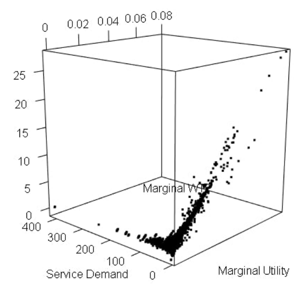
<table>
<tr><th>Service</th><th>View Angle 1</th><th>View Angle 2</th></tr>
<tr><td>Heating & Cooling</td><td>
</td><td>
</td></tr>
<tr><td>Lighting</td><td>
</td><td>
</td></tr>
<tr><td>Domestic Use</td><td>
</td><td>
</td></tr>
<tr><td>Business Use</td><td>
</td><td>
</td></tr>
</table>

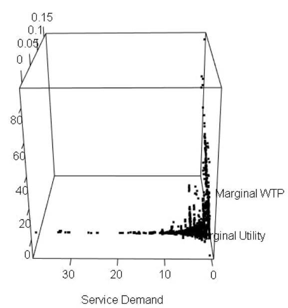
Table 17: 3D Scatter Plot of Marginal Utility, Service Demand, and Marginal Willingness to Pay, parametric functional form, filtered on service demand≥ 0.1 for all service categories

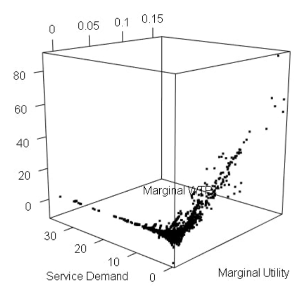
<table>
<tr><th>Service</th><th>View Angle 1</th><th>View Angle 2</th></tr>
<tr><td>Heating & Cooling</td><td>


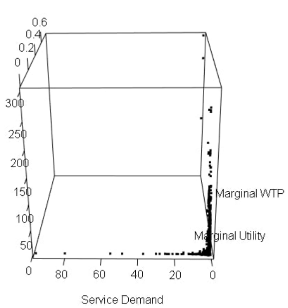
</td><td>


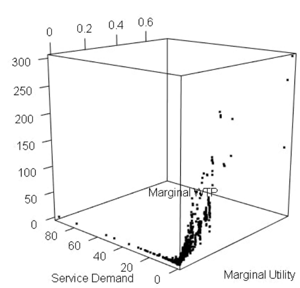
</td></tr>
<tr><td>Lighting</td><td>


</td><td>


</td></tr>
<tr><td>Domestic Use</td><td>


</td><td>


</td></tr>
<tr><td>Business Use</td><td>


</td><td>


</td></tr>
</table>

Table 18: 3D Scatter Plot of Marginal Utility, Service Demand, and Marginal Willingness to Pay, constrained kernel regression, filtered on service demand≥ 1 for all service categories

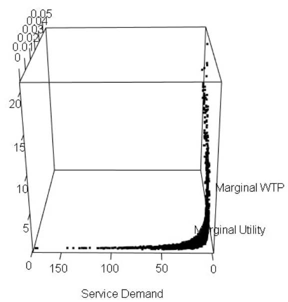
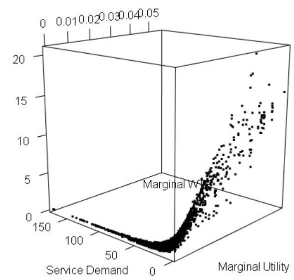
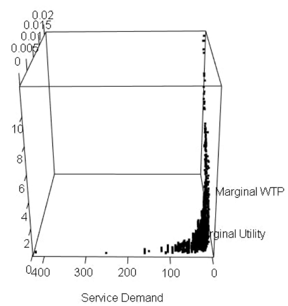
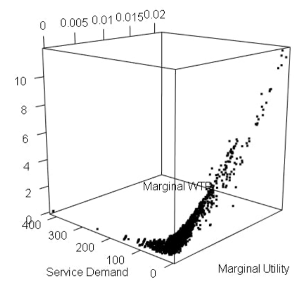
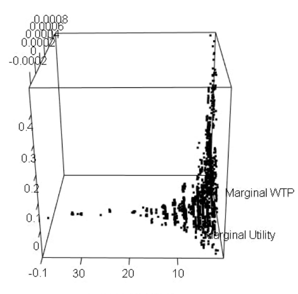
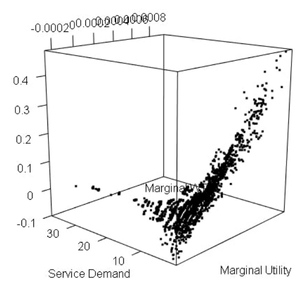
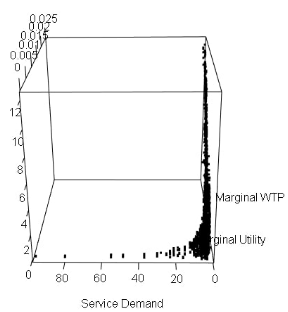
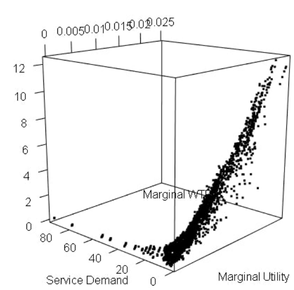
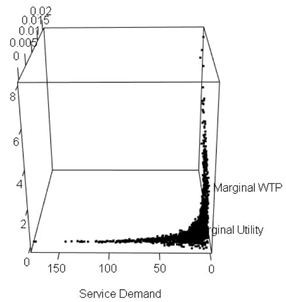
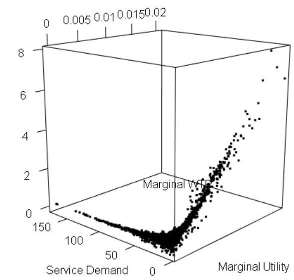
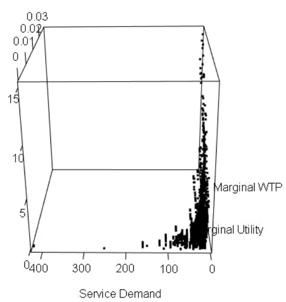
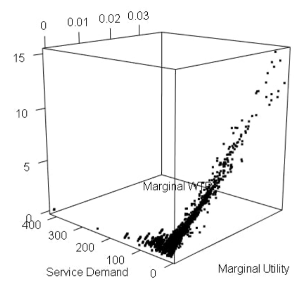
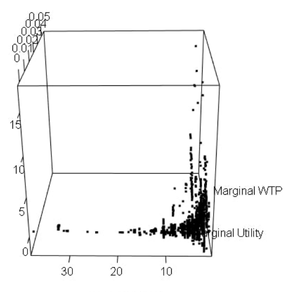
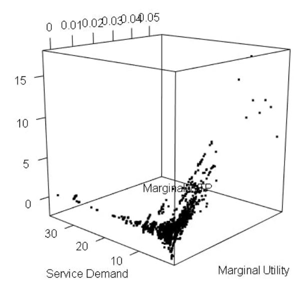
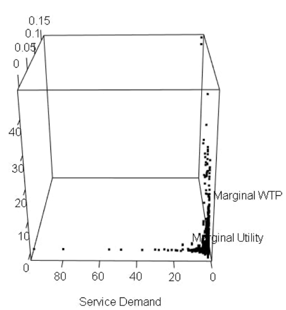
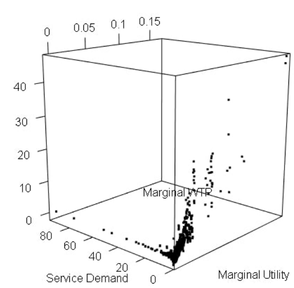
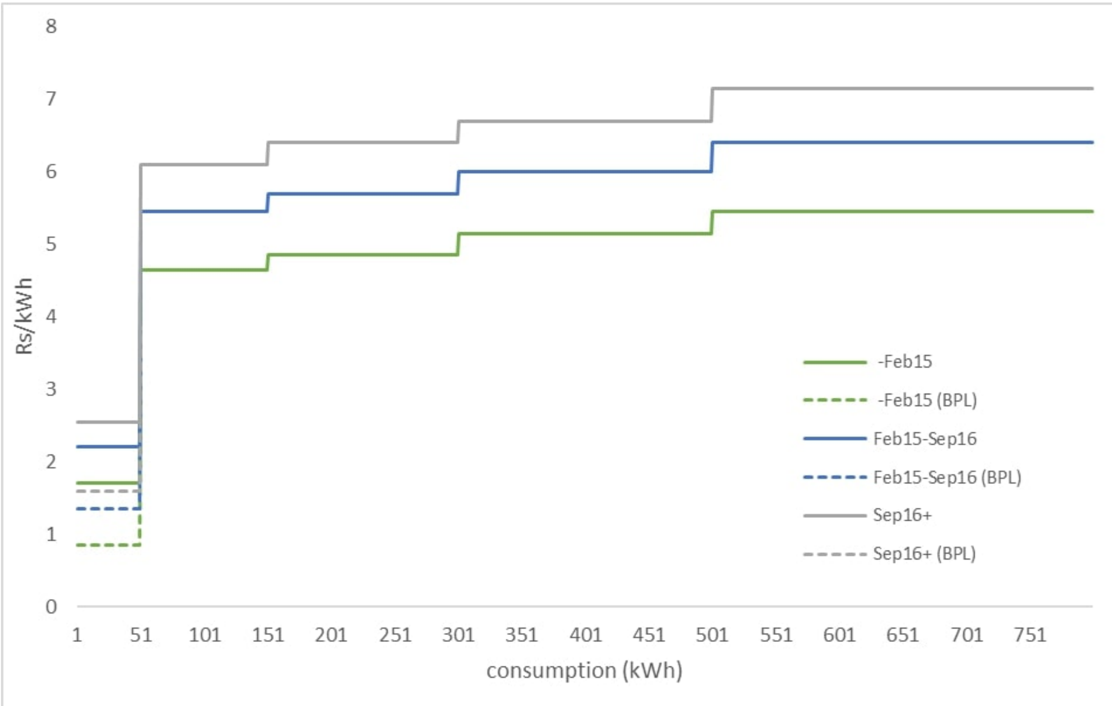
<table>
<tr><th>Service</th><th>View Angle 1</th><th>View Angle 2</th></tr>
<tr><td>Heating & Cooling</td><td>
</td><td>
</td></tr>
<tr><td>Lighting</td><td>
</td><td>
</td></tr>
<tr><td>Domestic Use</td><td>
</td><td>
</td></tr>
<tr><td>Business Use</td><td>
</td><td>
</td></tr>
</table>

Table 19: 3D Scatter Plot of Marginal Utility, Service Demand, and Marginal Willingness to Pay, parametric functional form, filtered on service demand≥ 1 for all service categories

Service	View Angle 1	View Angle 2
Heating & Cooling		
Lighting		
Domestic Use		
Business Use		

Figures

Figure 1: Energy Charges applicable on residential consumers of Rajasthan over the same period



Notes: Energy and fixed prices obtained from various tariff orders issued by JVVNL

Figure 2: Consumption distributions from NSS survey and administrative billing data

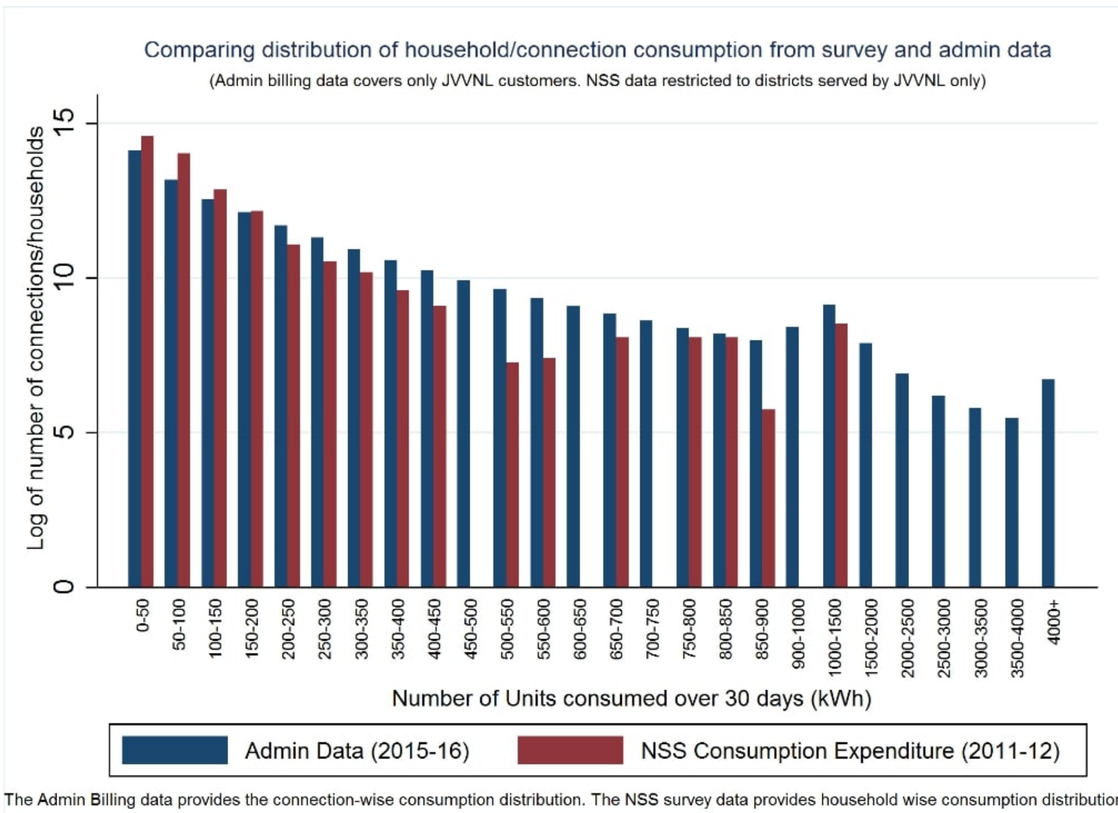
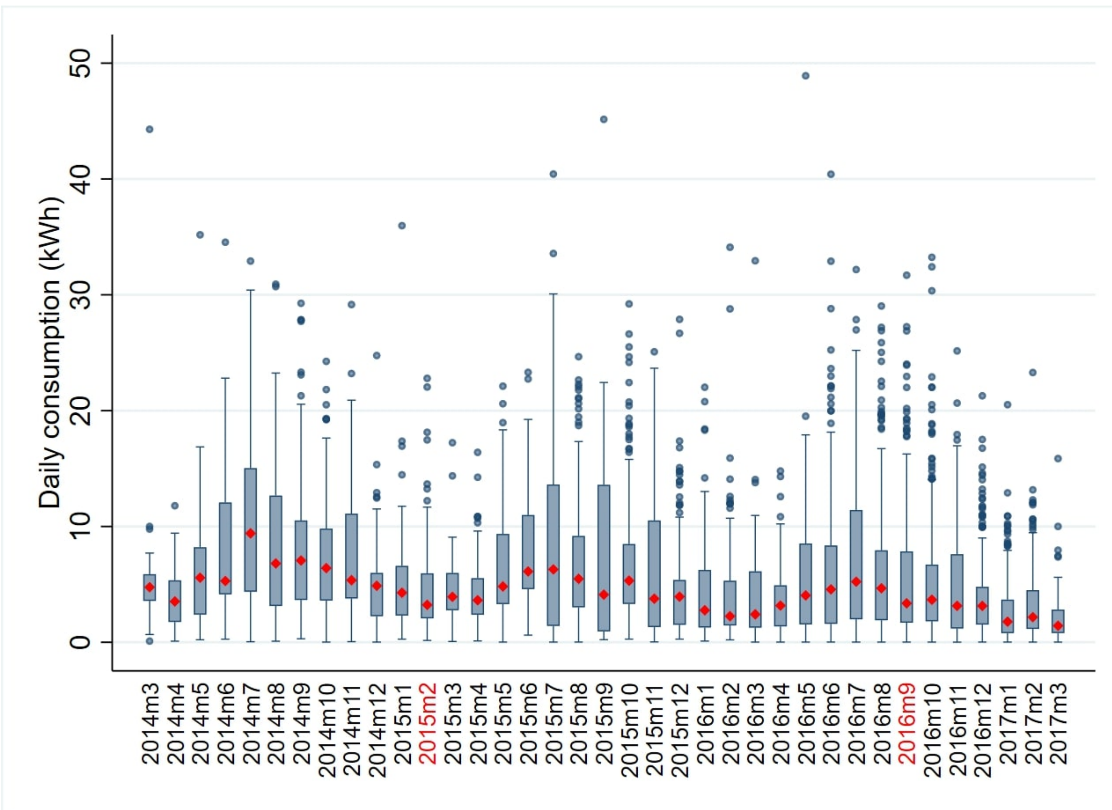


Figure 3: Daily consumption distribution by month – red labels indicate months in which price revisions occurred



Notes: Daily consumption is calculated as the total consumption over the billing cycle divided by the number of days in the billing cycle. The monthly distribution of daily consumption is weighted by the sampling probabilities. The labels on the horizontal axis denote the month of bill issuance. Red diamonds indicate the median daily consumption over the month and red monthly labels indicate the period at which a revision in tariff schedule occurred.

Figure 4: Histogram of service demand in terms of hours, for four categories of services

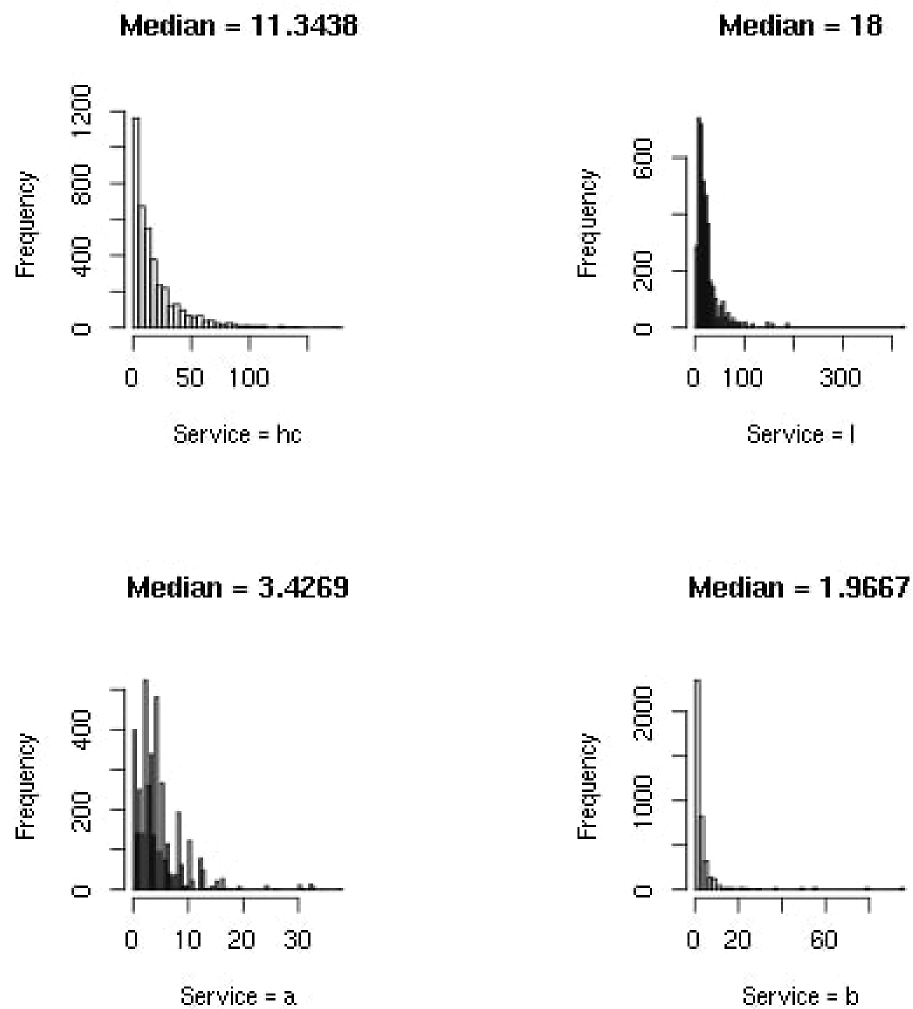
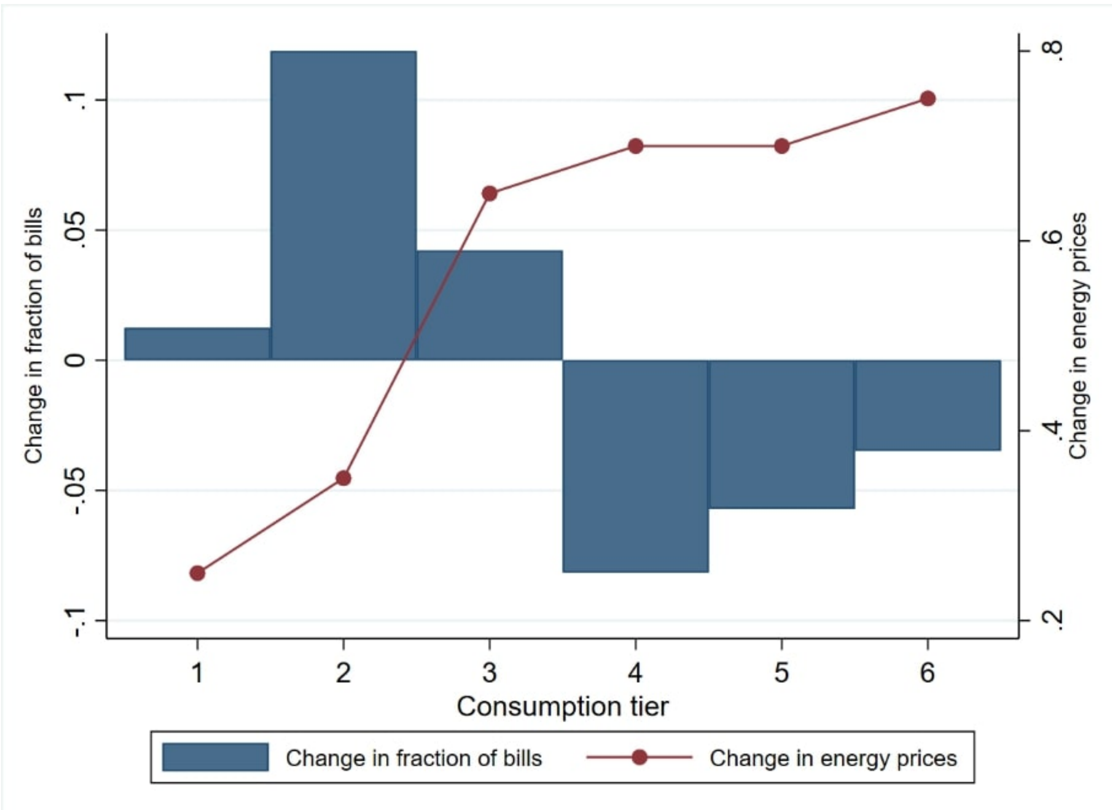
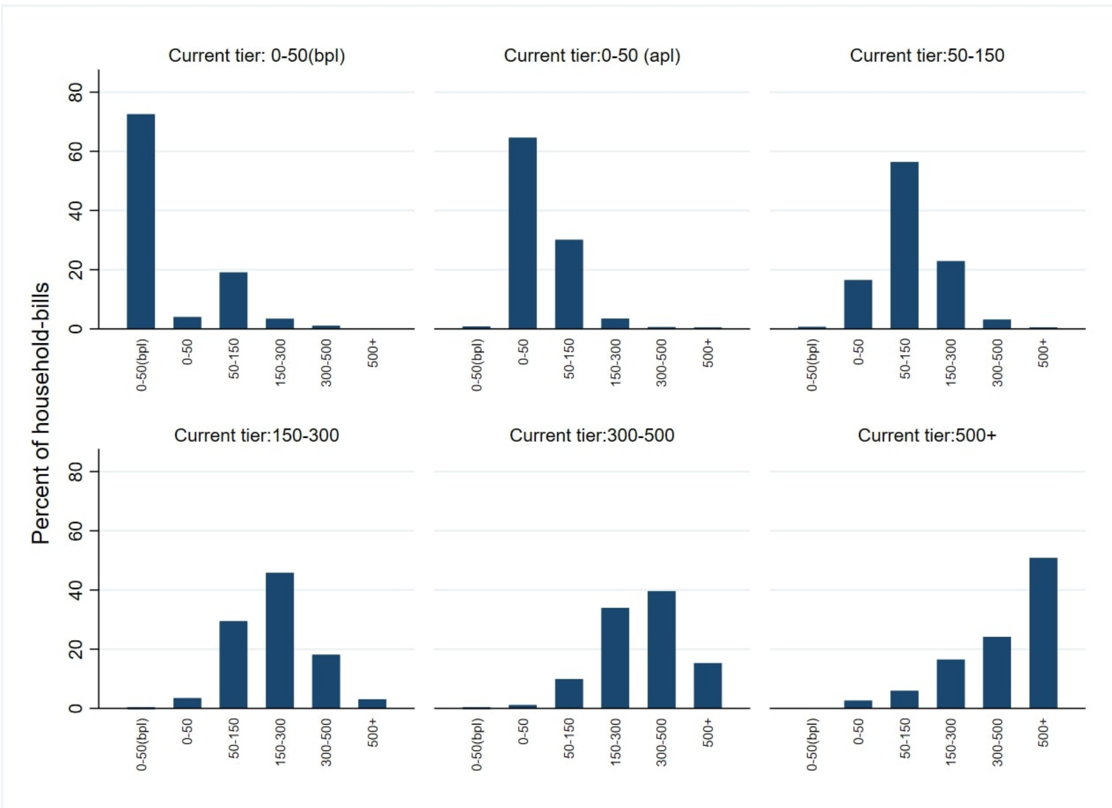


Figure 5: Changes in the fraction of household bills and energy prices by consumption tiers



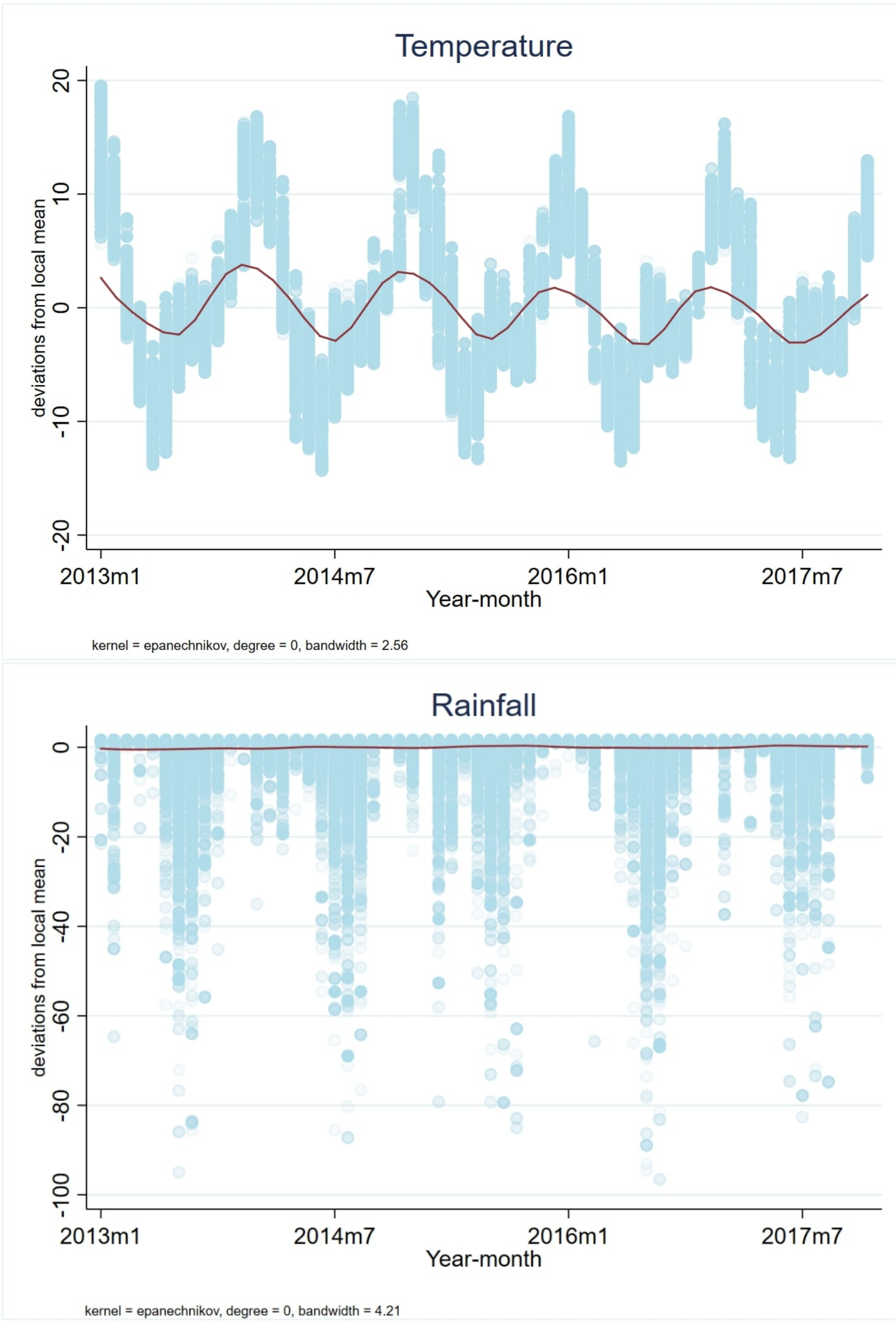
Notes: The graph indicates the correlation between changes in the fraction of bills by consumption tiers and the changes in energy prices. These changes are calculated over two periods: September-2016 to March-2017 and February-2015 to September 2016. The fraction of bills in a period is calculated as the number of bills in a consumption tier divided by the total number of bills across all tiers in that period. The price changes are calculated using the energy prices as prescribed in tariff schedules. The six consumption tiers indicated in the graph respectively correspond to 0-50 kWhs (BPL), 0-50 kWhs (APL), 51-150 kWhs, 151-300 kWhs, 301-500 kWhs, and greater than 500 kWhs.

Figure 6: Consumption transitions across tiers over consecutive billing cycle



Notes: The graph compares the probability of transitioning across consumption tiers over consecutive billing cycles. The consumption tiers from the previous period (t_0) are indicated in the horizontal axis. The percentage of household-bills is calculated as the total number of bills that were in consumption tier i in the previous period t_0 divided by all household-bills that are in consumption tier j in current period (t_1) (consumption tier j in period t_1 is denoted as "Current tier: X-Y" in the figure). The percentages are weighted by sampling weights.

Figure 7: Changes in temperature and precipitation from three year (2014-2017) average levels



Notes: The vertical axis shows the deviation of daily temperature and rainfall in a region from the local four-year period (2013-2017) means of temperature and rainfall

Figure 8: Histogram of household outage factor

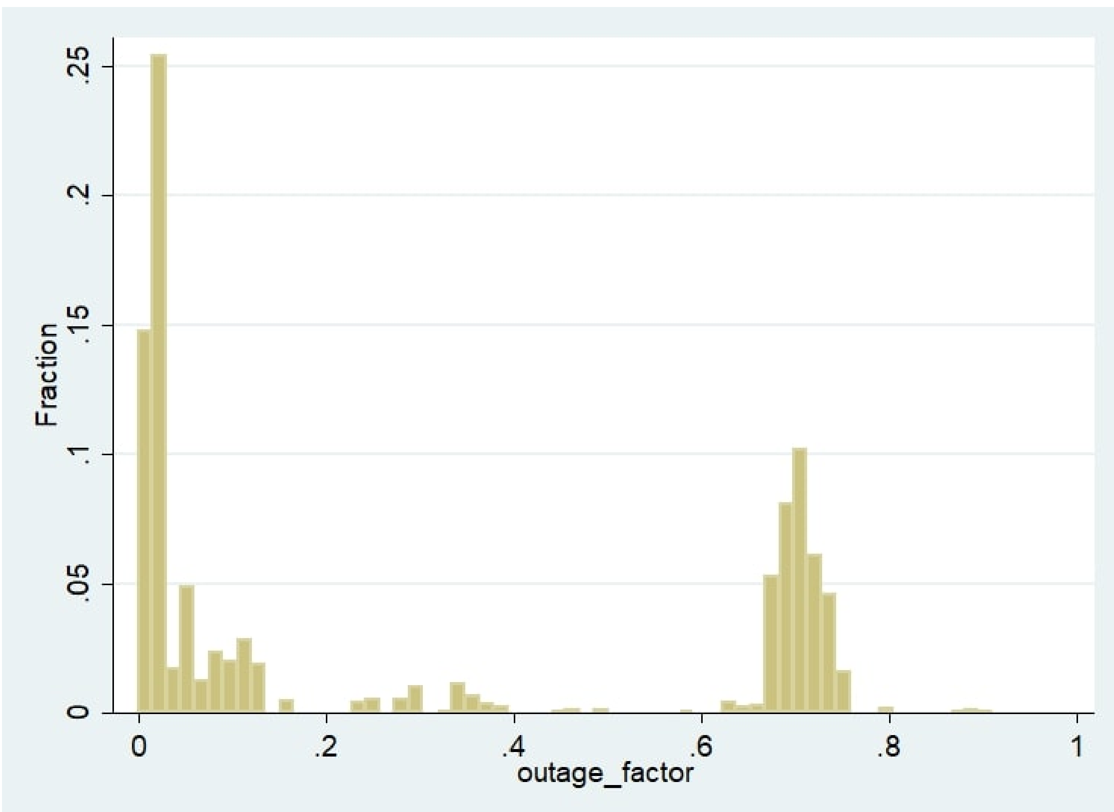


Figure 9: Histogram for outside good at log level

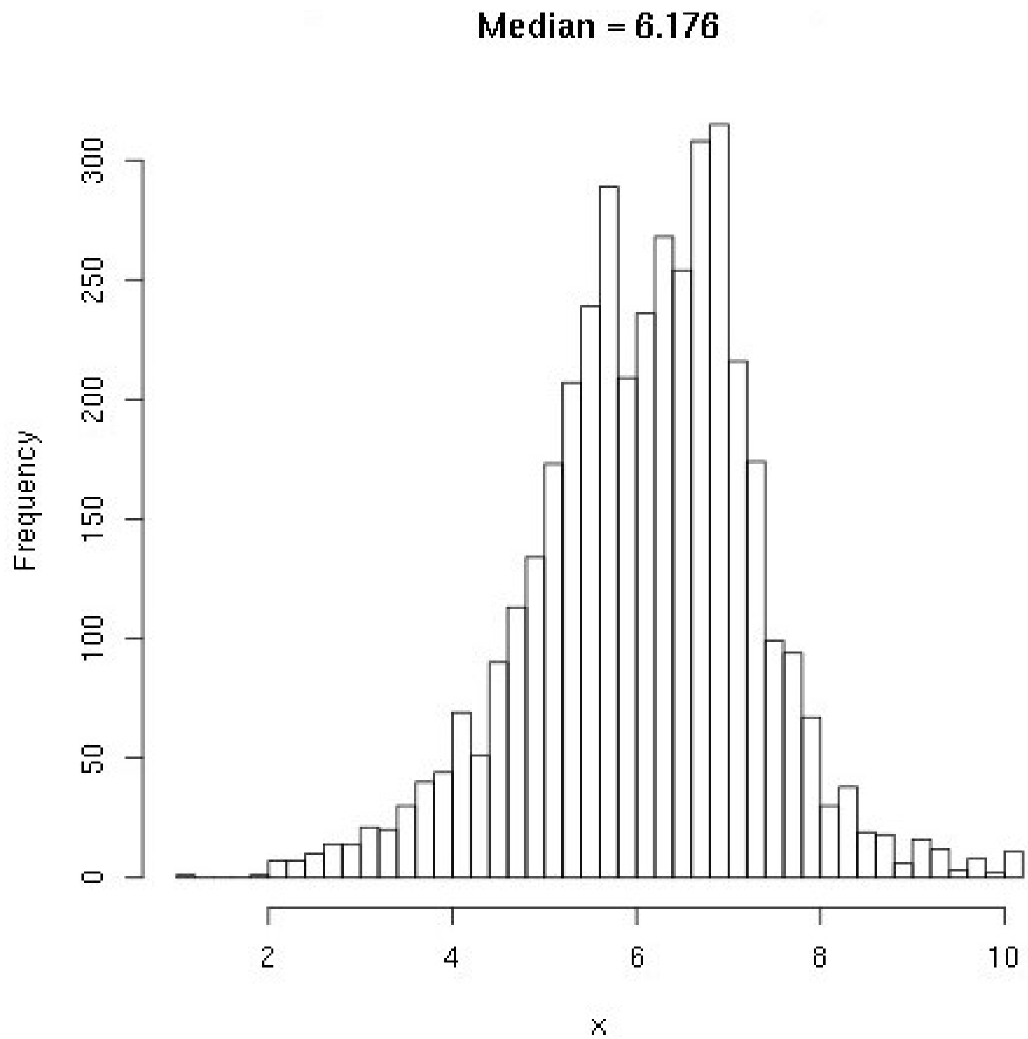


Figure 10: Histogram of gradients of electricity consumption function using constrained kernel regression

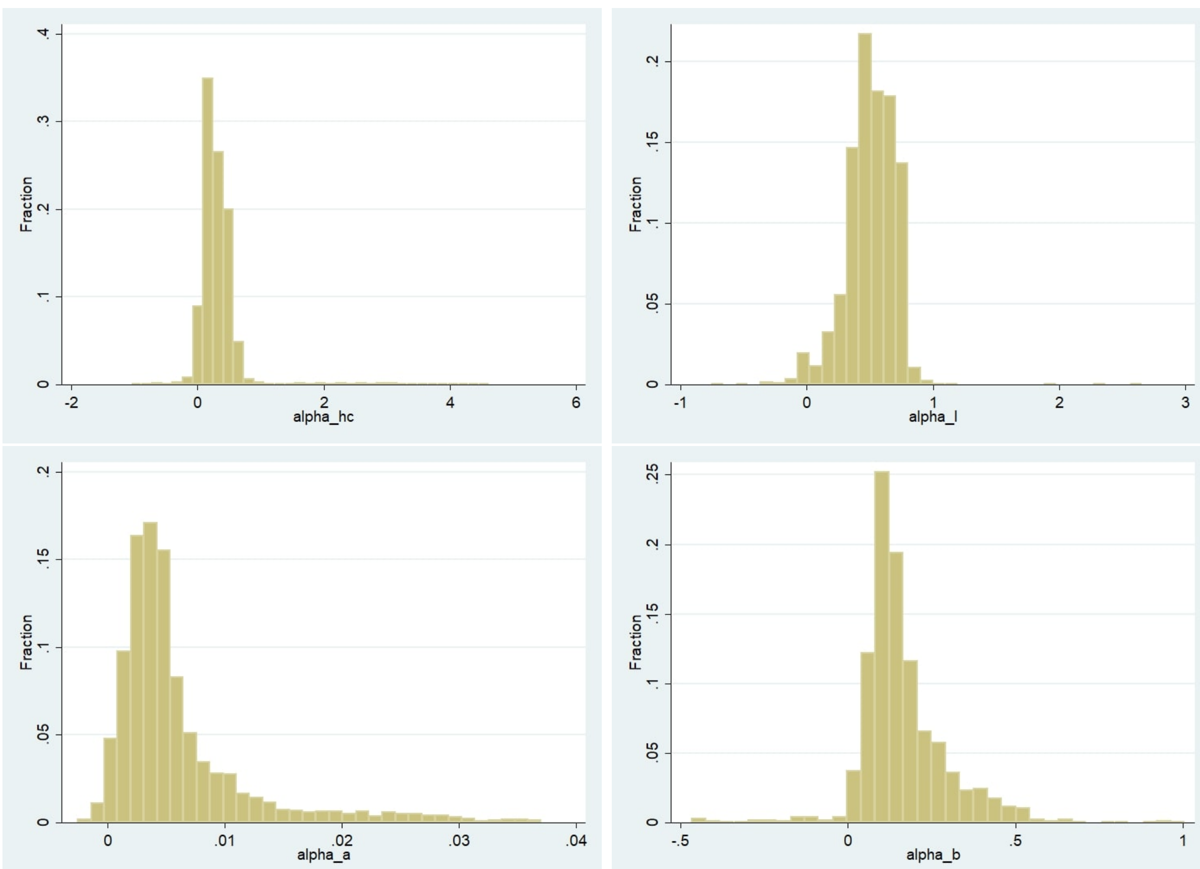


Figure 11: Histogram of gradients of electricity consumption function using specified functional form

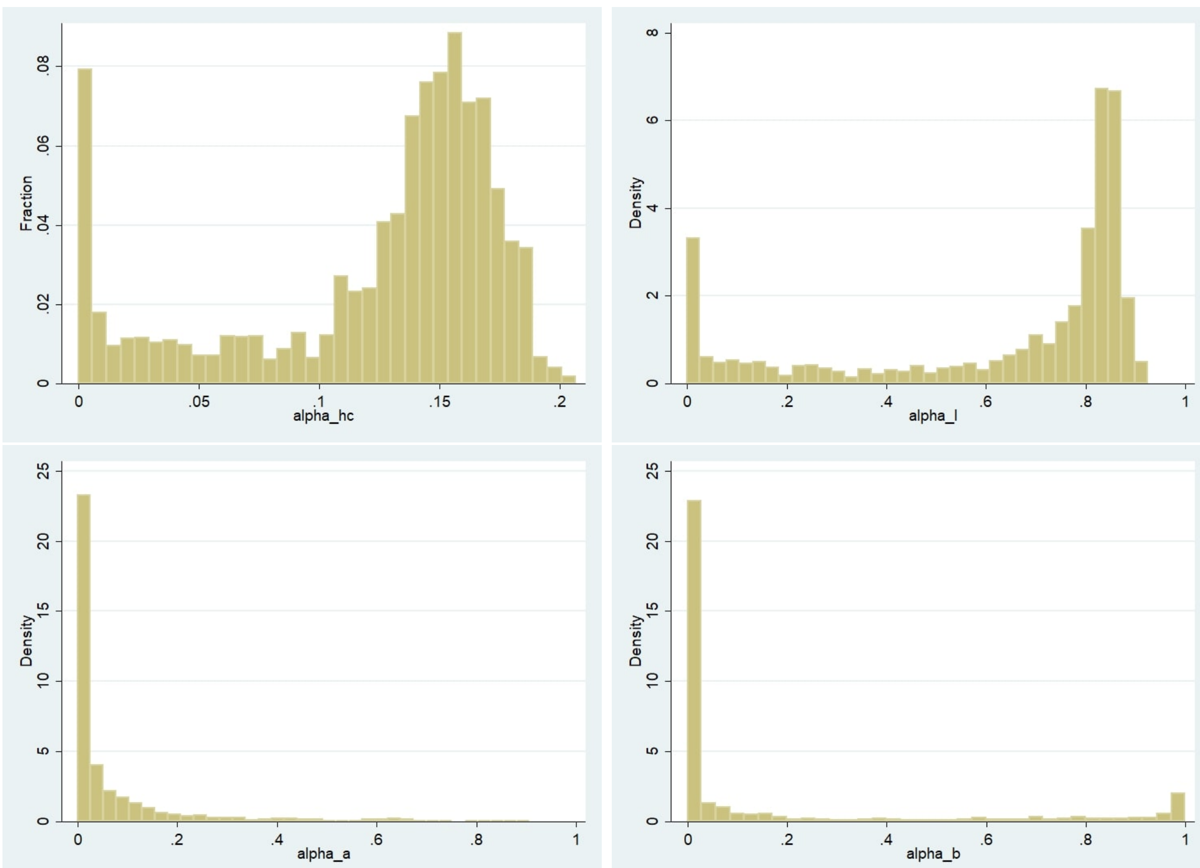
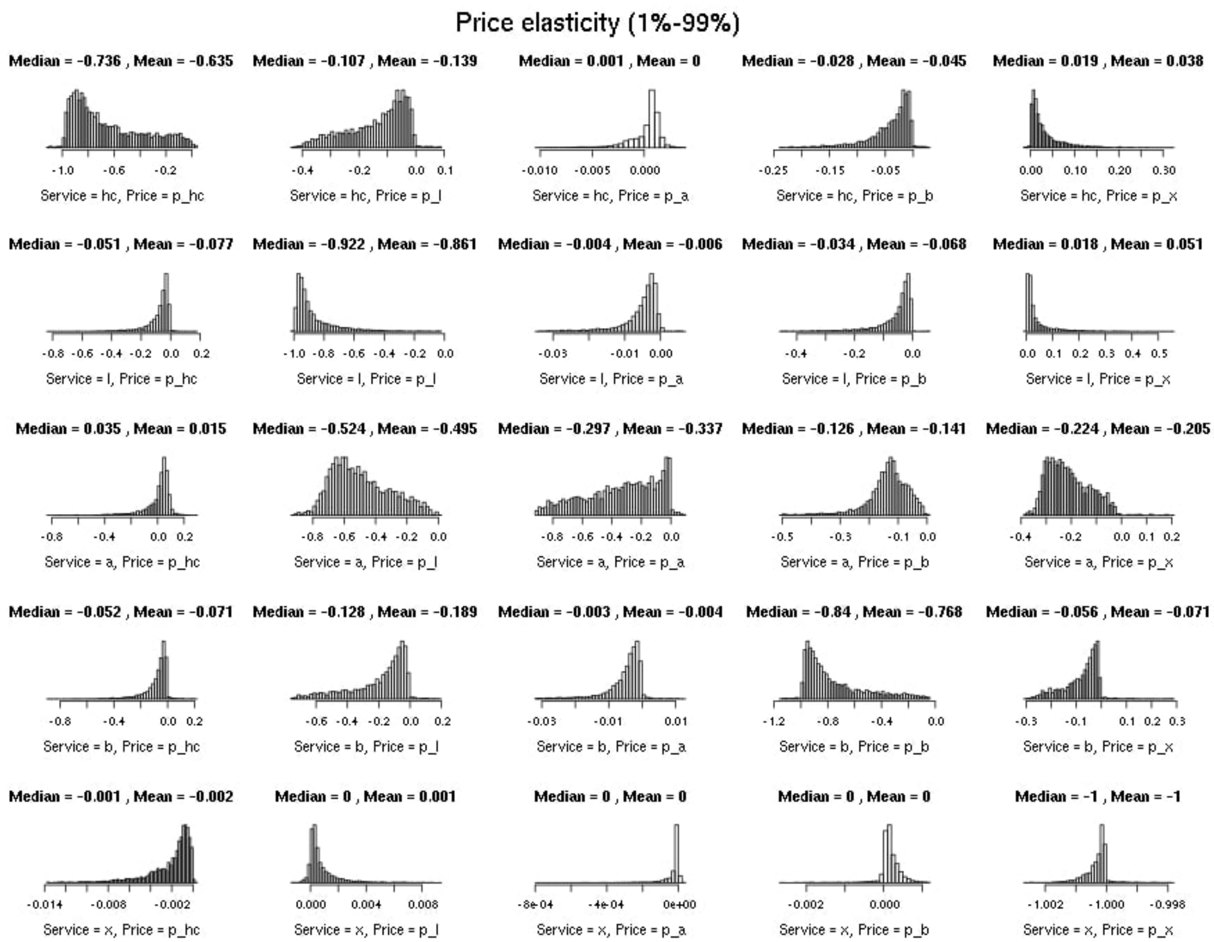
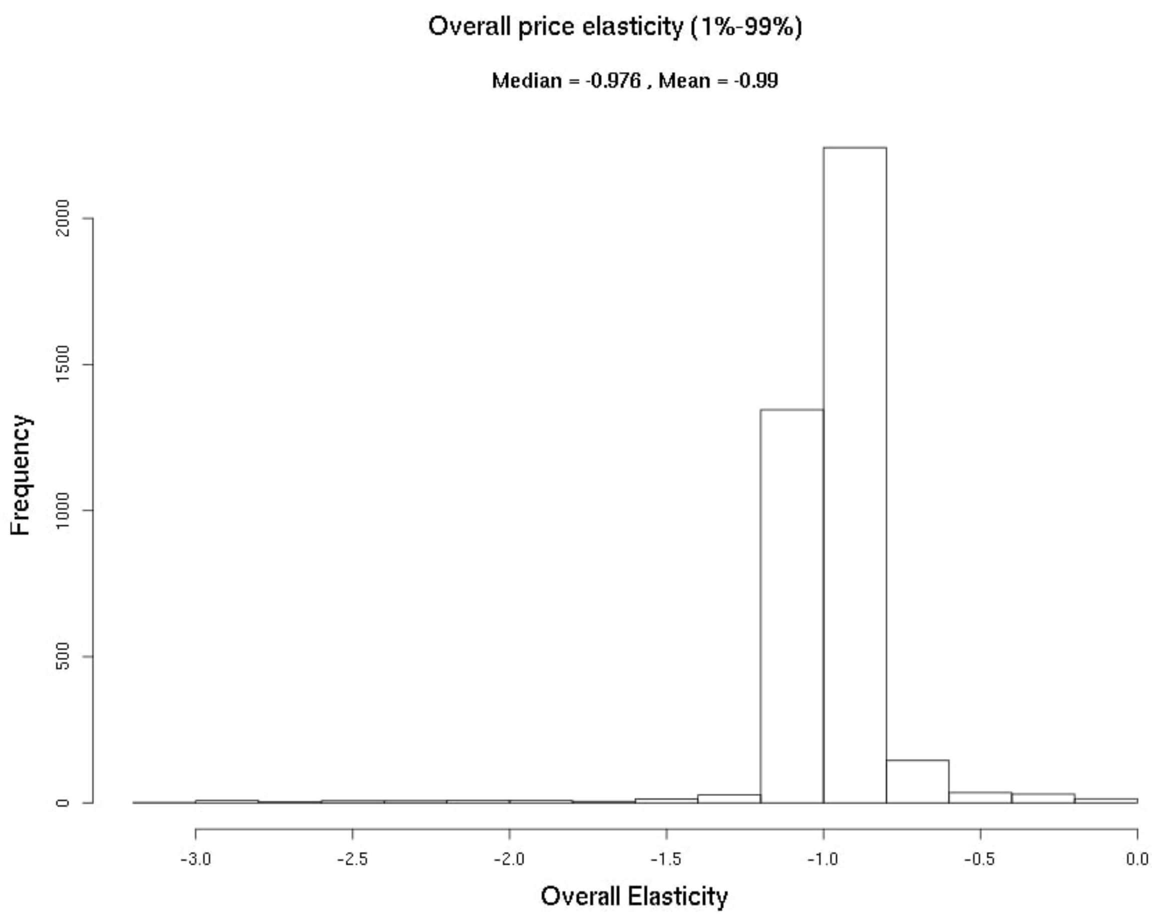


Figure 12: Histogram of price elasticity, using constrained kernel regression for electricity consumption function



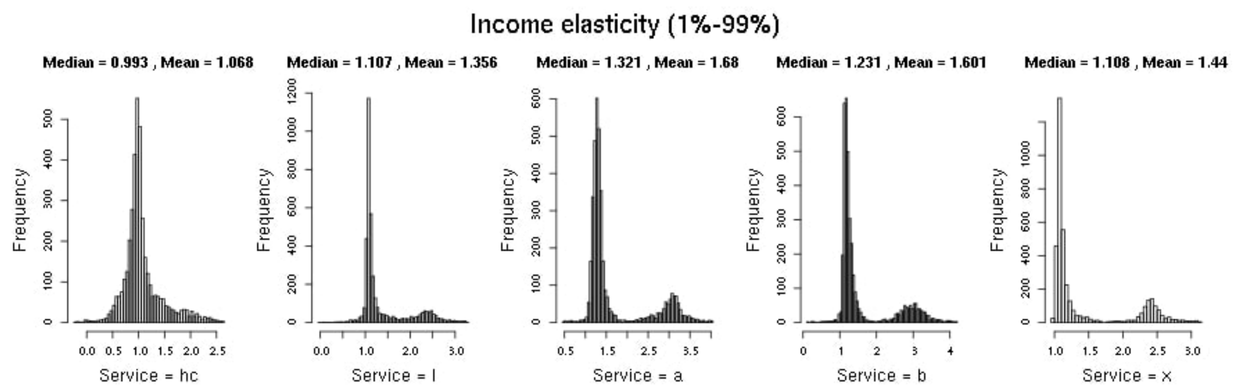
Note: To have a reasonable scale, we only present the data from 1st percentile to 99th percentile.

Figure 13: Histogram of overall price elasticity, using constrained kernel regression for electricity consumption function



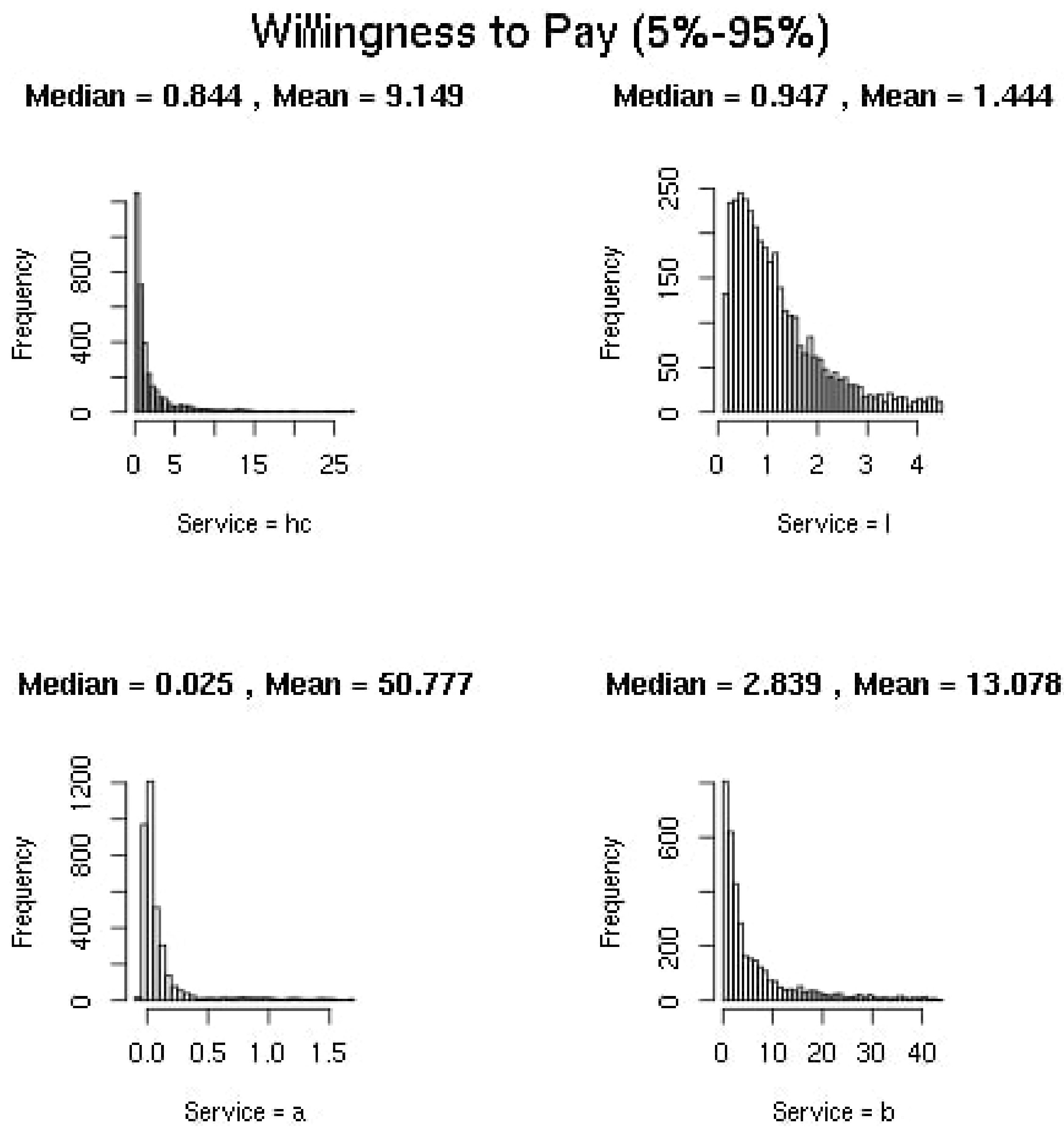
Note: To have a reasonable scale, we only present the data from 1st percentile to 99th percentile.

Figure 14: Histogram of income elasticity, using constrained kernel regression for electricity consumption function



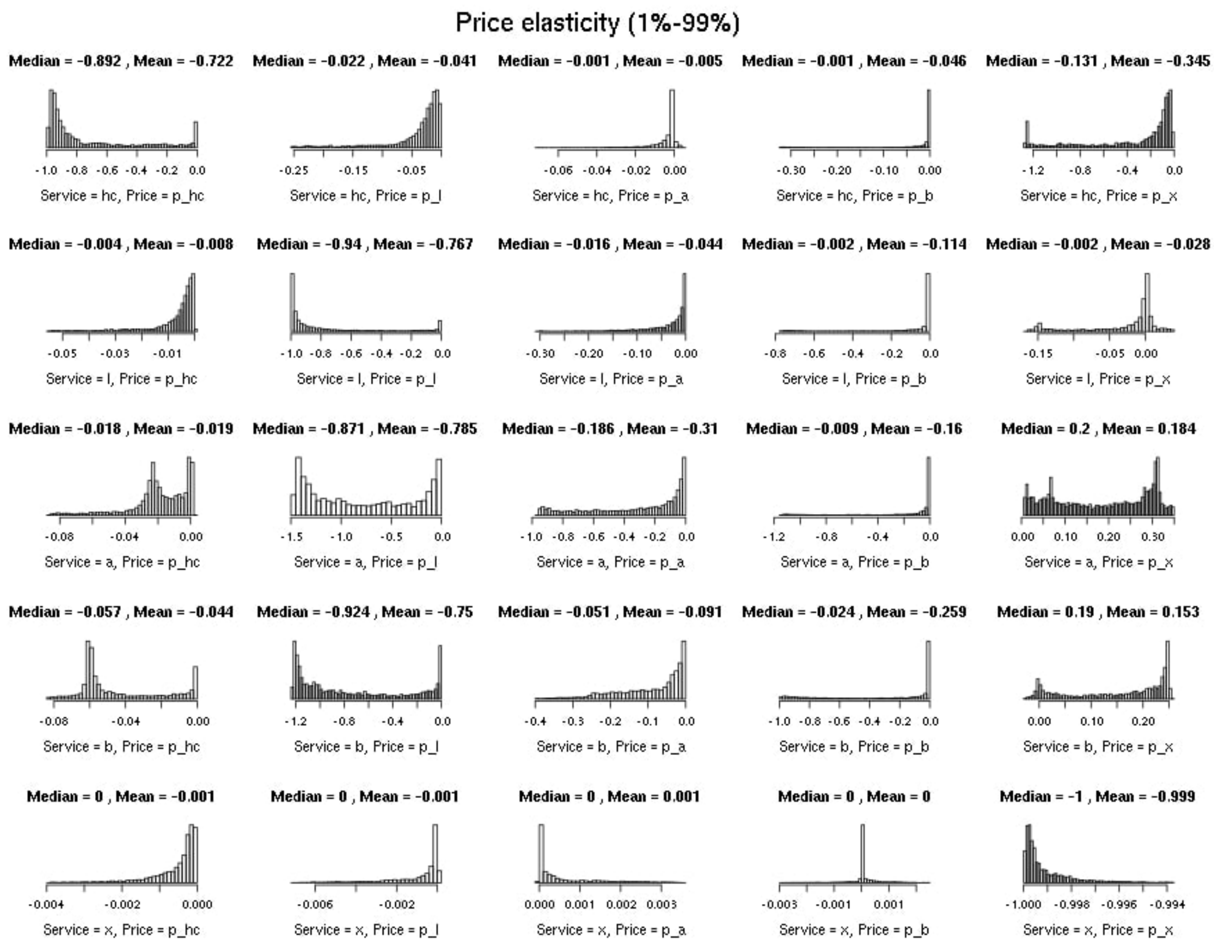
Note: To have a reasonable scale, we only present the data from 1st percentile to 99th percentile.

Figure 15: Histogram of willingness to pay, using constrained kernel regression for electricity consumption function



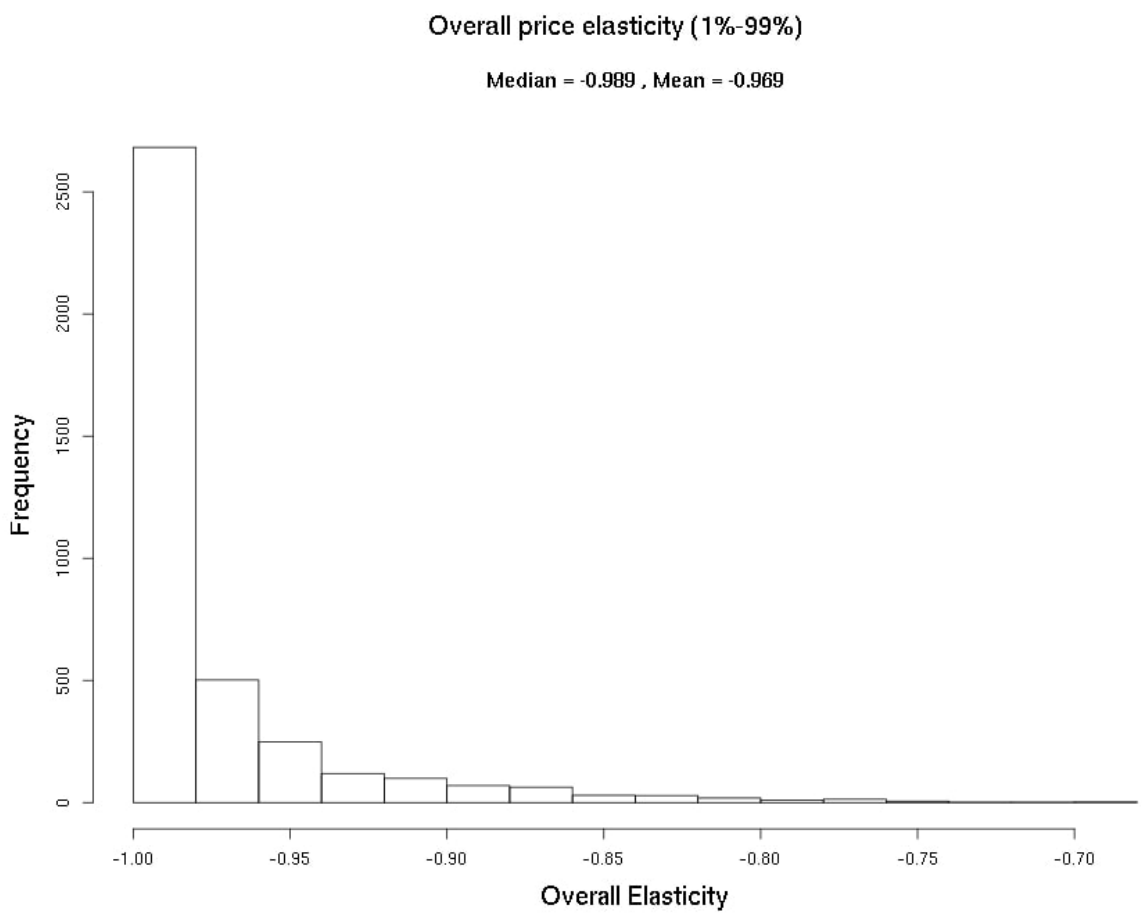
Note: To have a reasonable scale, we only present the data from 5th percentile to 95th percentile.

Figure 16: Histogram of price elasticity, using specified function form for electricity consumption function



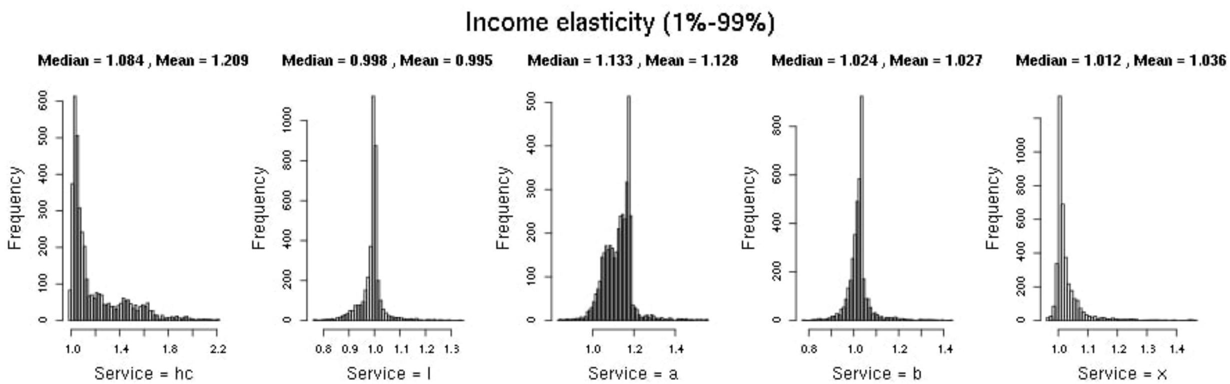
Note: To have a reasonable scale, we only present the data from 1st percentile to 99th percentile.

Figure 17: Histogram of overall price elasticity, using specified function form for electricity consumption function



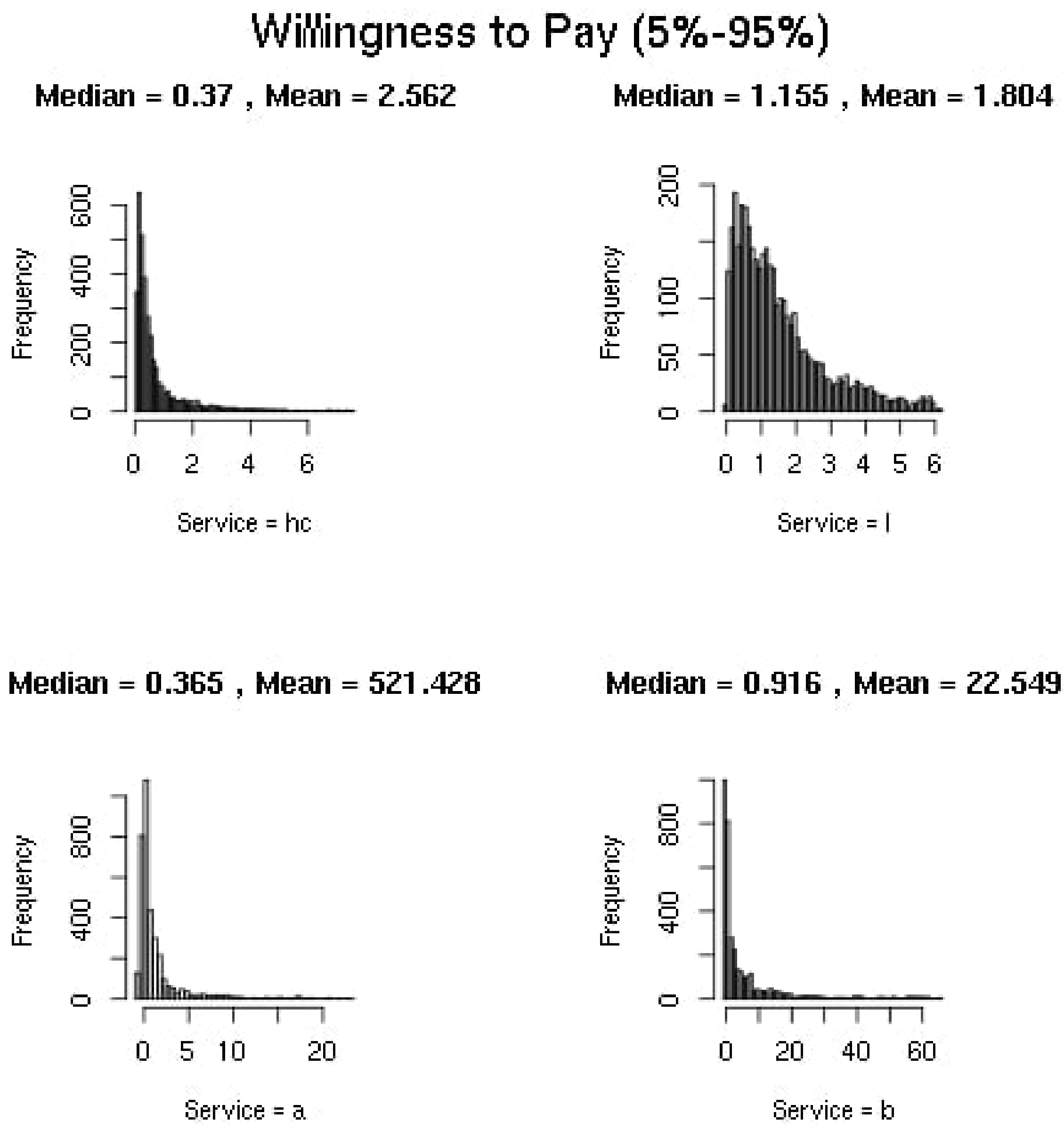
Note: To have a reasonable scale, we only present the data from 1st percentile to 99th percentile.

Figure 18: Histogram of income elasticity, using specified function form for electricity consumption function



Note: To have a reasonable scale, we only present the data from 1st percentile to 99th percentile.

Figure 19: Histogram of willingness to pay, using specified function form for electricity consumption function



Note: To have a reasonable scale, we only present the data from 5st percentile to 95th percentile.