Designing Minimum-Risk Nonlinear Price Schedules for Water Utilities*

by

Frank A. Wolak
Director, Program on Energy and Sustainable Development
Professor, Department of Economics
Stanford University
Stanford, CA  94305-6072
wolak@zia.stanford.edu
July 2015
Preliminary Draft
Please Do Not Distribute Beyond Conference Participants

*I would like to thank Michael Miller, Ivan Korolev, and Onder Polat for outstanding research assistance.
Abstract

This paper formulates and estimates a household-level demand for water under increasing block prices that accounts for the impact of monthly weather variation and customer-level heterogeneity in demand due to household demographics. The model utilizes PUMS-level data on the distribution of household demographics in the utility’s service territory to recover the impact of these factors on water demand. The household-level demand models can be used to compute the distribution of utility-level water demand and revenues for any possible price schedule. A number of counterfactual nonlinear pricing experiments are performed to assess the impact of resolving uncertainty about customer-level demand and the distribution of these demands throughout utility’s service territory on the system-wide sales and revenue uncertainty faced by the utility. Once set of counterfactuals finds the price schedule that reduces the system-wide demand for a water utility by 25 percent with a 95 percent probability subject while raising the same amount of system-wide expected revenue while sharing the burden of this consumption reduction across all demographic groups. Another counterfactual constructs nonlinear price schedules that yield no more than the same expected system-wide water sales and at least as much system-wide revenues as the utility’s existing price schedules, but also minimize the uncertainty in the utility’s annual revenues from water sales. These experiments demonstrate that knowledge of the customer-level demand can be used by the utility to design increasing block pricing plans that achieve any revenue or sales goals to optimize some policy function. Knowledge of how these demands different across customers based on observable household characteristics can allow the utility to reduce the utility-wide revenue or sales risk it faces for any pricing plan. Knowledge of how the structure of demand varies across customers can be used to design personalized (based on observable household demographic characteristics) increasing block price schedules to further reduce the risk the utility faces on a system-wide basis. For the utilities considered, knowledge of the customer-level demographics that predict demand differences across households reduced the uncertainty in the utility’s system-wide revenues from 85 to 95 percent. Further reductions in the uncertainty in the utility’s system-wide revenues in the, range of 5 to 30 percent, are possible by re-designing the utility’s nonlinear price schedules to minimize the revenue risk it faces given the distribution of household-level demand in its service territory.
1. Motivation for Research

An increasing number of urban water systems face the competing goals of encouraging water conservation and selling sufficient water to pay for the delivery infrastructure and increasing cost of the water sold to retail consumers. Historically, the major concern of urban water systems was ensuring adequate revenues to fund investments in new sources of water and the additional water delivery infrastructure needed to serve a growing demand. However, limited new sources of surface and groundwater and reduced reliability of existing sources have led to an increasing need for utilities to manage scarce water resources.

Nonlinear pricing is the standard approach used by utilities to balance these competing goals. Customers typically face rate schedules where the price charged for each additional unit, the marginal price, rises with the customer’s monthly consumption. The marginal price is fixed for a block or range of values of monthly consumption, but it increases across these blocks with increases in the value of monthly consumption. For this reason these nonlinear price schedules are called increasing block price schedules.

The form of the increasing block price schedule set by the utility impacts how much water each customer purchases and the revenues the utility receives from that customer. The form of the price schedules also impacts the amount of uncertainty the utility faces in the quantity of water it sells and the revenues it receives from each customer. This uncertainty in customer-level water sales and revenues to the utility is aggregated across customers to create uncertainty in the utility-level water sales and revenues. If a utility can accurately predict the customer-level demand for water for any possible nonlinear price schedule it can design increasing block price schedules to achieve any conservation or revenue goal while minimizing utility-level water sales or revenue risk. Increased information about the distribution of customer-level demand directly translates into reduced water sales and revenue risk associated with any rate design goal.

This paper formulates and estimates a household-level demand for water under increasing block prices that accounts for the impact of monthly weather variation and customer-level heterogeneity in demand due to observable demographic characteristics and other unobserved factors that differ across customers. This model can be used to construct an estimate of the distribution of each customer’s monthly demand and amount paid for water for any arbitrary nonlinear price schedule. Combined with data on the distribution on observable customer-level
heterogeneity in the utility’s service territory, these household-level demand models can be used to compute the distribution of aggregate water demand for any possible price schedule. This process also yields an estimate of the distribution of total utility-level revenues for any arbitrary nonlinear price schedule or set of nonlinear price schedules.

Consequently, this household-level water demand model can be used by the utility to design water rates to achieve a wide range of desired policy goals. The model can be used to design rate schedules that minimize the uncertainty in the utility’s total revenues subject to achieving given expected level of total utility revenues and water sales. It can also be used to design rate schedules that reduce expected payments by low-income consumers subject to achieving same level of total utility expected revenues as current nonlinear price schedule. The achievement of any aggregate water sales or revenues goals or distribution of water sales and revenues across customers can be optimized using this model applied to the utility’s customers.

The model assumes that water demand depends on the price schedule faced by the household, characteristics of household, such as household income, the size of the dwelling, size of the property, and number of adults living in dwelling, and weather conditions such as average daily temperature and rainfall during the customer’s billing cycle. These demand models are estimated for several water utilities charging increasing block price schedules using monthly billing cycle data for a sample of customers from each utility combined with data from the United States (US) Bureau of Census Public Use Microdata Sample (PUMS) of American Community Survey on the distribution of household demographic characteristics in the United States Postal Service (USPS) Zip Code and data from the National Oceanic and Atmospheric Administration (NOAA) on daily weather conditions in that Zip Code during the billing cycle.

Using the estimated demand models, a number of counterfactual nonlinear pricing experiments are performed to assess the impact of resolving uncertainty about customer-level demand and the distribution of these demands throughout utility’s service territory on the system-wide sales and revenue uncertainty faced by the utility. These counterfactuals construct nonlinear price schedules that yield no more than the same expected system-wide water sales and at least as much system-wide revenues as the utility’s existing price schedules, but also minimize the uncertainty in the utility’s annual revenues from water sales.

These counterfactual price schedules are constructed under two different sets of assumptions about what the utility is assumed to know about each customer and the form of their
demand for water. One set of counterfactuals assumes that utility only knows the distribution of demographic characteristics for customers in each Zip Code. The second set assumes the utility knows these demographic characteristic. Each counterfactual solves for the minimum system-wide revenue risk price schedule that obtains the no more than the same expected system-wide water sales and at least the same revenues as the current schedule. The difference in the system-wide revenue risk between no demographic information scenario and the one that assumes the utility knows each customer’s demographic information provides a metric for assessing the economic benefits to the utility from collecting this information from its customers.

These experiments demonstrate several sources of economic benefits to the utility from having more detailed knowledge individual customers. First, knowledge of the customer-level demand can be used by the utility to design increasing block pricing plans that achieve any revenue or sales goals with less revenue or sales risk. Second, knowledge of how these demands different across customers based on observable characteristics of the customers can allow the utility to significantly reduce the utility-wide revenue or sales risk it faces for any pricing plan. Third, knowledge of how the structure of demand varies across customers can be used to design personalized (based on observable household demographic characteristics) increasing block price schedules to further reduce the risk the utility faces on a system-wide basis. Because it is relatively straightforward for the utility to prevent resale of residential water service, utilities can set different increasing block price schedules for each customer based on its observable demographic characteristics. Finally, with detailed knowledge of how demands different across customers based on observable demographic characteristics, the utility can more accurately assess the likely to water sales and revenue impacts of changes in the number and types of customers in their service territory.

For the utilities considered, knowledge of the customer-level demographics that predict demand differences across households reduced the uncertainty in the utility’s system-wide revenues by almost 85 percent. Further reductions in the uncertainty in the utility’s system-wide revenues in the, range of 5 to 30 percent, are possible by re-designing the utility’s nonlinear price schedules to minimize the revenue risk it faces given the distribution of household-level demand in its service territory. This household-level demand information is also particularly important for assessing the economic benefits of proposed water infrastructure projects and designing the price
schedules necessary raise the revenue needed to pay for them with the least amount of water sales or revenue risk to the utility.

2. **Rate Design with Nonlinear Pricing**

The traditional rationale for increasing block pricing by water utilities is that this form of nonlinear prices balances two competing public policy goals. The first is to provide the “essential” amount of water a household needs for drinking, cooking, bathing, and other indoor use at a price that is affordable for virtually all households in the utility’s service territory. The second goal is to provide a financial incentive for households using more than the “essential” amount to reduce their demand for water. By this logic, the higher-priced steps in the increasing block price schedule beyond the initial baseline or essential consumption level are designed to discourage less essential water consumption. For example, the second price step might be intended for the demand to fill the household’s swimming pool. The third price step might be intended for the demand for watering the household’s outdoor trees, bushes, and shrubs. The fourth price step might be intended for the demand for watering the household’s lawn.

Another traditional argument in favor of increasing block pricing of water is that it recovers an increasing amount of the utility’s revenue from high demand customers, which are also likely to be the high income customers. Because higher income consumers typically consume more water, the highest marginal price they pay is typically greater than the highest marginal price low income consumers pay. For this reason, increasing block pricing implies that high income consumers pay a higher average price (total monthly payments divided by total monthly consumption) for their water consumption than low income consumers.

Increasing block pricing can also create revenue adequacy challenges for the water utility if the utility makes the length of the baseline level of demand too large. High demand households might consume along the baseline marginal price step as opposed to consuming at a higher marginal price step. Figure 1(a) illustrates case with $D_L(p)$, the demand curve for low-demand consumers, and $D_H(p)$, the demand curve for high income consumers. Both curves intersect the increasing block price schedule on the first price block, which raises significantly less revenue for the utility than would be the case if $D_H(p)$ intersected the price schedule on the higher-priced block. Setting the length of essential block too small can impose an excessive financial burden on low-demand, low-income consumers. Figure 1(b) illustrates this case where both demand curves intersect the increasing block price schedule on the higher-priced block.
From the perspective of achieving enough revenues to recover the utility’s cost and while selling no more than a certain amount of water to all customers, the design of a nonlinear pricing schedule amounts to choosing the length of each step to separate customer into distinct groups based on their willingness to pay for water. Moreover, if the utility has some uncertainty about the location and shape of each customer’s demand, then reducing this uncertainty could help the utility determine where to set the baseline demand level, $q_B$, shown in Figure 1(c). This figure shows the range of possible uncertainty (from the perspective of the utility) in $D_L(p)$ and $D_H(p)$. This is indicated by the dotted lines to the left and right of each demand curve. Note that $q_B$ has been chosen so that regardless of the realization of $D_L(p)$ and $D_H(p)$, each type of customer will continue to consume along the same step of the increasing block price schedule. Choosing the value of $q_B$ in this manner limits the amount of revenue variability facing the utility due to its uncertainty about the realized values of $D_L(p)$ and $D_H(p)$. The second part of the increasing block pricing design process must choose the levels of the first marginal price and second higher marginal price to recover sufficient revenues to cover the utility’s costs, which still achieving the goal of limiting the economic burden placed on low-income consumers to purchase their essential water needs.

One set of counterfactual pricing experiments uses the estimated demand models for a utility to determine the extent to which it is possible for that utility to re-design its increasing block price schedule to achieve at least as much expected revenue and expected water sales no larger than it does under the current rate schedule while facing less risk to its total revenues.

If the utility is able to sort households into different categories based on observable demographic characteristics, then it is possible to assign different increasing block price schedules to different households based on their observable demographic characteristics. In this case, the utility would like to achieve the outcome in Figure 1(c) for each set of observable demographic characteristics that predict differences in the form of the demand.

A second set of counterfactual pricing experiments will set separate increasing block prices schedules for households with different observable demographic characteristics to achieve at least as much expected revenue and no larger expected water sales than with the current increasing block price schedules used by the utility while facing the utility with less risk to its total revenues.
Both sets of counterfactual pricing experiments demonstrate that if a utility has more information about the demand for water of individual customers, it can significantly reduce the revenue or sales risk it faces in meeting a set of pricing goals.

3. Data Used in Analysis

Three datasets are used to estimate the customer-level demand model for each utility service territory. The first is billing cycle-level monthly water consumption data for a sample of households for at least one year in duration. The second dataset is composed of daily weather variables at the Zip Code level obtained from the National Oceanic and Atmospheric Administration (NOAA) for the utility’s service territory. The final dataset is the distribution of household-level demographic characteristics within each Zip Code in the utility’s service territory obtained from the US Bureau of the Census.

Monthly household-level water consumption is available from each utility at the billing cycle-level, along with the customer’s zip code, form of nonlinear price schedule faced by household, and other information necessary to compute customer’s monthly water bill. Although utilities typically bill their customers on a monthly basis, customers receive their bills at different times during the month. The time between consecutive billing dates is called the customer’s billing cycle depends on when the meter readers shows up at the customer’s premises to read its meter. For example, one customer might be billed on third day of every month, whereas another customer might be billed on a twentieth day of the month.

Having data available at the billing cycle level is important for accurately modeling the impact of weather conditions on a household’s demand for water. In terms of the above example, it might be the case that first two weeks of July are extremely hot so the water demand is extremely high, whereas the last two weeks of July are mild so that water demand is significantly lower. The customer with a billing cycle that starts on the third day of the month will have much higher weather-related demand than customer whose billing cycle begins on the 20th day of the month. Only by knowing the customer’s billing cycle is it possible to properly account for differences across customers in their weather-related demand for water.

The NOAA provides daily measures of rainfall and the maximum daily temperature at the Zip Code level for each utility service territory. The average value of the maximum daily temperature is computed as the average of the daily maximum temperature across all days in the billing cycle. The total amount of rainfall in that Zip Code during the billing cycle is also computed.
from this data. The inter-quartile range of the maximum daily temperatures and inter-quartile range of daily rainfall in the zip code during the billing cycle are also compiled. The inter-quartile range is the difference between the 75th percentile of the daily variables in the billing cycle and 25th percentile of this same distribution.

The distributions of household-level demographic variables for each Zip Code in each utility’s service territory are obtained from the US Bureau of Census Public Use Microdata Sample (PUMS) of American Community Survey. The demographic characteristics for each household surveyed in each Public Use Microdata Area (PUMA) are compiled along with the sampling weight for that household. These PUMAs can be matched to zip codes so that a distribution of household-level demographic variables in the Zip Code is available for all Zip Codes in the service territory.

The following utility service territories have provided customer-level billing cycle data: Cobb County, Georgia; Valley of Moon (near Sonoma), California; Phoenix, Arizona; Santa Rosa (County Seat of Sonoma County), California; Tacoma, Washington; and Washington, DC. Daily weather data has also been compiled for the time period that the customer-level billing cycle data is available for each utility service territory. The Zip Code-level distribution of household demographic data has also been compiled for the time period that the customer-level billing cycle data is available for each utility service territory.

All nominal prices are converted to 2012 dollars using the Federal Reserve Economic Database Gross Domestic Product (GDP) deflator from St Louis Federal Resource Bank.1

4. Econometric Model

This section describes the specification of the econometric model of the billing cycle-level and household-level model of water demand with increasing block prices that accounts for the weather facing that household during its billing cycle and differences in demographic characteristics across households. This model is derived from the assumption that households choose their level of water consumption to maximize a utility function that depends on their demographic characteristics and unobserved heterogeneity.

This economic model is used to derive the joint density of all billing cycle-level consumption choices for each household during the sample period conditional on the nonlinear price schedule the household faces, its demographic characteristics and the temperature and

---

1 Data available from https://research.stlouisfed.org/fred2/series/GDPDEF
rainfall distributions it was exposed each billing cycle during the sample period. Because the demographic characteristics of each household are unobserved, to arrive at the likelihood function used to estimate the parameters of the demand model, this conditional distribution of the household’s monthly billing cycle-level consumption choices is integrated with respect to the distribution of demographic characteristics in the Zip Code that contains that household. This yields a likelihood function that depends on observable data—the household’s vector of monthly water consumption choices, the vector of billing cycle-level monthly weather variables and the distribution of demographic characteristics for that household’s Zip Code.

4.1. Water Demand Model

Let \( U(x, w, A, Z, \epsilon, \beta) \) equal the utility function for household \( i \) over the \( N \)-dimensional vector of goods, \( x_i = (x_{i1}, x_{i2}, \ldots, x_{iN}) \), where \( x_{ik} \) is household \( i \)'s monthly consumption of good \( k \), and \( w_i \) is the household \( i \)'s monthly consumption of water. The utility function also depends on the household \( i \)'s demographic characteristics \( A_i \), the vector of weather variables faced by household \( i \), \( Z_i \), a vector of unobserved heterogeneity \( \epsilon_i \), and is parameterized by the vector \( \beta \). Let \( p_k \) equal the price of the \( k \)th element of \( x_i \), \( x_{ik} \). Let \( \theta_i(w) \) equal the increasing block price function that the household \( i \) faces for water. The value of this function at consumption level \( w \) is equal to, \( \theta_i(w) \), the marginal price. Figure 1(a) to 1(c) shows several increasing block price schedules with two price blocks.

If household \( i \) purchases \( w^* \) units of water during the month then its total bill is equal to \( R(\theta_i(w^*)) = \int_0^{w^*} \theta_i(s) \, ds \), which is equal to the area under the nonlinear price schedule up to the observed consumption level, \( w^* \). A household that consumes \( w \) units of water and the vector of other goods, \( x \), has a monthly spending on water and the \( N \) other goods equal to \( \sum_{k=1}^{N} p_k x_{ik} + R(\theta_i(w)) \). Under the assumption of utility-maximizing behavior, the household’s observed choices of \( x \) and \( w \) are assumed to be the solution to the following optimization problem:

\[
\max_{x \geq 0, w \geq 0} U(x, w | A_i, Z_i, \epsilon_i, \beta) \ \text{subject to} \ \sum_{k=1}^{N} p_k x_{ik} + R(\theta_i(w)) = M_i,
\]

where \( M_i \) is the household \( i \)'s monthly income. Solving problem (1) yields the household’s utility-maximizing choices for \( x \) and \( w \) as a function of the vector of prices, \( P = (p_1, p_2, \ldots, p_N) \) of the \( N \) other goods; the nonlinear price function, \( \theta_i(w) \); and total monthly expenditures, \( M_i \).

Let \( w(P, \theta_i, M_i, A_i, Z_i, \epsilon_i, \beta) \) equal the solution to the household-level optimization problem. It depends on the prices of other goods, \( P \); the nonlinear price schedule for water faced by household \( i \), \( \theta_i(w) \); household \( i \)'s total monthly income, \( M_i \); the vector of observed characteristics
of household i, $A_i$; the vector of unobserved characteristics of the household, $\epsilon_i$; the vector of weather variables faced by household i, $Z_i$; and the parameters of the household’s preference function, $\beta$.

Assuming a parametric joint density for $\epsilon$, $f(\epsilon|\delta)$, (where $\delta$ is the vector of parameters of this joint density) it is possible to derive the density of the household’s observed water consumption, $w$, $g(w|P, \theta, A, Z, \beta, \delta)$, which also equal to the conditional (on $A_i$) likelihood function for a single observation of monthly billing cycle-level consumption for household i.

### 4.2. Log-Likelihood Function

Let the subscript “t” denote the value of a variable for billing cycle t and $T(i)$ equal the number of monthly consumption observations for household i in the sample and $N$ is the total number of households in the sample. Let $W_i = (w_{i1}, w_{i2}, \ldots, w_{iT(i)})$ equal the $T(i)$ dimensional row vector monthly water consumption observations for household i. Let $W = (W_1, W_2, \ldots, W_N)$ equal the row vector of $N$ row vectors of monthly water consumption observations for all households in the sample. The first step in computing the likelihood function for the econometric model is to compute the joint density of $W_i$ for each household in the sample conditional on the household’s demographic characteristics and the $T(i)$ realizations of monthly weather conditions that they faced. In terms of the above notation, this joint density takes the form:

$$
\prod_{t=1}^{T(i)} g(w_{it} | P_{it}, \theta_{it}, A_n, Z_{it}, \beta, \delta) \tag{2}
$$

The PUMS data from American Community Survey can be used to compute the probability density functions for the vector of demographic characteristics for each Zip Code in the utility’s service territory. This dataset provides the sampling weights for each household in the American Consumer Survey and the vector of their demographic characteristics for each 5-digit Zip Code in the utility’s service territory. Let $(w_{i,n}, A_n)$ for $n=1, \ldots, L(i)$ equal the values of these sampling weights and associated vector of demographic characteristics for each sampled household in the Zip Code that contains household i. In terms of this notation, the log-likelihood function for single observation is equal to:

$$
L(W_i | \beta, \delta) = \ln[\sum_{n=1}^{L(i)} wti \prod_{t=1}^{T(i)} g(w_{it} | P_{it}, \theta_{it}, A_n, Z_{it}, \beta, \delta)]. \tag{3}
$$

Summing over all $N$ households in the sample yields log-likelihood function for the entire sample:

$$
L(W | \beta, \delta) = \sum_{i=1}^{N} \ln[\sum_{n=1}^{L(i)} wti \prod_{t=1}^{T(i)} g(w_{it} | P_{it}, \theta_{it}, A_n, Z_{it}, \beta, \delta)], \tag{4}
$$
4.3. Functional Forms

In order to implement the model empirically, it is necessary to choose functional forms for the household’s utility function, \( U(x_i, w_i, A_i, Z_i, \epsilon_i, \beta) \), which yields the functional form for the household’s demand function, \( w(P, \theta_i, M_i, A_i, Z_i, \epsilon_i, \beta) \). Because the distributions of monthly water consumption across both across households for the same month and for the same household over time are both positively skewed in the sense there many observations just below mean, but a few observations far above the mean, the appropriate variable to model is the logarithm of the household’s monthly demand for water.

This logic implies the following choice for the functional form \( w^*(\theta, M, A, Z, \epsilon, \beta) \), the observable portion household’s billing cycle-level monthly demand for water conditional on observing the household’s vector of demographic characteristics, \( A \):  
\[
\ln(w^*(p_w, V(A), A, Z, \epsilon, \beta)) = A'\beta_1 + Z'\beta_2 + \alpha(A)\ln(p_w) + \beta_3\ln(V(A)),
\]
where \( \alpha(A) = -exp(A'\beta_4) \). Define \( \beta = (\beta_1', \beta_2', \beta_3, \beta_4')' \) as the vector of parameters of the demand function. This functional form implies that the coefficient determining the price responsiveness of demand, \( \alpha(A) \), is minus 1 times the exponential function of a linear combination of some of the elements of the vector of household demographic variables. Where \( V(A) \) is the household’s monthly virtual income and it is written as a function of this vector of demographic characteristics to denote the fact that this variable depends on the elements of the vector of household characteristics that we “integrate” out with respect to in computing the likelihood function for household \( i \) in (4). The variable \( p_w \) is the marginal price of water for the step on the increasing block price schedule that the household is consuming at.

There are two sources of unobservables for each month and household \( \epsilon = (\eta, \nu) \), where \( \eta \sim N(0, \sigma^2_\eta) \) and \( \nu \sim N(0, \sigma^2_\nu) \) are independent random variables distributed independently across households and over time for the same household. This implies that \( \delta = (\sigma^2_\eta, \sigma^2_\nu)' \) in the notation of likelihood function (2). To construct the conditional (on demographic characteristics) likelihood function (2) for household \( i \), requires computing the density of the observed value of \( \ln(w_{it}) \), using the deterministic portion of the demand function and joint distribution of \( \epsilon \). The elements of \( \epsilon = (\eta, \nu) \) are called the unobserved household-level heterogeneity, \( \eta \), and the household-level optimization error, \( \nu \). The former is assumed to be observed by the household, but the latter is assumed to be unobserved by the household. Bother elements of \( \epsilon \) are unobserved by the econometrician.
To understand the mapping from \( \ln(w^*(p_w,V(A),A,Z,\beta)) \) to the logarithm of observed consumption of the household, consider the four-tier increasing block price schedule with a fixed charge. This implies \( p_1 < p_2 < p_3 < p_4 \). As shown in Figure 2(a), \((0,w_1^*)\) is the range of consumption where marginal price is \( p_1 \), \((w_1^*,w_2^*)\) is the range of consumption for the marginal price \( p_2 \), \((w_2^*,w_3^*)\) is the range of consumption for the marginal price \( p_3 \), \((w_3^*,\infty)\) is the range of consumption for the marginal price \( p_4 \), and FC equals the household’s fixed charge for the billing cycle. The increasing block price schedule implies the piece-wise linear budget set composed of the segments, BS1, BS2, BS3, and BS4 shown in Figure 2(a). For each segment of the increasing block price schedule, the segment of the household’s budget set becomes increasingly steep because the slope of each segment is equal to \(-p_k/p_o\), where \( p_k \) is the marginal price for the \( k \)th price block and \( p_o \) is the price of all other goods.

Define \( d_k \) as the difference between the cost of consumption level \( w \) (in the \( k \)th block of the increasing block price schedule) when all units are purchased at price \( p_k \), the marginal price for this block, and the actual cost purchasing \( w \) under increasing the block price schedule, so that

\[
d_k = p_kw - R(\theta(w)) = -FC - \sum_{j=1}^{k-1}(p_j - p_{j+1})w_j.
\]

For example, if the consumer is purchasing along the first price tier, then \( d_1 = -FC \) and if the consumer is purchasing along the second price tier, then \( d_2 = -FC - (p_2 - p_1)w_1 \).

Figure 2(a) shows a point of tangency between the household’s indifference curve and the nonlinear budget constraint. If \( \eta \) was the only unobservable in the household’s demand function the following statements would hold. If there is point of tangency between the household’s indifference curve and one of the segments of the piecewise linear budget set, there should be one value of \( \eta \) that yields this observed value of water consumption. However, as Figure 2(b) demonstrates, it is also possible that a point of tangency could occur at a kink point of the piecewise linear budget set. In Figure 2(b) the kink point is at the consumption level \( w_2^* \). In this case there would be a set of values of \( \eta \) such that the household consumes \( w_2^* \) because there are a number of possible values of \( \eta \) that shifts and rotates the household’s indifference curves which yields this point as the household’s utility maximizing consumption choice.

It is very unlikely that a household could actually manage its water consumption with so much precision as to end up exactly at a kink point on the piecewise linear budget set. Virtually all household uses of water involve uncertainty in the actual amount of water consumed that is unobservable to the household. A household taking bath or shower, filling their swimming pool,
washing their car, or watering their plants or lawn, in spite of their best intentions to use only a
certain amount of water for each these tasks, faces uncertainty in the exact amount of water they
use. For this reason, a second source of error, \( \nu \), the so-called optimization error, is introduced
into the demand model to account for uncertainty in actual amount of water consumed by the
household relative to their intended consumption level based on only \( \eta \).\(^2\)

The mapping from the realized values of the unobservables \((\eta, \nu)\) to the observed value of
the logarithm of the household’s monthly billing cycle-level consumption, \( \ln(w) \), for a K-step
increasing block price schedule takes the form

\[
\begin{align*}
\ln(w) &= \ln(w^*(p_1,V_1(A),Z,\beta)) + \eta + \nu \\
& \quad \text{if } \eta < \ln(w_1^*) - \ln(w^*(p_1,V_1(A),Z,\beta)) \\
\ln(w) &= \ln(w_1^*) + \nu \\
& \quad \text{if } \ln(w_1^*) - \ln(w^*(p_1,V_1(A),Z,\beta)) < \eta < \ln(w_1^*) - \ln(w^*(p_2,V_2(A),Z,\beta)) \\
\ln(w) &= \ln(w_2^*) + \nu \\
& \quad \text{if } \ln(w_2^*) - \ln(w^*(p_2,V_2(A),Z,\beta)) < \eta < \ln(w_2^*) - \ln(w^*(p_2,V_2(A),Z,\beta)) \\
& \quad \quad \quad \text{...} \\
\ln(w) &= \ln(w_{K-1}^*) + \nu \\
& \quad \text{if } \ln(w_{K-1}^*) - \ln(w^*(p_{K-1},V_{K-1}(A),Z,\beta)) \quad \text{< } \eta \quad \text{< } \ln(w_{K-1}^*) - \ln(w^*(p_{K},V_{K}(A),Z,\beta)) \\
\ln(w) &= \ln(w_{K}^*) + \nu \\
& \quad \text{if } \ln(w_{K}^*) - \ln(w^*(p_{K},V_{K}(A),Z,\beta)) < \eta \\
\end{align*}
\]

(6)

where \( V_k(A) = M(A) - d_k \) for \( k=1,2,\ldots,K \) and \( M(A) \) is household’s monthly income which is
written as a function of \( A \), the vector of household demographics, because the household’s monthly
income is one of the elements of \( A \).

In terms of this notation, the likelihood function conditional of \( A \), for single household and
billing cycle pair is equal to:

\[
\begin{align*}
\sum_{k=1}^{K} \phi \left( s_k \right) \left( \Phi \left( r_k \right) - \Phi \left( n_k \right) \right) + \sum_{k=1}^{K-1} \frac{\phi \left( u_k \right)}{\sigma^2} \left( \Phi \left( m_k \right) - \Phi \left( t_k \right) \right)
\end{align*}
\]

(7)

\(^2\) A similar error structure is employed by Hewitt and Hanemann (1995) and Olmstead, Hanemann, and Stavins
(2007) to derive the likelihood function for their demand models.
where \( t_k = \left[ \ln(w_k^*) - \ln(w^*(p_k, V_k(A), A, Z, \beta)) \right] / \sigma_{t_k}, \quad r_k = \frac{(t_k - \rho sk)}{\sqrt{1 - \rho^2}}, \quad \rho = \frac{\sigma^2_{w_i}}{\sqrt{\sigma^2_{w_i} + \sigma^2_{\beta}}}
\)

\( s_k = \left( \ln(w_{it}) - \ln(w^*(p_k, V_k(A), A, Z, \beta)) \right) / \sqrt{\sigma^2_{w_i} + \sigma^2_{\beta}}, \quad n_k = \frac{(m_{k-1} - \rho sk)}{\sqrt{1 - \rho^2}}, \quad m_k = \left( \ln(w_k^*) - \ln(w^*(p_{k+1}, V_{k+1}(A), A, Z, \beta)) \right) / \sigma_{t_k}, \quad u_k = \left( \ln(w_{it}) - \ln(w_k^*) \right) / \sigma_{w_i}.
\)

The multiplying this likelihood for billing cycle \( t \) for observation \( i \) by this same likelihood for all \( T(i) \) months for household \( i \) yields the likelihood function for observation \( i \) given in equation (2).

Maximizing this likelihood with respect to the parameters \((\beta', \delta')'\) yields the maximum likelihood estimates of this parameter vector. Two sets of standard errors for the parameter estimates are computed. The first set uses the inverse of the matrix of the sum of the outer products of gradient of the log-likelihood for each household evaluated at the maximum likelihood parameter estimates. The second set uses the White (1982) quasi-maximum likelihood estimate covariance matrix which is equal to the inverse of the matrix of second partial derivatives evaluated at the maximum likelihood parameter estimates pre- and post-multiplied by the matrix of the sum of the outer products of gradient of the log-likelihood function.

4.2. Estimation Results

Parameter estimates have been obtained for three utilities—Cobb County, Valley of the Moon (VoM) and Tacoma.\(^3\) Table 1 contains the estimation results for Cobb, Table 2 for VoM, and Table 3 for Tacoma. The coefficient estimates and the two sets of standard errors described above are reported for each region. The number of households in the sample is also reported for each region. There are different numbers of months of data for each household because of differences in billing cycles across households during the sample period for each utility.

The following variables make up \( Z_{it} \), the vector of weather characteristics that customer \( i \) was exposed to during billing cycle \( t \).\(^4\)

* Average high temperature: The average of the daily maximum temperature values in household \( i \)’s Zip Code during household \( i \)’s billing cycle.

* Inter-quartile range of maximum daily temperatures: The 75\(^{th}\) percentile of the daily maximum temperature values in household \( i \)’s Zip Code during household \( i \)’s billing cycle minus the 25\(^{th}\) percentile.

\(^3\) The estimation is still in progress for Phoenix. For Washington, DC there is not sufficient price variation within the region to estimate the model separately for these utilities, although the model that pools all of the utilities together can recover customer-level demand estimates for these regions.

\(^4\) All of the Zip Code-level weather data for each utility was obtained from the www.wunderground.com.
percentile of the daily maximum temperature values in household i’s Zip Code during household i’s billing cycle

*Total precipitation in billing cycle:* Sum of daily precipitation in inches during the billing cycle for the Zip Code containing household i.

*Interquartile range of daily precipitation:* The 75th percentile of the daily precipitation in household i’s Zip Code during household i’s billing cycle minus the 25th percentile of the daily precipitation in household i’s Zip Code during household i’s billing cycle

The household-level demographics variables, the vector A, all come from the PUMS data set. A subset of the available demographic variables that are thought to predict differences in water demand across households are included A.

*Monthly income of household:* Monthly household income in 2012 dollars. (Annual number reported in PUMS data divided by 12)

*Number of people over 18 years-old living in the household*

*Number of people under 18 years-old living in the household*

*House Size Indicators--House acreage between 1 and 10 acres.* House acreage above 10 acres.

*Number of bedrooms in the house*

*Number of vehicles owned by household*

As discussed earlier, for each household sampled by the US Bureau Census in that Zip Code, this demographic information is reported along with a sampling weight which gives the number of households in the Zip Code estimated to have this same vector of demographic characteristics as the sampled household. Dividing each sampling weight by the sum of the sampling weights for all households sampled in that zip code yields the weight, wt(i,n), used in the construction of the likelihood function.

Because the coefficient on the logarithm of the price in the demand function depends on the demographics of the household, each table also reports the value of the coefficient on price evaluated at the estimated population mean of the vector of demographic characteristics for each utility. The estimated population mean of the vector of demographics for the utility’s service territory is equal to

$$
\bar{A} = \sum_{i=1}^{N} \sum_{n=1}^{L(i)} wt(i, n) A_{in},
$$

(8)

where N is number of households in the sample and L(i) is the number of households sampled by the US Bureau of Census in the PUMS data in the Zip Code containing household i.
The price coefficient differs across households in the utility service territory, because the coefficient on the logarithm of price depends on \( A_{sn} \). Nonlinear pricing of water and the assumed stochastic structure described in the previous subsection that gives rise to the joint density of \( W_i \), the vector of billing cycle-level consumption values for household \( i \), implies that the coefficient on the logarithm of price for a given household cannot be interpreted as a price elasticity of demand. The same logic applies to the coefficient on logarithm of household-level income. Nevertheless, as shown in the following section, analogues to price and income elasticities can be computed with respect to the expected water demand of the household.

The model parameter estimates can be used to perform a number of counterfactual calculations to assist a water utility with designing increasing block price schedules that achieve water sales or revenue stabilization goals or other water pricing goals. In order to assess the extent to which knowledge of customer demographics can reduce both revenue risk and water sales risk for the utility, counterfactuals are run assuming the household’s demographics are known and the only source of randomness in the customer’s demand are the values of the customer-level heterogeneity (\( \eta \)) and optimization error (\( \nu \)).

Parameter estimates of the model and the likelihood function can be used to compute the posterior probability that household \( s \) has the vector of demographics \( A_{sn} \)

\[
p(A_{sn} | W) = \frac{\sum_{n=1}^{N(s)} wt(s,n) \prod_{t=1}^{T(s)} g(w_{st} | p_{st}, \theta_{st}, A_{sn}, Z_{mt}, \beta, \delta)}{\sum_{n=1}^{N(s)} wt(s,n) \prod_{t=1}^{T(s)} g(w_{st} | p_{st}, \theta_{st}, A_{sn}, Z_{mt}, \beta, \delta)},
\]

For each household in the sample, compute the \( L(s) \) values of \( p(A_{sn} | W) \) for \( s=1,2,\ldots,L(s) \). The value of \( A_{sn} \) that has the highest posterior probability for that household is assigned that vector of demographics for the purposes of computing the counterfactuals price schedules that assume the utility knows each household’s demographic attributes.

5. Using Model to Reduce Revenue and Quantity Risk

The estimates of the parameters of the household-level demand model given in Tables 1 to 3 make it possible to compute an estimate of the distribution of a household’s water consumption and monthly bill for any nonlinear price schedule either conditional on the household’s assigned demographic characteristics or without conditioning on the household’s demographic characteristics.

The expected value and variance of these magnitudes can be computed as follows. For counterfactual price schedule that could depend on the household’s demographic characteristics,
θ^C(w,A^*), a household with demographics A^* has expected consumption and the variance in this consumption equal to:

\[
E[w(P,θ^C,M,A^*,Z,ε,β)] = \int_{-\infty}^{\infty} w(P,θ^C,M,A^*,Z,s,β)f(s,δ)ds,
\]

\[
V[w(P,θ^C,M,A^*,Z,ε,β)] = \int_{-\infty}^{\infty} (w^*(P,θ^C,M,A^*,Z,s,β) - E[w^*(P,θ^C,M,A^*,Z,ε,β)])^2 f(s,δ)ds,
\]

where β and δ in the above expression are evaluated at the maximum likelihood estimates given in Tables 1 to 3. The expectations in the above expression are be taken with respect distributions of ε given A^* assigned by the rule based on equation (9). A household with assigned demographic characteristics A^* has an expected monthly water bill and the variance of its monthly water bill equal to:

\[
E[R(θ^C(w^*(P,θ^C,M,A^*,Z,ε,β),A^*))] = \int_{-\infty}^{\infty} R(θ^C(w^*(P,θ^C,M,A^*,s,β)) A^*)f(s,δ)ds,
\]

\[
V[R(θ^C(w^*(P,θ^C,M,A^*,Z,ε,β),A^*))] = \int_{-\infty}^{\infty} R(θ^C(w^*(P,θ^C,M,A^*,Z,s,β),A^*)) - E[R(θ^C(w^*(P,θ^C,M,A^*,Z,ε,β),A^*))] f(s,δ)ds.
\]

For the case that the household i’s demographics are assumed to be unknown, the household’s expected monthly water consumption and bill and the variance in its monthly water consumption and bill for the demographic characteristics-dependent increasing block price schedule, θ^C(w,A^*), are equal to:

\[
E[w^*(P,θ^C,M,A^*,Z,ε,β)] = \sum_{n=1}^{L(i)} w(t, n) \int_{-\infty}^{\infty} w^*(P,θ^C,M,A^*,Z,s,β)f(s,δ)ds,
\]

\[
V[w^*(P,θ^C,M,A^*,Z,ε,β)] = \sum_{n=1}^{L(i)} \int_{-\infty}^{\infty} (w^*(P,θ^C,M,A^*,Z,s,β) - E[w^*(P,θ^C,M,A^*,Z,ε,β)])^2 f(s,δ)ds,
\]

and

\[
E[R(θ^C(w^*(P,θ^C,M,A^*,ε,β),A))] = \sum_{n=1}^{L(i)} \int_{-\infty}^{\infty} wt(i,n)R(θ^C(w^*(P,θ^C,M,A^*,Z,s,β),A^*))f(s,δ)ds,
\]

\[
V[R(θ^C(w^*(P,θ^C,M,A^*,ε,β),A))] = \sum_{n=1}^{L(i)} \int_{-\infty}^{\infty} wt(i,n)(R(θ^C(w^*(P,θ^C,M,A^*,Z,s,β)) - E[R(θ^C(w^*(P,θ^C,M,A^*,Z,ε,β),A))])^2 f(s,δ)ds.
\]

The expectations in (14) to (17) are taken with respect to the distribution of ε = (η, ν) and the distribution of the demographic characteristics within the household’s Zip Code. The expectations in (10) to (13) are taken with respect to the distribution of ε = (η, ν) for the value of the household’s demographic characteristics assigned using the approach described above. Consequently,
comparing the variance of water consumption and total revenues given the assigned value of $A$
and the variance with respect to distributions of $\varepsilon$ and $A$ provides a measure of the value of
demographic information to utility.

These expressions in (10) and (14) can also be used to compute analogues to the price
elasticity and income elasticity of the demand for water. For the case of the price elasticity this
is computed as

$$\frac{\{E[w(P,0,M,A^*,Z,\varepsilon,\beta)] - E[w(P,\theta+,M,A^*,Z,\varepsilon,\beta)]\}/\{0.05*E[w(P,0^C,M,A^*,Z,\varepsilon,\beta)]\}}$$

where $\theta$ is the actual nonlinear price schedule charged by the utility and $\theta+$ is the actual nonlinear
price schedule with each price step multiplied by 1.05. This “price elasticity” is the percent change
in household $i$’s expected water consumption as a result of a 5 percent increase in all prices on the
nonlinear price function divided 0.05. Computing an “income elasticity” as the percent change in
expected consumption from a 5 percent increase in household $i$’s income divided by 5 percent
yields the coefficient on logarithm of income. Consequently, the model of demand with nonlinear
pricing and demographic characteristics in the price coefficient implies a different “price
elasticity” for each household, but the same income elasticity for each household.

The “price elasticities” can be computed conditional on the vector of the household’s
demographic characteristics or unconditional on the household’s vector of demographic
characteristics. The only differences in the two “price elasticities” is whether the expectations in
(18) are taken with respect to the distribution of $A$ or assume a fixed value of $A$.

It is also possible to compute the distribution of water consumption for all households in
the utility’s service territory and analogous aggregate demand elasticity estimates. Suppose there
are $J$ types of households, where households of type $j$ have a vector of observed attributes, $A_j$, and
$H_j$ is the number of type $j$ customers in the utility’s service territory. This implies that the expected
sales of water by the utility (summed across all customers) associated with rate schedule $\theta^C(w,A)$
is:

$$\text{Expected System-wide Water Sales} = \sum_{j=1}^{J} E[w^*(P, \theta^C, M, A_j, \varepsilon, \beta)]H_j$$

$$\text{Variance in System-wide Water Sales} = \sum_{j=1}^{J} \text{Var}[w^*(P, \theta^C, M, A_j, \varepsilon, \beta)]H_j.$$

Following the same procedure for system-wide revenues yields:

$$\text{Expected System-wide Revenues} = \sum_{j=1}^{J} E[R(\theta^C(w^*(P, \theta^C, M, A_j, \varepsilon, \beta), A)]C_j$$

$$\text{Variance in System-wide Revenues} = \sum_{j=1}^{J} \text{Var}[R(\theta^C(w^*(P, \theta^C, M, A_j, \varepsilon, \beta), A)]C_j.$$
Given the estimated distribution of $\varepsilon = (\eta, \nu)$ and the distribution of demographic attributes in each Zip Code within the utility’s service territory, other functions of the distribution of system-wide sales and revenues can be computed. The water utility or its regulatory body might be interested in the probability that system-wide sales or revenues exceed or fall below a pre-specified value for a prospective rate schedule. The model estimates can be used to compute that probability.

The aggregate or system-wide “price elasticity of demand” can be computed by finding the percentage increase in expected system-wide demand as a result of a 5 percent increase in all price steps faced by all customers divided by 5 percent. The aggregate “income” elasticity is the percentage increase in expected system-wide demand as a result of a 5 percent increase in all customer incomes divided by 5 percent. Because the “income elasticity” is the same for all households in a given utility service territory, the system-wide “income elasticity” is the same as the household-level “income elasticity”.

6. Counterfactual Price Schedules

This section reports on both the individual and system-wide “price” and “income” elasticities for each set of parameter estimates. It also reports on the computation of various counterfactual price schedules to achieve various policy goals. For example, for Cobb, I compute price schedules which yield the same or superior sales and revenue outcomes for these two utilities while subjecting the utility to less annual revenue risk. The counterfactual price schedule minimizes the standard deviation of utility-wide water revenues subject to the constraints that the utility expects to sell no more water than it does under the current price schedule and raises at least as much total revenue for the utility as the existing rate.

For VoM, I compute a counterfactual price schedule that is consistent with California’s current water demand reduction goals and is unlikely to run afoul of Proposition 218, which requires that municipal water customers only pay for the cost of the water that they consume. Specifically, I compute a price schedule which yields 25 percent less system-wide water sales than the existing price schedule with 95 percent probability and achieves the same system-wide expected revenue goals as the existing price schedule, while minimizing the financial burden of achieving these water consumption reduction goals across all classes of customers. The price schedule chosen minimizes the weighted sum of the squares of the difference between expected payments by each household under the counterfactual schedule and the current price schedule weighted by the inverse of that household’s expected payments under the current price schedule.
This objective function places the greatest burden to achieve water consumption reductions on households that currently have the largest water bills.

For each counterfactual price schedule I impose two additional constraints. First, the lowest marginal price in the counterfactual price schedule cannot be higher than the lowest marginal price in the actual price schedule. Second, the highest marginal price in the counterfactual price schedule cannot be higher than the highest marginal price in the actual price schedule. Each counterfactual price schedule is allowed to have as many marginal price steps as the actual price schedule subject to these two constraints. For Cobb, the counterfactual price schedules did not change, FC, the up-front fixed charge that the household faces. For VoM, FC was reduced to achieve the goal of maintaining system-wide expected revenues equal to those under the current price schedule.

Four conclusions emerge from this counterfactual price schedule design exercise:

1) The model of the household-level demand for a water utility can be used to reduce significantly the system-wide revenue or sales risk associated with achieving any water pricing goal.

2) By compiling information on the demographic characteristics of their customers and building this information into the utility’s model of household-level water demand, utilities can significantly reduce (up to 95% for the utilities considered) both the water sales and revenue risk associated with any expected water sales and revenue goals.

3) The customer-level model of demand incorporating demographic characteristics can also be used to design rates that allow the utility to achieve revenue and sales goals in response to changes in the number and distribution of customers in the utility’s service territory with minimal revenue risk.

4) The customer-level model of demand incorporating demographic characteristics can be used to design a menu of price schedules that can be offered to households (that allows them to select which specific price schedule they would like to be on) to achieve a given water pricing goal for the utility.

This price optimization framework can readily incorporate constraints like the majority of customers have the same or lower monthly water bills under the optimal price schedules compared to the current schedules. This framework can also be used to achieve water sales risk goals for utilities facing limited total water availability, such as designing a price schedule that makes the probability of a water shortfall (water demand from it customer less than the available supply) for
a utility less than a pre-specified amount subject to achieving certain revenue goals with some probability.

For both Cobb and VoM, a set of counterfactual price schedules are computed assuming that the household’s vector of demographic characteristics are unknown and a second set is computed assuming that each household is assigned a vector of demographic characteristics according to the rule described above. For the first set of counterfactual price schedules, the price schedule choice optimization problem assumes the demographics of a household are drawn from the same distribution that was used to estimate the parameters of the customer-level demand model. For each of the 50 vectors of realized demographic variables drawn from the distribution of demographic characteristics for household i’s Zip Code, the realized vectors of demographics characteristics for the household are kept fixed throughout the year across all billing cycles.

For each of these vectors of demographics, 50 values of ε are drawn from the estimated distribution of $\epsilon = (\eta, \nu)$. For any nonlinear price schedule that the household faces, the observable weather exposure variables for that household’s billing cycles, the drawn vectors of demographic characteristics, and billing cycle-specific shocks to water demand, and the demand model parameter estimates are used to simulate the household’s realized water demand. For each customer-billing cycle pair, there are $50 \times 50 = 2500$ simulated consumption values for each household in the sample under the chosen price schedule. Computing the sample average and variance of these simulated values yields an unbiased estimated of the mean and variance of the household’s consumption for the chosen price schedule for that billing cycle for the parameter values in Tables 1 to 3.

For the predicted demographics case, this mean and variance of billing cycle-level household consumption is computed conditional on the most likely vector of demographics in household i’s Zip Code for household i computed as follows. For the estimated parameters of the demand model, find the vector of household demographic characteristics in the Zip Code containing household i that has the highest posterior probability given the observed billing cycle-level water consumption levels and billing cycle-level monthly weather exposure variables for that household. This vector of demographics is assigned to that household and held fixed for all billing cycle-level water demand simulations. Fifty values of water consumption are simulated for each household-billing cycle combination by drawing 50 values from the estimated distribution of $\epsilon = (\eta, \nu)$. The sample mean and sample variance of these 50 simulated consumption values is an
unbiased estimate of the mean and variance of the household’s consumption for a billing cycle given the assumed vector of demographic characteristics for that household for the parameters values in Tables 1 to 3.

6.1. Cobb—Demographics Drawn

Figure 3 plots distribution of actual household-level billing cycle-level water consumption in thousands of gallons (TGAL) for all households and months during the sample period in the gray histogram. This figure also plots in blue the histogram of the simulated household-level billing cycle water consumption for each household in the sample obtained from our model of demand using the actual price schedule faced by that household. For each household in the simulated sample, 50 sets of demographics vectors for each household are drawn from the distribution of demographics for that household’s Zip Code. Each drawn vector of demographics for a customer is kept fixed throughout the year across all billing cycles. For each of these vectors of demographics, 50 values of ε are drawn for each billing cycle and month combination using the parameters estimates for the distribution of ε in Table 1. Using the actual price schedule faced by the household, the observed weather variables facing the household each billing cycle, the drawn vector of demographics characteristics and billing cycle-specific shocks to household's water demand, the household’s realized water demand is computed using the parameters estimates given in Table 1. For each household-billing cycle pair, we have $2500 = 50 \times 50$ simulated water consumption values for the current nonlinear price schedule. A comparison of the actual versus the simulated (at the actual price schedule) distribution of household-level billing cycle consumption provides a visual measure of the ability of the demand model to match the actual distribution of water demand across households and over time.

Figure 4 plots the distribution of the household-level “price elasticities” of demand computed as described above. These “elasticities” range from slightly above 1 in absolute value to slightly below 0.4 in absolute value, with a median of approximately 0.70. The figure also reports the aggregate “price elasticity” of 0.70, computed as described above. From Table 1, the household-level and aggregate the “income elasticity” of demand for Cobb is 0.44. Given the heterogeneity in demographic characteristics across households in Cobb, the wide range in “price elasticities” is not surprising. This result also points to the importance of knowing the structure of customer-level demand to achieving any revenue or sales goal for the utility with the least amount of revenue or sales risk. Changing the price schedule for a customer with a “price elasticity” of
-0.4 is likely to have a very different impact on the customer’s monthly consumption than the same price schedule change for a customer with -1.1 price elasticity.

The blue histogram in Figure 5 plots the expected billing cycle-level simulated household-level consumption for the actual price schedule facing the household. This figure also plots the expected billing cycle-level consumption for the price schedule that solves the following problem: Minimize the system-wide standard deviation of annual utility-level revenues from all households in our sample subject to the constraints that under the counterfactual price schedule the utility expects to sell no more water than it currently sells and it expects to earn at least as much revenue as it currently earns (as well as the constraint that the lowest counterfactual marginal price cannot be higher than the lowest actual marginal price and the highest marginal prices cannot be higher than the highest actual marginal price). The solution to this problem reduces the standard deviation of annual system-wide revenues from $14,365 to $12,685, approximately a 12 percent reduction. Figure 6 plots the actual price schedule faced by households and the counterfactual minimum standard deviation of system-wide revenue price schedule for Cobb computed by solving the optimization problem described above. Note that the counterfactual price schedule also reduces the standard deviation of system-wide water consumption from 1,449 to 1,204.

The remainder of this section derives optimal price schedules that depend on the demographic characteristics of the household. The thought experiment is that although the actual demographic characteristics of the household are still assumed to be unknown by the utility, it computes the optimal demographic characteristics-dependent price schedule given only knowledge of the distribution of demographic characteristics in each Zip Code. It is important to note that these demographic characteristics-dependent price schedules can be computed for each utility using the Zip Code-level demographic distribution data currently available in the PUMS data that is used to estimate the household-level demand model.

For the case of the number of bedrooms in the household’s dwelling, demographic characteristic-dependent price schedules are computed that minimize the standard deviation of utility-wide revenues subject to the constraints that under each demographic-specific counterfactual price schedule the utility expects to sell no more water than it currently sells and it expects to earn at least as much revenue as it currently earns as well as the constraint that the lowest counterfactual marginal price cannot be higher than the lowest actual marginal price and the highest marginal prices cannot be higher than the highest actual marginal price.
Figure 7 plots the histogram of expected household-level consumption under actual price schedule used by Cobb and the histogram of expected household-level consumption under the optimal (minimizing the standard deviation of system-wide revenues) price schedules conditional on less than 4 bedrooms in the household’s dwelling and greater than or equal to 4 bedrooms in the household’s dwelling. Figure 8(a) compares the actual price schedule to the optimal dwelling size-dependent price schedules. Figure 8(b) compares the optimal (minimum standard deviation of system-wide revenues) price schedule to the optimal dwelling size-dependent price schedules. Implementing a household-size dependent price schedule significantly reduces the standard deviation of system-wide revenues from $12,685 to $11,301, an 11 percent reduction.

Other optimal demographic variable-specific price schedules can be computed and these yield similar percentage reductions in the standard deviation of system-wide revenues. Even larger reductions in the standard deviation of the system-wide revenues of the utility are likely to be possible if the schedule could depend on combinations of elements of the vector of demographic characteristics. The optimal pricing framework used to develop the single demographic characteristic-based pricing could be extended to multi-dimensional demographic characteristic pricing. Solving this problem would require more computing time, but it is within the realm of technical feasibility and is clearly a topic for further research.

6.2. Cobb—Demographics Predicted

This section repeats the same set of optimal (minimum standard deviation of system-wide revenues) price schedule calculations as the previous section but under the assumption that the utility knows each household’s vector of demographic characteristics. To perform this analysis demographic characteristics are first assigned to each household based on what vector of demographic characteristics in that household’s Zip Code yields the highest posterior probability for the parameter estimates in Table 1 and given the household’s observed vector of water consumption and weather exposure variables. These assigned demographic characteristics are then held constant in simulating the distribution of the household’s billing cycle-level water consumption and payment obligations.

Figure 9 plots distribution of actual household-level billing cycle water consumption in thousands of gallons for all households and months in the sample period in the gray histogram. This figure also plots in blue the simulated household-level billing cycle water consumption for each household in the sample obtained from our model of demand using the actual price schedule
faced by that household. For each household in the simulated sample, the vector of demographics for each household are held constant for the entire sample period at the predicted value. For each household, 50 values of $\varepsilon$ are drawn for each billing cycle. Using the actual price schedule faced by the household, the vector of observed weather variables facing the household each billing cycle, the predicted vector demographic characteristics and billing cycle-specific shocks to household's water demand, the parameter estimates in Table 1 are used to compute the household’s realized water demand. This figure shows that standard deviation of system-wide revenues using the actual price schedule drops from $14,365 when the vector of demographic characteristics for each household is assumed to be unknown to $2,098 when the vector of demographic characteristics is assumed to be known, or a reduction of more than 85 percent.

Figure 10 plots the distribution of the household-level “price elasticities” of demand conditional on the household’s predicted demographics. This distribution has a far larger support that the distribution of “price elasticities” which does not assume the household’s demographic characteristics are known. This result occurs because the “known” demographic case conditions on an actual vector of demographic characteristics in the support of the distribution of the vector of demographic characteristics for each household, whereas the “unknown” demographic characteristics case integrates with respect to the distribution of these demographic characteristics in the household’s Zip Code. This result provides further evidence of the importance of knowledge of the customer-level demographic characteristics in achieving any pricing goal for the utility with minimum sales or revenue risk.

Figure 11 plots histogram of expected consumption under the actual price schedule and the schedule that minimizes the standard deviation of system-wide revenues. In both cases, each household’s demographic characteristics are assumed to be known and equal to the value of the vector of demographic characteristics assigned using the procedure described above. Figure 12 plots the actual price schedule in Cobb and the schedule that minimizes the standard deviation of system-wide revenues given that each household’s demographic characteristics are “known”. The optimal schedule reduces the standard deviation of system-wide revenues from $2,098 to $2065, a reduction of 1.5 percent.

The optimal demographic characteristic-dependent price schedule computation is repeated assuming that each household’s demographic characteristics are known. Figure 13 plots the histogram of expected consumption under the actual price schedule and the two price schedules
that depend on the household’s dwelling characteristics (number of bedrooms) that minimize the standard deviation of system-wide revenues. Optimal price schedules that depend on the number of bedrooms in the household’s dwelling reduces the standard deviation of system-wide revenues from 2,098 to 2,006, a reduction of 5 percent. Figure 14(a) plots the actual price schedule and the two optimal demographic characteristic-dependent price schedules. Figure 14(b) plots the optimal (minimum of the standard deviation of system-wide revenues) price schedule and the optimal number of bedroom-dependent price schedules. These results demonstrate that even if the customer’s demographic characteristics are assumed to be known, further reductions in the system-wide revenue risk are possible from choosing these demographic characteristic-dependent price schedules optimally. However, comparing the results in Figure 13 to those in Figure 7 demonstrates that the major reductions in system-wide revenue and sales risk occur from knowledge of each household’s demographic characteristics.

6.3. VoM—Demographics Drawn

This section considers a set of counterfactual price schedule choices that reflect policy goals and constraints relevant to California during the summer of 2015, the fourth consecutive summer of low water availability in the state. As a consequence, in the spring of 2015, Governor Jerry Brown issued an executive order requesting a 25 percent reduction in urban water consumption state-wide relative to 2013. A pre-existing legal constraint has further complicated the ability of municipal utilities to achieve this goal.

Proposition 218 (The Right to Vote on Taxes Initiative) requires that municipal utility consumers only pay what it costs to provide them with the water than they consume. AB 2882 (Allocation-based conservation water pricing), signed into law in 2008, attempts to clarify how nonlinear pricing of water can be implemented to avoid running afoul of Proposition 218. However, a recent lawsuit filed by customers of the municipal utility in San Juan Capistrano and the resulting decision which struck down the utility’s increasing block rate structure has led to considerable uncertainty over the use of nonlinear pricing of water in California (Stephens, 2015). Because the vast majority of the cost of urban water delivery in California (as high as 95 percent in a number of municipalities) is the fixed cost of the water storage and delivery infrastructure, determining whether prices charged are only recovering the cost of providing the service is likely to be an extremely challenging task.
One possible solution to this problem is to determine a system-wide average cost of delivering a thousand gallons of water for the utility and then set a nonlinear price schedule so that the revenues recovered from each type of household (as determined by their demographic characteristics) equals this average cost times the amount of water they consume. Because this average cost information is not available for VoM, an aggregate revenue constraint is imposed that households in the utility service territory do not pay more under the new schedule than they pay under the existing price schedule. (The constraint implicitly assumes that the utility was only recovering the cost of the water supplied under the existing schedule.) The other constraint on the counterfactual price schedule is that it reduces system-wide water consumption by 25 percent relative to expected consumption under the existing schedule with at least a 95 percent probability. The objective function for the optimal tariff design problem is to minimize the weighted sum of squared differences between each household’s expected monthly bill under the current price schedule and the household’s expected monthly bill under the counterfactual price schedule, where the weight applied to each household-level squared difference is the inverse of that household’s expected monthly bill under the current price schedule. This objective function is designed to obtain the largest revenue increases from households with the largest current water bills and the smallest revenue increases from households with the smallest current water bills. This finding this price schedule requires solving the following optimization problem:

\[
\min \theta(w) \sum_{h=1}^{H} \left[ \frac{E(R_h(\theta(w)) - E(R_h(\theta_e(w)))^2}{E(R_h(\theta_e(w)))} \right] \\
\text{subject to } \text{Prob}\left( \sum_{h=1}^{H} q_h(\theta(w)) < Q(\text{water}) \right) \geq 0.95 \,
\]

\[
E(\sum_{h=1}^{H} R_h(\theta(w)) - R_h(\theta_e(w))) \leq 0 
\]

where \( \theta(w) \) is the price schedule being solved for, \( \theta_e(w) \) is the existing price schedule, \( R_h(\theta(w)) \) is the revenue received from household \( h \) under the price schedule \( \theta(w) \), \( q_h(\theta(w)) \) is the quantity demanded by household \( h \) under the price schedule \( \theta(w) \), and \( E(\cdot) \) is the expectation operator.

There are many other possible objective functions to optimize to obtain Governor Brown’s desired 25 percent reduction in system-wide water consumption with a high probability. This one has the desirable property of putting less of the burden on households that are currently spending less on water.

Figure 15 plots the histogram of actual household-level consumption and the histogram of expected consumption under the current price schedule for VoM simulated using the parameter estimates in Table 2. Figure 16 plots the distribution of the household level “price elasticities of
demand”. The support of absolute value of these household-level “price elasticities” ranges from 0.2 to 1.2, with a median in the range of 0.7. The absolute value of the aggregate demand “price elasticity” is equal to 0.61. From Table 2, the household-level and aggregate “income elasticity” is 0.05, which is considerably smaller than the value found for Cobb. The substantial amount of household-level heterogeneity in the “price elasticities” suggests that knowledge of household-level demographic characteristics is likely to enable the utility to achieve virtually any pricing goal with significantly less risk.

Figure 17 plots that actual price schedule charged by VoM and the price schedule that solves the optimization problem described above subject to the constraint of achieving no more same expected revenues as the existing price schedule and reducing system-wide water demand by 25 percent with at least a 95 percent probability. As noted on the figure, the counterfactual price schedule reduces the monthly fixed charge by 60 percent to achieve the constraint that total system-wide revenue does not increase as a result of the tariff change. Figure 18 shows the histogram of expected consumption under the actual price schedule and the histogram of expected consumption under the counterfactual price schedule. The histogram of expected consumption under the counterfactual price schedule has significantly high probability at low levels consumption, which is consistent with it reducing the system-wide consumption by 25 percent.

6.4. VoM—Demographics Predicted

Figure 19 repeats Figure 15 assuming that household-level demographic characteristics are known and equal to the values assigned using the algorithm described above. The histogram of actual household-level consumption and the histogram on expected household-level consumption for the actual price schedule and assigned demographic characteristics are shown in Figure 19. Figure 20 plots the histogram of “price elasticities” for the case of assigned demographic characteristics. Similar to the case of Figure 10 for Cobb, the support of the histogram of household-level “price elasticities” is significantly larger than for the case in Figure 16, which is consistent with these elasticities being based on actual vectors of household-level demographic characteristics, whereas the ones in Figure 16 integrate out with respect to the distribution of these demographic characteristics.

Figure 21 plots the actual price schedule in VoM and the price schedule that minimizes the objective function described above subject to the constraints that system-wide water demand falls by 25 percent with at least 95 probability and expected total system-wide revenues do not increase.
Note that in this case the fixed charge only needs to be reduced by 31 percent to ensure that expected system-wide revenues do not increase. Comparing the standard deviation of system-wide revenues under the optimal price schedule in Figure 17 that assumes household-level demographic characteristics are unknown of $48,583 to the standard deviation of system-wide revenues under the optimal price schedule in Figure 21 that assumes household-level demographic characteristics are known of $2,568 further illustrates the massive reduction in revenue uncertainty to the utility in achieving any pricing goal from knowing their customers’ demographic characteristics. Knowledge of the customer’s demographic characteristics reduces the utility’s revenue uncertainty from optimal pricing by 95 percent.

Figure 22 plots the histogram of expected household-level consumption and the histogram of expected household-level consumption under the optimal price schedule that achieves the water savings and revenue goals in Figure 21. For both household-level expected consumption levels plotted, the household’s demographic characteristics are fixed at the value assigned using the algorithm described above. The distribution of expected consumption under the optimal price schedule is significantly more concentrated at low levels of consumption than the actual price schedule, which is consistent with it achieving the 25 percent system-wide consumption reduction with a 95 percent probability.

Proposition 218 may eventually prevent water utilities from setting different price schedules for different customers because this would result in different households paying different monthly bills for the same amount of water, even though one might argue both households have the same cost of supply. One way to achieve water pricing goals without running afoul of Proposition 218 could be for the water utility to offer all customers the same menu of price schedules and allow them to select the price schedule that best suits them. Although it would still result in different households paying different monthly bills for the same monthly water consumption, all customers would have the option to pay the same monthly bill for the same monthly consumption.

Figure 23 plots the optimal menu of two price schedules that could be offered to households in VoM that assume each household will chose the schedule that yields the highest expected utility given its household-level indirect utility function derived from the parameter estimates in Table 2. Specifically, these price schedules minimize the weighted sum of squared deviations between the customer’s consumption under the current price schedule and the price schedule they ultimately
chose subject to the constraints that system-wide demand is reduced by 25 percent with at least 95 percent probability and expected system-wide revenues does not increase relative to amount paid under the current price schedule. The actual price schedule in VoM is also plotted on Figure 23. The magenta-colored price schedule is selected by 31 percent of the households in VoM and with the green-colored price schedule selected by the remaining households. The fixed charge is reduced by 44 percent for the green-colored price schedule and by 66 percent for the magenta-colored price schedule. In this case, allowing household’s to select from a menu of price schedules increases the system-wide revenue volatility relative to a single optimal price schedule. The standard deviation of system-wide revenues increases from $2,568 to $3,153 as result of offering a menu of price schedules. However, the standard deviation of system-wide consumption falls from 391 to 359. Consequently, depending on the utility’s preferences for system-wide revenue versus quantity risk, it may find offering a menu of price schedules to be a superior approach to obtaining its water pricing goals.

Figure 24 plots the histogram of expected consumption under the current price schedule and the histogram of expected consumption under menu of two price schedules in Figure 23. In this case, the constraint that the highest priced tier had to be less than the highest price tier on the current price schedule was relaxed for both price schedules. This significantly reduced the incidence of high levels of expected consumption under the menu of price schedules solution.

6.5. Tacoma—Demographics Drawn

This section discusses the estimation results for Tacoma. Because Tacoma only has a single monthly connect charge and a single per unit price, I do not perform counterfactual price schedule calculations. Figure 25 plots the distribution of actual consumption and expected consumption under the assumption that household-level demographic characteristics are unknown for the parameter estimates in Table 3. Figure 25 plots the histogram of household-level “price elasticities” for Tacoma. These are much more concentrated and significantly more inelastic than was the case for either Cobb or VoM. The aggregate “price elasticity” is 0.19 in absolute value. However, the “income elasticity” from Table 3 is 0.67, which is considerably larger than the values for Cobb or VoM.

6.6. Tacoma—Demographics Predicted

Figure 27 computes the histogram of expected consumption in Tacoma under the current price schedule assuming each household’s demographic characteristics are those assigned by the
algorithm described above and the histogram of actual consumption in Tacoma. Figure 28 plots the histogram on household-level “price elasticities” conditional on each household’s assigned demographic characteristics. The support of these “price elasticities” expands considerably, similar to the case of Cobb and VoM.

6.7. Other Possible Counterfactual Price Schedules

This model of the household-level billing cycle level demand for water can be used to compute price schedules that achieve a wide-variety of water pricing goals. For example, a utility with a finite amount of water resources available on an annual basis because of drought conditions, may wish to design price schedules that do not have more than a five percent probability of selling more than the available water, Q(water), and achieve at least the required revenue of TR(required) with at least a probability of 95 percent, while minimizing the sum of squares of the differences between the expected revenue each household will have to pay under the new price schedules relative to the existing price schedules.

Another possible objective for managing this finite quantity of water and achieving revenue adequacy would be to minimize sum of squared deviation between the customer’s expected bill under the existing price schedule and the new price schedule weighted by a decreasing function of the household’s annual consumption during the previous year. Suppose that f(qlag, h) is a positive, monotone decreasing function in qlag, h’s consumption of water in the previous 12 months. The new optimization problem would be:

$$\min_{\theta(w)} \sum_{h=1}^{H} f(qlag, h) \left[ E(R_h(\theta(w))) - E(R_h(\theta_e(w))) \right]^2$$

subject to

$$\text{Prob} \left( \sum_{h=1}^{H} q_h(\theta(w)) > Q(water) \right) \leq 0.05$$

$$\text{Prob} \left( \sum_{h=1}^{H} R_h(\theta(w)) > TR(required) \right) \geq 0.95$$

Weighting the squared deviations by f(qlag, h) would ensure that the average deviation between household’s expected bill at the existing price schedule and the new price schedule is smaller for a household with a low value of qlag than it typically is for a household with a high value of qlag. The resulting optimal rate schedule would reward those households that consume less water in the past with a smaller bill increase than the households that consumed more water in the past year.

This model could also be used to compute probability that a utility will experience a water supply shortfall to its customer relative to some level of available water for any rate schedule. If Q(water) is the amount of available water, then this probability equal to

$$\text{Prob} \left( \sum_{h=1}^{H} q_h(\theta_e(w)) > Q(water) \right).$$
A similar calculation is possible for the probability of a total revenue shortfall. If TR(required) is the critical amount of revenue. Then the probability of a revenue shortfall for the existing price schedule, $\theta_e(w)$, is equal to:

$$\text{Prob}(\sum_{h=1}^{H} R_h(\theta_e(w)) < TR(required)).$$

Consequently, this model can be a very effective tool for managing the risk of water shortfalls during drought conditions or revenue shortfalls under all conditions.

Because the model provides the analyst with an estimate of the distribution of household-level water demand for any nonlinear price schedule and the distribution of demographic characteristics in the utility’s service territory can be computed from the PUMS data, any optimal counterfactual price schedule that relies on this information can be computed.

7. Conclusions

This model of the household-level billing cycle demand for water under increasing block pricing can be estimated from the sample billing cycle level household water bills and price schedules. Weather variables and the distribution of demographic characteristics at the Zip Code level can be used to account for the impact of weather conditions and household demographic characteristics on the demand for water.

Model of demand can be used simulate the distribution of the customer-level billing cycle level household demand for water for any increasing block price schedule. This model can then be used to simulate the distribution of the system-wide demand for water for any nonlinear price schedule. The model then be use price schedules that achieve a wide range of water supply risk or revenue risk management goals in the utility’s rate design process.

The model was used to compute two sets of counterfactual price schedules. The set first did not condition on observable demographic characteristics, and it was shown to be possible to reduce the system-wide revenue variance associated with achieving the utility’s current expected water sales and expected revenues goals. The second set of counterfactual price schedules assumed the demographic characteristics of each household was known to the utility. In this case it was also possible to reduce variance of system-wide revenues and still achieve the utility’s current expected sales and revenues goals.

An important implication of these two sets of counterfactuals was to demonstrate the tremendous reduction in revenue risk facing the utility if it is has the information on the demographic characteristics of its households. For the case of Cobb, the measure of the variance
of system-wide revenue conditional assumed knowledge of the vector of demographic characteristics was roughly 15% of the measure of the variance of system-wide revenues assuming only the distribution of demographic in each Zip Code in the utility’s service area was known.

The model was used to show that further revenue variance reductions could be achieved by demographics-based price schedules. The household-level water demand model was used to solve for the optimal (minimum system-wide revenues) demographic-based price schedules. Again, significant variance reductions were possible without be used to compute the distribution expected demand and variance in demand conditional on demographics. The model can even be used to assist the utility in managing water shortfall and potential revenue shortfalls.

The results presented here demonstrate that there is significant value to utility from understanding distribution of household level demand to design price schedules to achieve competing policy goals. In particular, by compiling demographic data from customers and using in customer-level models of demand, utilities can be used to significantly reduce the variance in both the system-wide revenues and the amount of water sold in achieving any price schedule design process. This results implies up to roughly 85% reduction in the revenue risk that the utility faces from knowledge of the demographic characteristics of its customers is possible, suggests significant economic benefits to water utilities from collecting demographic data on its customers and formulating household-level demand models for price schedule design.
References


Table 1: Model Parameter Estimates and Standard Errors—Cobb County, Georgia

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Income</td>
<td>0.4406</td>
<td>0.0129</td>
<td>0.0476</td>
</tr>
<tr>
<td>Constant Term in price elasticity formula</td>
<td>-0.4644</td>
<td>0.1332</td>
<td>0.1907</td>
</tr>
<tr>
<td>Std.dev. of household heterogeneity, ( \eta )</td>
<td>0.0203</td>
<td>0.0302</td>
<td>0.0130</td>
</tr>
<tr>
<td>Std.dev. of optimization error, ( \nu )</td>
<td>0.3615</td>
<td>0.0018</td>
<td>0.0056</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.3033</td>
<td>0.2134</td>
<td>0.3704</td>
</tr>
<tr>
<td>Average high temp in billing cycle</td>
<td>-0.0146</td>
<td>0.0019</td>
<td>0.0022</td>
</tr>
<tr>
<td>75th - 25th percentile of temperature in billing cycle</td>
<td>0.0017</td>
<td>0.0015</td>
<td>0.0014</td>
</tr>
<tr>
<td>Total precipitation in billing cycle</td>
<td>-0.0034</td>
<td>0.0079</td>
<td>0.0122</td>
</tr>
<tr>
<td>75th - 25th percentile of precipitation in billing cycle</td>
<td>-0.4129</td>
<td>0.1574</td>
<td>0.1484</td>
</tr>
<tr>
<td>Number of people over 18 in house</td>
<td>0.2241</td>
<td>0.0280</td>
<td>0.1618</td>
</tr>
<tr>
<td>Number of people under 18 in house</td>
<td>0.0472</td>
<td>0.0231</td>
<td>0.1071</td>
</tr>
<tr>
<td>House acreage above 1 acre</td>
<td>-0.2635</td>
<td>0.0353</td>
<td>0.0752</td>
</tr>
<tr>
<td>Number of bedrooms in house</td>
<td>0.6559</td>
<td>0.0339</td>
<td>0.0415</td>
</tr>
<tr>
<td>Price*temp</td>
<td>-0.0145</td>
<td>0.0008</td>
<td>0.0015</td>
</tr>
<tr>
<td>Price*precip</td>
<td>-0.0128</td>
<td>0.0037</td>
<td>0.0057</td>
</tr>
<tr>
<td>Price* # of adults</td>
<td>-0.0203</td>
<td>0.0109</td>
<td>0.0433</td>
</tr>
<tr>
<td>Price* # of children</td>
<td>0.0659</td>
<td>0.0072</td>
<td>0.0250</td>
</tr>
<tr>
<td>Price* # of bedrooms</td>
<td>0.2656</td>
<td>0.0156</td>
<td>0.0348</td>
</tr>
<tr>
<td>Number of customers</td>
<td>1034</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Price coefficient at mean of weather and demographic variables: -0.629
### Table 2: Model Parameter Estimates and Standard Errors—Valley of the Moon, California

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Income</td>
<td>0.0536</td>
<td>0.0186</td>
<td>0.0176</td>
</tr>
<tr>
<td>Constant Term in price elasticity formula</td>
<td>-0.4375</td>
<td>0.0557</td>
<td>0.1145</td>
</tr>
<tr>
<td>Std.dev. of household heterogeneity, eta</td>
<td>0.2813</td>
<td>0.0113</td>
<td>0.0290</td>
</tr>
<tr>
<td>Std.dev. of optimization error, epsilon</td>
<td>0.3031</td>
<td>0.0094</td>
<td>0.0238</td>
</tr>
<tr>
<td>Constant</td>
<td>-4.7368</td>
<td>0.1971</td>
<td>0.4994</td>
</tr>
<tr>
<td>Average high temp in billing cycle (temp)</td>
<td>0.0780</td>
<td>0.0019</td>
<td>0.0050</td>
</tr>
<tr>
<td>75th - 25th percentile of temperature in billing cycle</td>
<td>-0.0257</td>
<td>0.0021</td>
<td>0.0025</td>
</tr>
<tr>
<td>Total precipitation in billing cycle (precip)</td>
<td>-0.0055</td>
<td>0.0060</td>
<td>0.0005</td>
</tr>
<tr>
<td>75th - 25th percentile of precipitation in billing cycle</td>
<td>-0.8159</td>
<td>0.3947</td>
<td>1.4497</td>
</tr>
<tr>
<td>Number of people over 18 in house</td>
<td>0.8145</td>
<td>0.0222</td>
<td>0.0429</td>
</tr>
<tr>
<td>Number of people under 18 in house</td>
<td>0.1711</td>
<td>0.0146</td>
<td>0.0373</td>
</tr>
<tr>
<td>House acreage above 1 acre</td>
<td>0.4373</td>
<td>0.0180</td>
<td>0.0292</td>
</tr>
<tr>
<td>Number of bedrooms in house</td>
<td>-0.0137</td>
<td>0.0077</td>
<td>0.0057</td>
</tr>
<tr>
<td>Price*temp</td>
<td>0.0384</td>
<td>0.0024</td>
<td>0.0023</td>
</tr>
<tr>
<td>Price*precip</td>
<td>-0.0128</td>
<td>0.0070</td>
<td>0.0033</td>
</tr>
<tr>
<td>Price* # of adults</td>
<td>0.0572</td>
<td>0.0132</td>
<td>0.0343</td>
</tr>
<tr>
<td>Price* # of children</td>
<td>0.2946</td>
<td>0.0089</td>
<td>0.0303</td>
</tr>
<tr>
<td>Price* # of bedrooms</td>
<td>-0.5591</td>
<td>0.0157</td>
<td>0.0460</td>
</tr>
<tr>
<td>Number of customers</td>
<td>2232</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Price coefficient at mean of weather and demographic variables: -0.646
Table 3: Model Parameter Estimates and Standard Errors—Tacoma, Washington

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Income</td>
<td>0.6759</td>
<td>0.0136</td>
<td>0.0512</td>
</tr>
<tr>
<td>Constant Term in price elasticity formula</td>
<td>-1.4057</td>
<td>0.1489</td>
<td>0.1760</td>
</tr>
<tr>
<td>Std.dev. of household heterogeneity, eta</td>
<td>0.4808</td>
<td>0.0533</td>
<td>0.2355</td>
</tr>
<tr>
<td>Std.dev. of optimization error, epsilon</td>
<td>0.3643</td>
<td>0.0693</td>
<td>0.3065</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.6685</td>
<td>0.1518</td>
<td>0.3884</td>
</tr>
<tr>
<td>Average high temp in billing cycle (temp)</td>
<td>0.0164</td>
<td>0.0016</td>
<td>0.0031</td>
</tr>
<tr>
<td>75th - 25th percentile of temperature in billing cycle</td>
<td>-0.0110</td>
<td>0.0046</td>
<td>0.0036</td>
</tr>
<tr>
<td>Total precipitation in billing cycle (precip)</td>
<td>-0.0082</td>
<td>0.0026</td>
<td>0.0077</td>
</tr>
<tr>
<td>75th - 25th percentile of precipitation in billing cycle</td>
<td>-0.4318</td>
<td>0.1058</td>
<td>0.1206</td>
</tr>
<tr>
<td>Number of people over 18 in house</td>
<td>-0.2723</td>
<td>0.0272</td>
<td>0.0974</td>
</tr>
<tr>
<td>Number of people under 18 in house</td>
<td>-0.4638</td>
<td>0.0225</td>
<td>0.1704</td>
</tr>
<tr>
<td>House acreage above 1 acre</td>
<td>-0.1369</td>
<td>0.0802</td>
<td>0.1413</td>
</tr>
<tr>
<td>Number of bedrooms in house</td>
<td>-0.3215</td>
<td>0.0196</td>
<td>0.0569</td>
</tr>
<tr>
<td>Price*temp</td>
<td>0.0073</td>
<td>0.0020</td>
<td>0.0132</td>
</tr>
<tr>
<td>Price*precip</td>
<td>-0.0015</td>
<td>0.0058</td>
<td>0.0444</td>
</tr>
<tr>
<td>Price* # of adults</td>
<td>0.5237</td>
<td>0.0329</td>
<td>0.0595</td>
</tr>
<tr>
<td>Price* # of children</td>
<td>0.4503</td>
<td>0.0229</td>
<td>0.0598</td>
</tr>
<tr>
<td>Price* # of bedrooms</td>
<td>0.2989</td>
<td>0.0134</td>
<td>0.0343</td>
</tr>
<tr>
<td>Number of customers</td>
<td>2231</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Price coefficient at mean of weather and demographic variables: -0.245
Figure 3: Actual versus Simulated Consumption

Figure 4: Distribution of Price Elasticities
Figure 5: Consumption Under Optimal Prices That Do Not Depend on Consumer Demographics

Cobb: Demographics Drawn, prices not based on demographics

Prices in Cobb:
\[ \mathbb{E}[\text{Rev per bill}] = \$56.32, \text{sd}[\text{Rev system}] = \$14365, \mathbb{E}[\text{Cons per bill}] = 5.46 \text{TGAL}, \text{sd}[\text{Cons system}] = 1449 \text{TGAL} \]

Alternative prices to minimize standard deviation of systemwide revenue:
\[ \mathbb{E}[\text{Rev per bill}] = \$56.32, \text{sd}[\text{Rev system}] = \$12685, \mathbb{E}[\text{Cons per bill}] = 5.46 \text{TGAL}, \text{sd}[\text{Cons system}] = 1204 \text{TGAL} \]

Figure 6: Optimal Prices versus Actual Price Schedule

Cobb: Demographics Drawn, Prices not based on demographics

Prices in Cobb:
\[ \mathbb{E}[\text{Rev per bill}] = \$56.32, \text{sd}[\text{Rev system}] = \$14365, \mathbb{E}[\text{Cons per bill}] = 5.46 \text{TGAL}, \text{sd}[\text{Cons system}] = 1449 \text{TGAL} \]

Prices to minimize standard deviation of systemwide revenue:
\[ \mathbb{E}[\text{Rev per bill}] = \$56.32, \text{sd}[\text{Rev system}] = \$12685, \mathbb{E}[\text{Cons per bill}] = 5.46 \text{TGAL}, \text{sd}[\text{Cons system}] = 1204 \text{TGAL} \]
Figure 7: Consumption Under Optimal Prices Based on Number of Bedrooms

Prices in Cobb:
E[Rev per bill]=56.32, sd[Rev system]=14365, E[Cons per bill]=5.46TGAL, sd[Cons system]=1449TGAL

Alternative prices to minimize standard deviation of systemwide revenues:
E[Rev per bill]=56.32, sd[Rev system]=111301, E[Cons per bill]=5.46TGAL, sd[Cons system]=1301TGAL

Figure 8a: Optimal Prices Based on Number of Bedrooms versus Actual Price Schedule

Prices in Cobb:
E[Rev per bill]=56.32, sd[Rev system]=14365, E[Cons per bill]=5.46TGAL, sd[Cons system]=1449TGAL

Alternative prices to minimize standard deviation of systemwide revenues:
E[Rev per bill]=56.32, sd[Rev system]=111301, E[Cons per bill]=5.46TGAL, sd[Cons system]=1301TGAL
Figure 8b: Optimal Prices Based on Number of Bedrooms versus Non-differentiated Optimal Prices

Figure 9: Actual versus Simulated Consumption
Figure 10: Distribution of Price Elasticities

Cobb: Demographics Predicted

Prices in VOM: $E[Revenue per bill] = 55.01$, $E[Consumption per bill] = 5.34$TGPL
All prices up by 5%: $E[Revenue per bill] = 56.39$, $E[Consumption per bill] = 5.15$TGPL
Aggregate price elasticity: 0.74

Figure 11: Consumption Under Optimal Prices That Do Not Depend on Consumer Demographics

Cobb: Demographics Predicted, Minimizing standard deviation of systemwide revenues

Prices in Cobb:
$E[Rev per bill] = 55.01$, $sd[Rev system] = 2096$, $E[Cons per bill] = 5.34$TGPL, $sd[Cons system] = 211$TGPL

Alternative prices to minimize standard deviation of systemwide revenues:
$E[Rev per bill] = 55.05$, $sd[Rev system] = 2085$, $E[Cons per bill] = 5.34$TGPL, $sd[Cons system] = 212$TGPL
Figure 12: Optimal Prices versus Actual Price Schedule

Cobb: Demographics Predicted, Minimizing standard deviation of systemwide revenues

Prices in Cobb:
E[Rev per bill]=55.01, sd[Rev system]=2096, E[Cons per bill]=5.34TGAL, sd[Cons system]=211TGAL

Alternative prices to minimize standard deviation of systemwide revenues:
E[Rev per bill]=55.05, sd[Rev system]=2065, E[Cons per bill]=5.34TGAL, sd[Cons system]=212TGAL

Figure 13: Optimal Pricing Based on Number of Bedrooms

Cobb: Demographics Predicted, Bedroom Pricing

Prices in Cobb:
E[Rev per bill]=55.01, sd[Rev system]=2096, E[Cons per bill]=5.34TGAL, sd[Cons system]=211TGAL

Alternative prices to minimize standard deviation of systemwide revenues:
E[Rev per bill]=55.01, sd[Rev system]=2096, E[Cons per bill]=5.34TGAL, sd[Cons system]=2097GAL
Figure 14a: Optimal Prices Based on Number of Bedrooms versus Actual Prices

Figure 14b: Optimal Prices Based on Number of Bedrooms versus Non-differentiated Optimal Prices
Figure 15: Actual versus Simulated Consumption

Figure 16: Distribution of Price Elasticities
Figure 17: Optimal Price Schedule to Save 25% Water vs. Actual Price Schedule

Figure 18: Consumption Under Optimal Price Schedule to Save 25% Water vs Consumption Under Actual Price Schedule
Figure 19: Actual versus Simulated Consumption

Figure 20: Distribution of Price Elasticities
Figure 21: Optimal Price Schedule to Save 25% Water vs. Actual Price Schedule

VOM: Demographics Predicted, Optimal prices to reduce water consumption by 25% w.p. 0.95

Prices in VOM:
E[Rev per bill]=36.31, sd[Rev system]=2176, E[Cons per bill]=7.81T GAL, sd[Cons system]=449T GAL

Alternative prices to reduce water consumption by 25% with probability 0.95:
E[Rev per bill]=36.32, sd[Rev system]=2598, E[Cons per bill]=5.59T GAL, sd[Cons system]=391T GAL

Minimized welfare loss measure: $8.01 per bill

This price schedule involves a 0.31 discount on all fixed charges.

Figure 22: Consumption Under Optimal Price Schedule to Save 25% Water vs Consumption Under Actual Price Schedule

VOM: Demographics Predicted, Optimal prices to reduce water consumption by 25% w.p. 0.95

Prices in VOM:
E[Rev per bill]=36.31, sd[Rev system]=2176, E[Cons per bill]=7.81T GAL, sd[Cons system]=449T GAL

Alternative prices to reduce water consumption by 25% with probability 0.95:
E[Rev per bill]=36.32, sd[Rev system]=2598, E[Cons per bill]=5.59T GAL, sd[Cons system]=391T GAL

Minimized welfare loss measure: $8.01 per bill
Figure 23: Saving Water by Letting Customers to Choose from an Optimally Prepared Menu of Price Schedules:

VOM: Demographics Predicted, Optimal menu of prices to reduce water cons. by 25% w.p. 0.95

Prices in VOM:
E[Rev per bill]=363.31, sd[Rev system]=2176, E[Cons per bill]=7.81TGAL, sd[Cons system]=449TGAL
Alternative prices to reduce water consumption by 25% with probability 0.95:
E[Rev per bill]=363.58, sd[Rev system]=3153, E[Cons per bill]=5.90TGAL, sd[Cons system]=359TGAL
Minimized welfare loss measure: $11.46 per bill

31% of customers select into the magenta schedule.

Green price schedule involves a 44% discount on fixed charges.
Magenta price schedule involves a 66% discount on fixed charges.

Figure 24: Consumption Under Optimal Menu of Price Schedules
Figure 25: Actual versus Simulated Consumption

Figure 26: Distribution of Price Elasticities
Figure 27: Actual versus Simulated Consumption

Figure 28: Distribution of Price Elasticities