Residential Water Demand

Estimating Price Elasticity Under an Increasing Block Pricing System

October 17, 2006

Shanthi Nataraj[1]

Abstract

This paper investigates the price elasticity of residential water demand in the City of Santa Cruz, California using household-level data from 1989 to 2003. Residential water consumers faced an increasing two-block pricing system from 1989 to 1994, and an increasing three-block pricing system from 1995 to 2003. We estimate a conditional demand function using a two-stage-least squares (2SLS) approach, and find a price elasticity of approximately -0.8. The conditional demand model assumes that a consumer will remain in the price block in which she was observed, even if prices change. We also estimate unconditional price elasticity using a discrete-continuous choice (DCC) model, which allows consumers to switch between blocks when prices change. Unlike most previous studies, the DCC model predicts that demand is price elastic, with an unconditional elasticity estimate of -1.04. The results of the DCC model indicate that water demand (especially in the long run) may be more responsive to price changes than the current literature suggests.

I. Introduction

In recent years, many water utilities have turned to increasing block pricing (IBP) as a means of encouraging water conservation. Under an IBP system, quantity consumed is divided into "blocks"; in our study, the first 8 units of water consumed make up the first block, units 9-40 make up the second block, and units above 40 make up the third block. Users are charged a certain marginal price for water consumed within the first

[1]Ph.D. student, Department of Agricultural and Resource Economics, University of California, Berkeley, 207 Giannini Hall #3310, Berkeley, CA 94720-3310. E-mail: shanthi@are.berkeley.edu.

I would like to thank Michael Hanemann, Zhen Lei, Clair Null, Elisabeth Sadoulet, Hugo Salgado, and Alix Peterson Zwane for their valuable comments. I am indebted to Sheila Olmstead for providing the GAUSS code that served as a starting point for the DCC model programming in this paper. The Santa Cruz Water Department provided consumption and pricing data, and I would especially like to thank Toby Goddard at the Water Department for his help in understanding how the price increases occurred. Any remaining errors are mine alone.

This material is based upon work supported under a National Science Foundation Graduate Research Fellowship. Any opinions, findings, conclusions or recommendations expressed in this publication are those of the author and do not necessarily reflect the views of the National Science Foundation.

block, and subsequently higher marginal prices for water consumed in higher blocks. Consumers therefore make both a discrete choice (which block to consume in, and thereby which marginal price they face) and a continuous choice (how much water to consume, conditional on being in a certain block). Conventional econometric methods for estimating price elasticity, such as ordinary least squares (OLS), do not account for the simultaneous choice of marginal price and consumption, and thus produce biased (and often positive) results.

In this paper, we estimate price elasticity using a two-stage-least-squares (2SLS) approach, which estimates conditional water demand, and a discrete-continuous choice (DCC) maximum likelihood model, which estimates unconditional demand. The conditional demand (2SLS) model takes a consumer's choice of block, and thus her choice of marginal price, as given. The elasticity estimate, therefore, assumes that even if prices change, each consumer will remain in the same block in which she is observed. This assumption may be plausible for small price changes, and for consumers located far from the "kinks" (the points at which the price changes). However, it is less plausible for larger price changes, and for consumers near the kinks. The unconditional demand (DCC) model explicitly looks at the consumer's choice of both block and quantity. Like a discrete choice model, the unconditional demand model allows each consumer some probability of choosing each block, no matter which block she is observed in. The probability that she optimizes in any given block can therefore change when prices change. The DCC model can thus be used to answer a far richer set of policy questions than a 2SLS model. For example, it can be used to predict consumer reactions to a shift in the quantity at which a block begins, or the addition of a block to the pricing structure. It might also be used to design the optimal IBP structure for a water manager's objective, such as maximizing revenue.

This paper is one of only a handful to estimate unconditional demand for residential water, and makes the following contributions to the literature. First, we use what may be the most extensive data set in the residential water demand literature, including all households served by the Santa Cruz Water Department over a 14-year period, with variation in price levels both during and after the California drought. The cross-sectional variation allows us to estimate how household characteristics such as housing density and household size affect water use. The panel is important because it allows us to avoid the omitted variable bias generally associated with cross-sectional studies. Most previous studies of water demand use price variation across cities to identify elasticity. However, we might expect cities that are more likely to have water shortages to set higher prices, and to have more conservation-minded consumers; this correlation would cause our elasticity estimates to be biased upwards. Of the papers that do consider price changes over time within a city, many

of these study price changes that occur during droughts, which makes it difficult to separate the impacts of non-price related conservation measures (such as rationing) from price effects. In contrast, our data set includes price increases and a change in the block pricing structure after demand has stabilized following a drought, allowing us to more cleanly identify price elasticity. Second, the IBP structure in this paper affects most households. The second block of many IBP systems is set so high that the average consumer is not affected by the block structure. In contrast, the average household in our data set consumes 17 units of water every two months, and therefore falls into the second block. The unconditional model is likely to be more valuable for a data set in which most of the households are affected by the block structure. Third, of the few papers that model unconditional residential water demand, this is the only paper to fully exploit the panel nature of the data. We use the panel to validate our assumptions about potential household fixed effects in the conditional demand model before applying them in the unconditional demand model.

The rest of this paper is organized as follows: Section 2 reviews the residential water pricing literature; Section 3 presents an economic model of demand under an IBP system; Section 4 describes the data; Section 5 develops the econometric models and presents results; and Section 6 concludes.

II. Previous Literature

The literature on residential water demand is extensive, and many of the empirical studies are summarized in two meta-analyses by Espey et al. (1997) and Dalhuisen et al. (2003). Most studies have found that water demand is inelastic to price changes, and the conventional wisdom has been that moderate price increases cannot provide adequate demand management. However, IBP structures have become increasingly popular as a means of encouraging water conservation.

Block pricing structures are also found in the labor and electricity markets, and much of the literature on IBP systems originated in these two fields. One of the first issues to arise in estimating price elasticities under an IBP system is whether marginal price or average price is more relevant in explaining demand. In a seminal article on block pricing, Taylor (1975) proposes that both the marginal price of the last unit consumed, and either average price or total price for the infra-marginal units, should be included in a specification for electricity demand. A 1976 comment by Nordin argues that the theoretically correct specification should include the marginal price of the last unit consumed and a "difference" variable. The difference variable is defined as the difference between what a consumer actually paid, and what she would have paid had every unit been at the marginal price of the last unit consumed. The Taylor/Nordin specification of marginal

price and difference prompted a debate over whether consumers react to marginal price and difference, or average price. Two tests (Opaluch, 1982 and 1984; Shin, 1985) have been developed to empirically test which specification is most appropriate. Unfortunately, the test proposed by Opaluch does not account for the simultaneous choice of marginal price and quantity, and therefore produces biased results. Shin's test, which was originally developed for a decreasing block pricing system, does account for this simultaneity using an instrumental variable (IV) method; he estimates a "perceived price parameter" to test how the price perceived by the consumer relates to the average and marginal prices she faces. Nieswiedomy and Molina (1991) modify the interpretation of Shin's test for an IBP system. However, they point out that the results of the test can be inconclusive. If the consumer responds to only marginal price or only average price, then the perceived price parameter should equal 0 or 1, respectively. However, if the consumer responds to marginal price and difference, then the model makes no predictions about what the sign or magnitude of the perceived price parameter should be. This paper conducts a similar test in Section 5, but the results are inconclusive. Since the test fails to reject the Taylor/Nordin specification, we follow most of the recent water demand literature, and proceed by using marginal price and difference.

Another issue to arise in the estimation of water demand under an IBP system is the simultaneous choice of marginal price and quantity. A consumer's choice of quantity also dictates her "choice" of marginal price. Due to the nature of the IBP system, quantity and marginal price are positively correlated; therefore, regressing quantity on marginal price will produce results that are biased upwards. Several IV[2] methods have been used in an attempt to address this simultaneity bias (Deller et al. 1986). One of the more common methods is to use all of the block prices as instruments for the marginal price, and then to regress consumption on estimated marginal price in the second stage. Although IV methods can correct the simultaneity bias, they have several shortcomings. First, even if prices change, the consumer is assumed to remain within the block in which she is observed, so the estimated elasticity is conditional on the discrete choice. Second, observations at the "kinks" (points at which the price changes) must be assigned to a block or dropped. Third, and perhaps most importantly for policymakers, IV methods cannot predict the effect of changing the block pricing structure (Olmstead et al. 2005).

The DCC model was developed to explain the simultaneous choices of block and consumption in the labor literature (Burtless and Hausman 1978), and was first applied to water pricing in a Ph.D. thesis (Hewitt 1993). The theory of the DCC model is discussed in Sections 3 and 5. Only a handful of papers have previously used the DCC model to estimate residential water demand (Hewitt and Hanemann 1995; Olmstead et al.

[2]We use the term "IV" here to refer to both instrumental variables and two-stage-least-squares (2SLS) methods.

2005; Pint 1999; Rietveld et al. 1997) because the model requires household-level consumption data and is computationally expensive.

A meta-analysis of water pricing literature indicates that the DCC model may provide higher elasticity estimates than IV models (Dalhuisen et al. 2003). For example, Hewitt and Hanemann (1995) apply the DCC model to water use data in Denton, Texas, and estimate a price elasticity of approximately -1.5. A slightly different version of the data set was previously used in an IV analysis by Nieswiedomy and Molina (1989), who find that demand is price inelastic.

IBP structures are also correlated with higher elasticity estimates. In their meta-analysis, Espey et al. (1997) find that the absolute value of price elasticity is higher for cities with IBP structures than for cities with flat or decreasing block rate structures. Olmstead et al. (2005) explicitly consider this link by using data from 11 cities in the United States and Canada with flat and IBP structures. The average price elasticity for the IBP cities is -0.61, while the average price elasticity for the flat rate cities is -0.33. However, a causal link between IBP structures and higher elasticities has not been established.

As discussed above, this paper differs from the previous studies that model unconditional demand in several ways. First, none of the previous studies have included both the rich cross-sectional and time variation found in our data set. Rietveld et al. (1997) estimate elasticity based on a cross-section of 220 households, and Olmstead et al. (2005) estimate short-run elasticity using data from 1,082 households over a 1-year period. Our study avoids the typical omitted variable bias associated with cross-sectional studies, because it includes price changes within one city over time. Hewitt and Hanemann (1995) use data from 121 households over a 5-year period, and only include houses with outdoor water use. Pint (1999) has a longer and larger panel, covering 599 households over 11 years. However, the price variation in her data set occurs largely during the 1987-1992 California drought. Since she does not account for drought-related, non-price controls, such as rationing, that occur at the same time as the price increases, it is difficult to distinguish the effects of the price increases from the effects of the non-price control measures. The data set in this paper includes several price increases that occurred after demand had stabilized following the California drought, and purposefully excludes the drought years in several specifications in order to avoid potential bias from non-price measures.

Second, the average household in our data set consumes in the second block, and is therefore affected by the block pricing structure. The block structure in Hewitt and Hanemann (1995) begins at 27 CCF per month, while the structure in Pint (1999) begins at 14 CCF per month. Neither of these block rates would have affected the average household (which consumed 20 CCF and 13.5 CCF per month, respectively).

Finally, of the studies that model unconditional demand, this is the only one that fully exploits the panel nature of the data. Rietveld et al. (1997) use a cross-section. While Olmstead et al. (2005) have multiple observations from households over time, their data only cover two two-week periods during one year, and their identification is largely based on cross-sectional variation in prices. Hewitt and Hanemann (1995) and Pint (1999) have longer panels, but neither takes advantage of the panel to test for household fixed effects. Pint (1999) does include a fixed effects model; however, her fixed effects regression does not account for the simultaneous choice of marginal price and quantity, so the results are biased and reflect a positive correlation between price and consumption over much of the price range. The nature of the DCC model (specifically, the assumption of normally distributed errors) precludes the use of fixed effects in estimating unconditional demand. However, we estimate conditional demand using both random effects and fixed effects assumptions. The results of a Hausman test comparing the random and fixed effects specifications support the random effects assumptions, which we carry over to the unconditional demand model.

III. Deriving the Demand Function

This section briefly presents an economic theory of demand under an IBP system, and is largely based on Hanemann (1984), Hewitt (1993, Chapter 3) and Moffitt (1986).

Suppose that water has an increasing block pricing schedule, with K blocks (k=1,....,K) and (K-1) "kinks" (quantities at which the price increases), as follows:

$$C(Q) = \begin{cases} P_1 Q & 0 < Q < Q_1 \\ P_1 Q_1 + P_2(Q - Q_1) & Q_1 \leq Q < Q_2 \\ ... & ... \\ \sum_{j=1}^{K-1} P_j(Q_j - Q_{j-1}) + P_K(Q - Q_{K-1}) & Q \geq Q_{K-1} \end{cases} \tag{1}$$

where
Q=quantity consumed
$C(Q)$=total cost of quantity Q
Q_k=quantity at kink k
P_k=marginal price in block k

We can rewrite the payment schedule in terms of the difference variable, which was introduced in Section 2. Each block or kink determines part of the budget constraint, so we denote consumption in block or kink k as consumption in budget subset k. Conditional on locating in budget subset k, the payment schedule is

$$C_k(Q) = \sum_{j=1}^{k-1} P_j(Q_j - Q_{j-1}) + P_k(Q - Q_{k-1}) = \sum_{j=1}^{k-1}(P_j - P_{j+1})Q_j + P_kQ \tag{2}$$

Note that the left hand side of the above equation is the total amount paid by the consumer, while the second term on the right hand side is the amount the consumer would have paid had all units been priced at the marginal price of the last unit consumed. We define the difference variable for budget subset k as

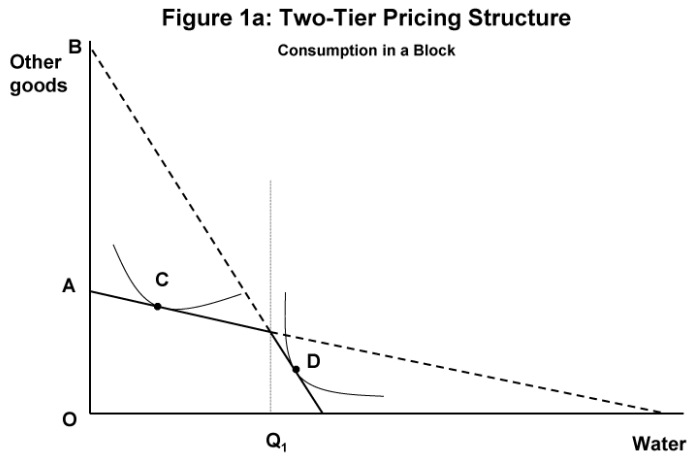
$$D_k \quad = \quad P_kQ - C_k(Q) = \sum_{j=1}^{k-1}(P_{j+1} - P_j)Q_j \tag{3}$$

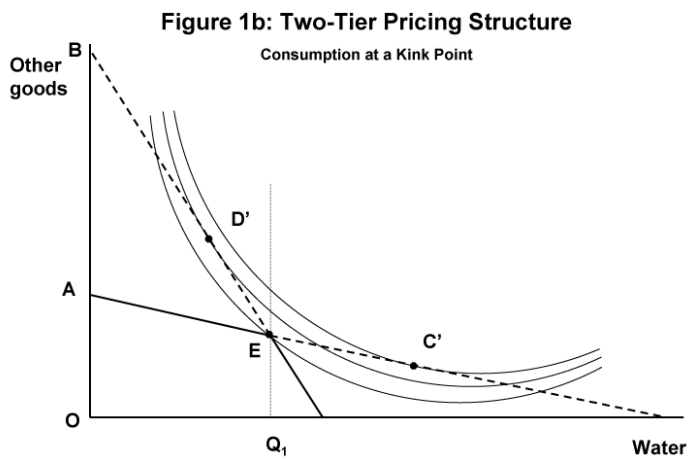
We can now rewrite the conditional payment schedule as

$$C_k(Q) = P_kQ - D_k \tag{4}$$

The difference variable acts as a lump-sum change in income, and we define virtual income to be the sum of actual income (Y) and difference (D_k) for budget subset k.

Figure 1a shows an example of a budget constraint for a 2-block system. The actual budget constraint is shown in dark lines. The kink (Q_1) occurs at the consumption level where the marginal price increases. The slope of the dark line below Q_1 is based on the (lower) marginal price in block 1, while the slope of the dark line above Q_1 is based on the (higher) marginal price in block 2. The dashed lines extend the prices for blocks 1 and 2 beyond the quantities at which they are relevant, for illustrative purposes. Income can be seen graphically as the vertical distance OA. The difference variable for a consumer in block 2 is represented by AB, and virtual income is given by OB.





The consumer optimizes over water and a composite good. Normalizing the price of the composite good to one, and assuming the budget constraint is binding, the consumer's utility-maximization problem is

$$\max U(Q, X) \text{ subject to } C(Q) + X = Y \tag{5}$$

where
$Q =$ quantity of water consumed
$C(Q) =$total cost of water
$X =$composite of other goods
$Y =$income

As illustrated in Figure 1, the consumer still faces a convex, though oddly shaped, budget constraint. As usual, she optimizes where her indifference curve is tangent to the budget constraint. In order to represent

her optimization problem in a mathematically tractable manner, though, we break down her decision into a discrete and a continuous choice. The discrete choice is which budget subset to locate in, and the continuous choice is how much water to consume, given a budget subset. We first consider the continuous choice, conditional on the discrete choice. Suppose the consumer has chosen budget subset k, and thus faces the price P_k and the virtual income $Y + D_k$. The utility maximization then becomes:

$$\max U(Q, X) \text{ subject to } P_k Q + X = Y + D_k \tag{6}$$

Note that we have not constrained the conditional demand to be in the budget subset k where the price P_k and the virtual income $Y + D_k$ are relevant. Therefore, solving this utility maximization problem yields an unconstrained (and therefore possibly infeasible) conditional demand function:

$$\tilde{Q_k} = \begin{cases} \tilde{Q}(P_k, Y + D_k; \theta) & \text{for block } k \\ Q_k & \text{for kink point } k \end{cases} \tag{7}$$

where θ represents other characteristics that affect demand, such as demographic or weather variables.

We now turn to the discrete choice. Each of the conditional demand functions has an associated conditional indirect utility function:

$$\tilde{V_k} = \begin{cases} \tilde{V}(P_k, Y + D_k; \theta) & \text{for block } k \\ U(Q_k, Y + D_k - P_k Q_k) & \text{for kink point } k \end{cases} \tag{8}$$

The consumer chooses to consume in budget subset k if $\tilde{V_k} > \tilde{V_j} \, \forall j \neq k$ and if the conditional demand is feasible (i.e., if the conditional demand for budget subset k, $\tilde{Q_k}$, falls within the range where the price P_k and the virtual income $Y + D_k$ are relevant). Moffitt (1986) shows that we can combine the discrete and continuous choices to get the constrained (feasible) conditional demand for budget subset k:

$$Q_k = \begin{cases} Q_{k-1} & \text{if } \tilde{Q}(P_{k,} Y + D_k; \theta) < Q_{k-1} \\ Q_k & \text{if } \tilde{Q}(P_{k,} Y + D_k; \theta) > Q_k \\ \tilde{Q}(P_k, Y + D_k; \theta) & \text{otherwise} \end{cases} \tag{9}$$

Combining the feasible conditional demands for each of the k budget subsets, we get the unconditional demand function (Hewitt 1993):

$$Q = \begin{cases} \tilde{Q}(P_1, Y + D_1; \theta) & \text{if } \tilde{Q}(P_1, Y + D_1; \theta) < Q_1 \\ Q_1 & \text{if } \tilde{Q}(P_2, Y + D_2; \theta) < Q_1 < \tilde{Q}(P_1, Y + D_1; \theta) \\ \tilde{Q}(P_2, Y + D_2; \theta) & \text{if } Q_1 < \tilde{Q}(P_2, Y + D_2; \theta) < Q_2 \\ ... & ... \\ \tilde{Q}(P_K, Y + D_K; \theta) & \text{if } \tilde{Q}(P_K, Y + D_K; \theta) > Q_K \end{cases} \tag{10}$$

The above derivation can be illustrated using the two-block IBP system shown on Figures 1a and 1b. Consider a consumer faced with the price of block 1 for all units of water, not just those below Q_1. Her budget constraint is represented by the dark line below Q_1 (where there price of block 1 is the marginal price) and the dashed line above Q_1 (where a higher marginal price is actually in effect). She will choose a quantity in block 1 (point C on Figure 1a) if the usual optimum conditions are met at a point below Q_1. Similarly, a consumer faced with the price of block 2 for all units of water will choose a quantity in block 2 when the usual optimum conditions are met at a point above Q_1 (point D on Figure 1a). Figure 1b shows what happens when neither of these conditions is met. If the optimal consumption at the lower price (point C' on Figure 1b) is greater than Q_1, and the optimal consumption at the higher price (point D' on Figure 1b) is less than Q_1, the consumer will choose to consume at point E, the kink (Moffitt 1986).

IV. Data

The primary data set consists of bi-monthly, household-level water consumption and price data for all single-family households served by the Santa Cruz Water Department from December 1989 to November 2003. The primary data set is combined with demographic and weather data.

Consumption and Price Data

The Santa Cruz Water Department served 18,402 single-family households between December 1989 and November 2003. Approximately two-thirds of the single family households live in the City of Santa Cruz, while the remaining one-third live in surrounding suburbs and unincorporated areas. The Water Department charges higher rates to customers outside the City. It provides combined water, sewer, and garbage services to residents inside the City, but only water services to those outside. Since sewer rates are flat charges (not based on volume), we do not consider them in this study. Households are billed every two months; approximately one-half of the households are billed in any given month.

There were a substantial number of zero values for consumption (89,150 out of nearly 1.5 million observations, or almost 6%). According to the water conservation manager at the Water Department, a consumption

value of zero indicates that no one was present in the house at the time. Some families only live in Santa Cruz during certain months of the year, or the house may have been unoccupied while on the market, or not yet constructed (Goddard 2006). Therefore, we drop observations that had zero consumption values.

Although the data allow us to construct a panel of water use for a given house over time, we are not able to identify whether the same family lived at the same address over the entire period. Our assumptions about unobserved household heterogeneity (discussed in detail in Section 5) may be less appropriate if different families lived in the house during the period of interest. To investigate the potential extent of this problem, we turn to block-group-level data[3] from the 2000 Census. The Census indicates that between 57% and 91% of households in the relevant block groups lived in the same house in 2000 as they did in 1995. Among renters, the average move-in date was 1995. However, we might expect that most single-family homes are owned; among homeowners, the average move-in date was 1987. Although these data are not conclusive, they imply that the average single-family homeowner in Santa Cruz did not change residence over the period of interest. We therefore treat the data as a panel of households, ignoring the fact that the occupants may have changed over time.

From 1989 to 1994, Santa Cruz had a 2-block pricing structure for water. Customers were billed bi-monthly, and there was a fixed charge. The marginal price in block 1 (units 1-8, where 1 unit=100 cubic feet=1 CCF) was lower than the marginal price in block 2 (units 9 and above). Prices for both blocks, as well as the fixed charge, were increased in 1991, 1993, and 1995. In 1995, a third block was introduced for revenue reasons, and to encourage water conservation. The kink between blocks 1 and 2 remained at 8 units, and the third block started at 40 units. The marginal price in block 3 was approximately twice the marginal price in block 2. The rate increase was part of a three-year plan, which also included increases in the prices of all three blocks in 1996 and 1997. There were no nominal price changes between between 1998 and 2003.

In terms of price variation, there is a trade-off to be made when using panel versus cross-sectional data. On one hand, the price variation in cross-sectional data is likely to be much greater. Even though an individual city may change its prices, the variation over time is not likely to be as great as the variation across cities. In our data set, the price of block 1 varies by 36%, while the price of block 2 varies by 76%. Block 3 is introduced, and the price varies by approximately 22% afterwards. We might therefore be concerned that our identification is based on a limited variation in price. On the other hand, using a panel from one city allows us to solve the usual endogeneity problem that plagues cross-sectional studies. We would expect cities that

[3] A Census block is a cluster of households. A block group includes all blocks that start with the same 4-digit Census identification number. Each relevant 2000 Census block includes (on average) 20 households in our sample, while each relevant 2000 Census block group includes (on average) 160 households in our sample.

are more conservation-oriented, due to weather or other factors, to have higher prices; we would also expect citizens in those cities to adopt water-saving measures. Interpreting the variation in demand as caused by the variation in prices across cities is therefore likely to overstate price elasticity. We argue that the ability to avoid such endogeneity by using a panel from one city outweighs the drawback of having a lower degree of price variation.

The bi-monthly consumption and price data are summarized in Table 1. Prices are deflated to January 1990 US$ using the Bureau of Labor Statistics' Consumer Price Index (CPI) for the San Francisco-Oakland-San-Jose area, the nearest large metropolitan area to Santa Cruz (Bureau of Labor Statistics 2006). Table 1 also includes a variable called "jurisdiction," which takes on a value of 1 if the household is inside the City of Santa Cruz, and 0 if it is outside.

Demographic Data

Using ArcInfo Geographic Information System^{TM} (GIS) software and TIGER/Line® maps, each household was spatially matched, or geocoded, to the appropriate 1990 and 2000 Census blocks and block groups. Out of 18,402 households, 265 could not be located and were dropped from the sample, leaving a total of 18,137 households in the sample[4].

The Census data in this paper are from two types of surveys: Summary File 1 (SF1) and Summary File 3 (SF3). The SF1 survey covers 100% of the population and includes basic information such as age, race, and household size. The SF3 survey uses a 1-in-6 sample, and contains more detailed demographic and housing information. For the 2000 Census, SF1 data are available at the block level, while SF3 data are available at the block group level. The 1990 Census provides both SF1 and SF3 data at the block group level. The 18,137 households belong to 951 blocks and 114 block groups in the 2000 census, and 121 block groups in the 1990 Census. Since the consumption data span a period from 1989-2003, we use the average of the 1990 and 2000 Census values. Table 1 summarizes the following Census variables:

Number of Residents. We include the median number of people in a household and expect, a priori, that larger households will consume more water.

Age of Residents. We include the median age, as families at different stages of life (e.g., families with small children vs. single people vs. the elderly) may have different water use patterns.

Household income. We construct bi-monthly household income by dividing median annual income by six. Assuming water is a normal good, we expect income to be positively correlated with water use.

[4]Comparing water consumption of the 265 dropped households to water consumption of the remaining households did not indicate a significant difference between the two groups.

House age. We construct this variable by subtracting the median year built from 1995[5]. Modern houses may use less water due to newer pipes and plumbing systems; however, modern houses may also have larger appliances that encourage more water use.

Number of rooms and bedrooms. The median numbers of rooms and bedrooms are likely to be correlated with house size and number of residents. We expect that water use will be positively correlated with the number of rooms and bedrooms.

Housing density. Large lots can significantly increase outdoor water use, especially during the summer months. However, obtaining lot size data for over 18,000 households is prohibitively expensive. Therefore, we use housing density as a proxy for lot size. We construct this variable by dividing the number of households in a block group by the area, which is estimated using ArcInfo GIS ^{TM} software.

Population density. We construct this variable by dividing the number of people in a block group by the area. Higher population densities are likely to be correlated with smaller lot sizes, indicating lower water use, as well as with larger numbers of residents, indicating higher water use.

Fraction Owned. Although we might expect higher rates of home ownership among single-family households than in the general population, we include this variable as an imperfect measure of ownership to allow renters and owners to have different water use patterns. It is constructed by dividing the number of households in a block group that are owned by the total number of households.

Weather Data

Temperature, evapotranspiration (ETo), and rainfall data were collected from the California Irrigation Management Information System (CIMIS). Data from the DeLaveaga station, located near the center of Santa Cruz, are used from September 1990 to November 2003. The DeLaveaga station was not active before September 1990; therefore, data from the nearest station, Watsonville, located about 20 miles southeast of Santa Cruz, are used from December 1989 to August 1990. We expect temperature to be positively correlated with water use. ETo is a measure of how much water is lost to the atmosphere through evaporation and transpiration. ETo and rainfall are likely to affect outdoor water use, with demand increasing in ETo and decreasing in rainfall.

We also construct a different measure of temperature, in terms of degree days, from daily temperature data. We calculate each day's contribution to the total number of degree days that month by subtracting a baseline temperature (50°F) from the average daily temperature. This measure, which accounts for particularly hot days during a month, may provide a better gauge of water needs than the monthly average

[5]The year 1995 is used because it is halfway between the two Census surveys.

Table 1: Summary Statistics

	Mean	Std. Dev.	Min	Max
Consumption and Price Variables				
Consumption (CCFs)	17.16	13.01	1	887
Block 1 Price (Jan. 1990 $)	0.60	0.09	0.43	0.77
Block 2 Price (Jan. 1990 $)	1.38	0.24	0.79	1.83
Block 3 Price* (Jan. 1990 $)	2.54	0.26	2.13	3.03
Fixed Cost (Jan. 1990 $)	14.54	3.10	8.75	20.57
Demographic Variables				
Jurisdiction (=1 within City)	0.63	0.48	0	1
No. HH Members	2.11	0.35	0.5	4
Age of All HH Residents	34.96	5.10	9.5	57
Bi-Monthly Income (thousands of $)	6.34	1.92	1.81	16.67
House Age (Years)	33.12	11.46	8	56
No. Rooms	4.76	0.78	2	7
No. Bedrooms	2.37	0.49	0.5	3
Housing Density (houses/1000 m^2)	1.08	0.64	0.001	7.12
Population Density (people/1000 m^2)	2.61	1.43	0.002	14.69
Fraction Owned	0.59	0.19	0.01	0.95
Weather Variables				
Average Temperature (degrees F)	55.75	4.83	45.9	73.43
Maximum Temperature (degrees F)	67.66	5.95	58.15	88.58
Degree Days (degrees F)	319.47	115.58	154.96	841.84
Evapotranspiration (inches)	3.49	1.39	0.99	5.94
Rainfall (inches)	2.67	2.92	0.00	14.16
No. Observations*	1,470,062			

* Block 3 did not exist prior to June 1995; therefore the number of observations is 930,304.

temperature.

Summary statistics for the weather data are presented in Table 1.

V. Model Estimation and Results

Due to the complex nature of the unconditional demand function (Equation 10), most of the water pricing literature has focused on estimating conditional demand functions. This paper begins by estimating a conditional demand function with a 2SLS procedure frequently used in previous studies. We then estimate unconditional demand using a DCC model.

Estimating the Conditional Demand Function - A Two-Stage-Least-Squares Model

We use a semi-log form for the conditional demand function for the ith household in time t:

$$Q_{kit} = P_{kit}^{\alpha} e^{\kappa + \beta D_{kit} + X_i'\gamma + W_t'\delta + \omega_1 d_y + \omega_2 d_s + \nu_{it}} \tag{11}$$

where

Q_{kit}=quantity of water demanded by household i in time t, conditional on being in the kth block
P_{kit}=marginal price faced by household i in time t, conditional on being in the kth block
D_{kit}=difference for household i in time t, conditional on being in the kth block

X_i=vector of covariates of household i that do not vary over time (e.g., household size)
W_t=vector of covariates in time t that do not vary across households (e.g., ETo)
d_y=dummy vector for year y
d_s =dummy vector for season s[6]
ν_{it}=error term for household i in time t

We begin by assuming that the error term (ν_{it}) is made up of: (1) time-invariant, household-level heterogeneity (c_i) and (2) a component that affects different households differently over time (ε_{it}). Breaking down the error into these two components, taking logs, and dropping the subscript k for notational ease, we have:

$$\ln(Q_{it}) = \kappa + \alpha \ln(P)_{it} + \beta D_{it} + X_i'\gamma + W_t'\delta + \omega_1 d_y + \omega_2 d_s + c_i + \varepsilon_{it} \tag{12}$$

We begin with an OLS analysis, which requires that $E(\varepsilon_{it}|P_{it}, D_{it}, X_i, W_t, d_y, d_s, c_i) = 0 \forall t = 1, ..., T$ and $E(c_i|P_{it}, D_{it}, X_i, W_t, d_y, d_s) = E(c_i) = 0$ for consistency. Specification (1) in Table 2 presents the OLS results. The positive coefficient on price (1.23) is significant and indicates that higher marginal prices are correlated with higher consumption. However, as discussed in Section 2, higher draws of ε_{it} are correlated with higher marginal prices. The strict exogeneity assumption is therefore violated.

To overcome this problem, we use a 2SLS method that is originally due to Wilder and Willenborg (1975), and that has been used in several previous papers on water demand (see, e.g., Hewitt and Hanemann 1995; Nieswiedomy and Molina 1989). This method uses the fixed costs and block prices set by the Water Department as instruments for marginal price and difference. The fixed costs and block prices are correlated with the marginal price and difference faced by a household. However, following the arguments of Taylor (1975) and Nordin (1976), they should not affect demand other than through marginal price and difference.

At this point, we return to the question of whether the Taylor/Nordin specification is appropriate, or if consumers react to average price, as discussed in Section 2. We follow Nieswiedomy and Molina (1991) and implement Shin's (1985) test for whether consumers react to average or marginal price under an IBP system. Our results show that the perceived price parameter is greater than one. This implies that the consumer does not react solely to marginal price; however, it is not clear whether she reacts to average price, or to marginal price and difference. We therefore continue to follow the Taylor/Nordin specification of marginal price and difference, although further work in this area may be justified.

[6]We include time dummies for each year (d_y, to allow a time trend) as well as for each season, across all years (d_s, to account for seasonality). However, due to the nature of the price changes, we cannot include a time dummy for every time period (e.g., spring 1994). This issue is discussed in more detail below.

Table 2: OLS and 2SLS Parameter Estimates

	(1) OLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
Marginal Price	1.232***	-0.292***	-0.775***	-0.764***	-0.821***	-0.809***
	(0.005)	(0.042)	(0.103)	(0.087)	(0.139)	(0.144)
Difference	0.0029***	-0.022***	-0.006	-0.007	-0.012**	-0.012**
	(0.0001)	(0.004)	(0.004)	(0.006)	(0.005)	(0.006)
Degree Days	-0.00003	0.00036	0.0002***	0.0002***	0.0002***	0.0002***
	(0.0001)	0.0004	(0.00003)	0.00004	0.00004	0.00004
Eto	0.065***	0.111***	0.122***	0.123***	0.11***	0.11***
	(0.001)	(0.003)	(0.006)	(0.007)	(0.007)	(0.007)
Rainfall	-0.004***	-0.012***	-0.011***	-0.011***	-0.014***	-0.013***
	(0.0002)	(0.0006)	(0.0006)	(0.0005)	(0.0009)	(0.0009)
Jurisdiction	0.26***	-0.093***	-0.267***	-0.228***	–	–
	(0.005)	(0.012)	(0.023)	(0.013)		
No. Residents	–	–	–	0.035**	0.028	–
				(0.017)	(0.021)	
Resident Age	–	–	–	0.006***	0.007***	–
				(0.001)	(0.001)	
Income	–	–	–	-0.0003	0.013	–
				(0.011)	(0.022)	
House Age	–	–	–	-0.004***	-0.001	–
				(0.0004)	(0.001)	
No. Rooms	–	–	–	0.038*	0.02	–
				(0.02)	(0.035)	
No. Bedrooms	–	–	–	0.009	0.048	–
				(0.021)	(0.026)	
Housing Density	–	–	–	-0.31***	-0.261***	–
				(0.02)	(0.043)	
Pop. Density	–	–	–	0.11***	0.091***	–
				(0.008)	(0.017)	
Fraction Owned	–	–	–	0.491***	0.524***	–
				(0.042)	(0.082)	
Constant	2.514***	2.049***	2.37***	1.747***	1.234***	–
	(0.007)	(0.057)	(0.051)	(0.154)	(0.138)	
Time Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Drought Excluded	No	No	Yes	Yes	Yes	Yes
City Only	No	No	No	No	Yes	Yes
Fixed Effects	No	No	No	No	No	Yes
Sample Size	1,470,062	1,470,062	1,171,347	1,171,347	742,241	742,241
F-Statistic	6,918.61	4,918.66	3,964.26	3,071.33	1,928.32	471.60
p-value	0.00	0.00	0.00	0.00	0.00	0.00

Dependent variable is log of consumption.
***, **, and * represent 1%, 5%, and 10% significance levels, respectively.
Standard errors (in parenthesis) are clustered at the household level in Specification 1. Standard errors are calculated by a non-parametric block bootstrap, which resamples across households in Specifications 2, 3, and 6, and across block groups in Specifications 4 and 5.

Table 3: 2SLS First-Stage Estimates

	(2) 2SLS		(3) 2SLS		(4) 2SLS		(5) 2SLS		(6) 2SLS	
	ln(MP)	Diff	ln(MP)	Diff	ln(MP)	Diff	ln(MP)	Diff	ln(MP)	Diff
Block 1 Price	1.390***	-15.116***	0.782***	-16.499***	0.792***	-15.813***	0.780***	-22.72***	0.78***	-22.993***
	(0.010)	(0.259)	(0.022)	(0.552)	(0.023)	(0.548)	(0.0517)	(2.2571)	(0.026)	(0.663)
Block 3 Price	-0.027***	2.238***	0.011***	1.528***	0.011***	1.503***	0.006	1.748***	0.006	1.739***
	(0.004)	(0.112)	(0.004)	(0.112)	(0.004)	(0.112)	(0.008)	(0.382)	(0.004)	(0.127)
Degree Days	0.0001***	0.007***	0.00007***	0.006***	0.00007***	0.006***	0.00007***	0.006***	0.00007***	0.006***
	(0.000005)	(0.0002)	(0.000005)	(0.0002)	(0.0002)	(0.0002)	(0.0003)	(0.0011)	(0.000007)	(0.0003)
Eto	0.025***	0.498***	0.029***	0.61***	0.029***	0.621***	0.027***	0.468***	0.027***	0.461***
	(0.001)	(0.02)	(0.001)	(0.025)	(0.001)	(0.025)	(0.002)	(0.098)	(0.001)	(0.031)
Rainfall	-0.005***	-0.031***	-0.004***	-0.016***	-0.004***	-0.009**	-0.004***	-0.019**	-0.004***	-0.023***
	(0.0001)	(0.004)	(0.0002)	(0.004)	(0.0002)	(0.004)	(0.0003)	(0.007)	(0)	(0.005)
Jurisdiction	0.078***	-0.419***	-0.072***	-0.325***	-0.064***	-0.035	–	–	–	–
	(0.0044)	(0.113)	(0.0063)	(0.1629)	(0.0067)	(0.168)				
No. Residents	–	–	–	–	0.006***	0.225	0.011	0.192	–	–
					(0.006)	(0.139)	(0.022)	(0.697)		
Resident Age	–	–	–	–	0.002***	0.066***	0.002	0.062*	–	–
					(0.00049)	(0.010)	(0.00145)	(0.03341)		
Income	–	–	–	–	-0.011***	-0.742***	0.002	-0.445**	–	–
					(0.003)	(0.06)	(0.006)	(0.213)		
House Age	–	–	–	–	-0.001***	-0.036***	-0.0002	-0.03*	–	–
					(0.0002)	(0.004)	(0.001)	(0.017)		
No. Rooms	–	–	–	–	0.021***	2.256***	0.005	1.397*	–	–
					(0.007)	(0.176)	(0.024)	(0.723)		
No. Bedrooms	–	–	–	–	0.014**	-0.08	0.013	0.025	–	–
					(0.006)	(0.141)	(0.02)	(0.461)		
Housing Density	–	–	–	–	-0.086***	-1.364***	-0.073**	-1.368**	–	–
					(0.01)	(0.188)	(0.031)	(0.584)		
Pop. Density	–	–	–	–	0.034***	0.291***	0.027**	0.215	–	–
					(0.004)	(0.085)	(0.013)	(0.292)		
Fraction Owned	–	–	–	–	0.186***	1.227***	0.164*	0.79	–	–
					(0.019)	(0.388)	(0.089)	(1.528)		
Constant	0.599***	-20.767***	0.496***	-22.472***	0.304***	-29.303***	0.195*	-31.018***	–	–
	(0.006)	(0.161)	(0.01)	(0.251)	(0.025)	(0.63)	(0.1)	(3.726)		
Time Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Drought Excluded	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City Only	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Fixed Effects	No	No	No	No	No	No	No	No	Yes	Yes
Sample Size	1,470,062	1,470,062	1,171,347	1,171,347	1,171,347	1,171,347	742,241	742,241	742,241	742,241
F-Stat (all instr.)	7,805.13	2,503.08	1,829.24	619.16	1,275.47	435.47	1,200.21	184.23	1,059.92	343.91
p-value (all instr.)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F-Stat (excluded instr.)	12,869.20	1,838.58	89.26	47.16	861.72	416.55	332.80	131.96	560.92	619.55
p-value (excluded instr.)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Dependent variables are ln(marginal price) and difference.
***, ** and * represent 1%, 5% and 10% levels of significance, respectively.
MP=Marginal price, Diff=Difference
Standard errors (in parentheses) are clustered at the household level in Specifications 1, 2, 3 and 6, and at the block group level in Specifications 4 and 5.

There are four potential excluded instruments (the price for each of the three blocks and the fixed costs) and only two endogenous regressors (marginal price and difference). Therefore, it would be ideal to use all four instruments and check the validity of the the exclusion restrictions with an over-identification test. Unfortunately, three of the instruments (the prices for blocks 1 and 2, and fixed costs) are highly correlated. When all four instruments are included, the first stage suffers from multicollinearity, and the matrix of instruments cannot be inverted. Therefore, we only use the prices for blocks 1 and 3 as excluded instruments[7].

We define a vector of instruments $Z_{it} = (P_{1t}, P_{3t}, X_i', W_t', d_y, d_s)'$ where P_{1t} and P_{3t} are the prices for blocks 1 and 3 at time t, respectively. For notational ease, let $Z_i = (Z_{i1,...,}' Z_{iT}')'$. In the first stage, we estimate $\widehat{P_{it}}$ and $\widehat{D_{it}}$ by the following regressions:

$$P_{it} = \kappa_1 + Z_{it}\Pi_1 + \zeta_{it} \text{ and } D_{it} = \kappa_2 + Z_{it}\Pi_2 + \upsilon_{it} \tag{13}$$

where we assume that $E(\zeta_{it}|Z_i) = 0$ and $E(\upsilon_{it}|Z_i) = 0 \forall t = 1, ..., T$. In the second stage, we assume that $E(\varepsilon_{it}|, Z_i, c_i) = 0 \forall t = 1, ..., T$ and $E(c_i|Z_i) = E(c_i) = 0$ and use the estimated values from the first stage regressions:

$$\ln(Q_{it}) = \kappa + \alpha \ln(\widehat{P_{it}}) + \beta \widehat{D_{it}} + X_i'\gamma + W_t'\delta + \omega_1 d_y + \omega_2 d_s + c_i + \varepsilon_{it} \tag{14}$$

Specification (2) in Table 2 shows the results of the basic 2SLS, random effects regression. Now the coefficient on marginal price is significant and negative (-0.29). Since the conditional demand is in semi-log form, this implies that a 1% increase in marginal price would produce a 0.29% decrease in consumption. All of the 2SLS regressions show first strong stage correlations between the excluded instruments and the endogenous regressors; results are presented in Table 3.

Specification (2) includes data from December 1989 to November 2003. California experienced a severe drought from 1987 to 1992. During this time, the Water Department used various voluntary and mandatory controls, in addition to price increases, to reduce water consumption. At the height of the drought in 1990, a rationing program was implemented. Under this program, a typical single family was allocated 17 CCF every two months; units above 17 CCF were subject to a \$5/unit surcharge for up to 10% over the limit, and a \$25/unit surcharge above that. In 1992, an excess use fee was charged for units over 55 CCF[8]. Other

[7] The first-stage partial F-statistics and partial R-squared values were calculated for each of the three potential instrument combinations (blocks 1 and 3, blocks 2 and 3, fixed cost and block 3). The F-statistic and R-squared values for the block 1/block 3 combination were slightly higher than the others; therefore, this combination of instruments was chosen.

[8] Although these programs used price incentives to control water use, they were not considered by the Water Department

mandatory and voluntary non-price water conservation programs were also implemented (Goddard 2006).

Total monthly consumption dropped steeply from mid-1989 to the end of 1990. Water use returned to mid-1989 levels, and began to stabilize, after the drought ended in 1992 (Maddaus Water Mgmt. 1998). The drought recovery coincided with a steep rate increase. Therefore, although much of the increase in water use during this time might be attributed to a decline in non-price drought policies, our analysis will include a positive correlation between marginal price and consumption during these years, and we would expect the coefficient on marginal price to be attenuated. Specification (3) therefore re-estimates the model using only the data from 1993-2003. As expected, the price elasticity is now more negative, at -0.78.

Specifications (4), (5), and (6) further refine the 2SLS model. Specification (4) includes Census block-group level data, correcting the standard errors for possible intra-block-group correlation. Specification (5) excludes households outside the City of Santa Cruz. Households outside the City faced a higher price than households inside the City; we might be concerned that price differences are correlated with unobserved differences between these two populations. Once we restrict ourselves to households within the City, we are no longer concerned that the excluded instruments - the prices of blocks 1 and 3 - are correlated with unobserved, time-invariant heterogeneity because all households within the City faced the same price schedule at the same time. However, we might still be concerned that some of our included instruments - specifically, the Census block-group level data on household characteristics - could be correlated with unobserved household heterogeneity. Specification (6) addresses this issue with a fixed effects model, which allows us to drop the assumption that $E(c_i|Z_i) = E(c_i) = 0$. The price elasticity estimate is robust to all of these specifications, and remains between -0.76 and -0.82. We also perform a Hausman test to compare the fixed effects estimate in Specification (6) to the random effects estimate in Specification (5)[9]. Note that Specifications (5) and (6) are the same, expect that the time-invariant demographic characteristics drop out of the fixed effects model. The Hausman test statistic is 0.00003, so we fail to reject the null hypothesis that the household fixed effects are uncorrelated with the instruments. We are therefore confident in ignoring potential fixed effects in our unconditional demand model.

In Specifications (2)-(6), the coefficients on the other parameters generally have the expected signs, except for the difference and income variables. Assuming that water is a normal good, we would expect both

to be rate increases, nor were they permanent, like the 1991 and 1993 price changes. The 1990 program, in particular, was implemented as a rationing program, not a price increase. We therefore distinguish between the 1991/1993 rate increases and the 1990/1992 fees in this paper.

[9]The random effects estimator in Specification (5) is unlikely to be fully efficient, so a simple Hausman test statistic is invalid. To account for the intra-household correlation in both models, as well as the intra-block-group correlation in Specification (5), the variance of the difference between the coefficients is estimated by a nonparametric bootstrap. Re-sampling is performed over block groups for the random effects model and over households for the fixed effects model. The usual Hausman test statistic is then calculated using the bootstrapped variance matrix (Cameron and Trivedi 2005, p. 718).

difference and income to be positively correlated with consumption. The coefficient on difference is negative in all of our 2SLS regressions, and significant in three of them. The coefficient on income is not significant in any specification, and switches sign. However, we should interpret the coefficients on these variables with caution. IBP theory predicts that the difference variable should act as a lump-sum change in income, and thus only virtual income (income plus difference) should enter the demand function. We separate the two variables in this study because it may be inappropriate to combine the difference variable, which varies by household and month, with our proxy for income, which varies by block-group and does not change over time. Since this decomposition is not entirely consistent with utility theory (Hewitt 2000), the coefficients on difference and income may not be economically meaningful.

Specifications (5) and (6) address correlations between the instruments and time-invariant, household heterogeneity. We now consider potential correlations between the instruments and unobserved, time-varying components of the error term. While our time dummies control for the average consumption in a given year, as well as the average consumption in any given season, across all years, we cannot control for an unobserved shock during every time period. Including a time dummy for every season (such as spring 1994) incorporates so much of the variation in our excluded instruments that the first stage regression is degenerate.

As discussed in the previous section, the most important series of price increases studied in this paper took place between 1995 and 1997. The rate increases for 1995, 1996, and 1997 were approved on April 1, 1995, and took place as scheduled in June of each year. The fact that the rate increases were instituted in advance by the Water Department gives us confidence that they were not imposed in reaction to unobserved factors that affected demand. However, the rate increases may still be (unintentionally) correlated with other factors that shifted demand.

The most likely candidate for such an unobserved correlation is weather. Figures 2a and 2b show that the summer of 1995 was wetter than, but about as warm as, the average summer. Part of the lower-than-average summer consumption in 1995 is likely due to lower outdoor watering needs, and we might be concerned that part of the lower summer 1995 consumption may be inappropriately attributed to the 1995 rate increase. However, in order for such a correlation to bias our results, two things would have to occur. First, the correlation between price changes and weather-induced consumption changes would have to be consistently positive or negative over time. We have no reason to think this is the case; the summers of 1996 and 1997, during which additional price increases occurred, were hotter and dryer than other summers, respectively. Second, our weather variables would have to consistently over- or underestimate the true weather. As a robustness check, we run Specifications (5) and (6) using three different measures of the temperature -

average temperature, maximum temperature, and degree-days - to see if our results change. The coefficient on marginal price varies slightly across these specifications, but is generally robust to all three measures of temperature. This result, and the robustness of our estimates in Specifications (3)-(6), give us confidence in our estimate of conditional price elasticity.

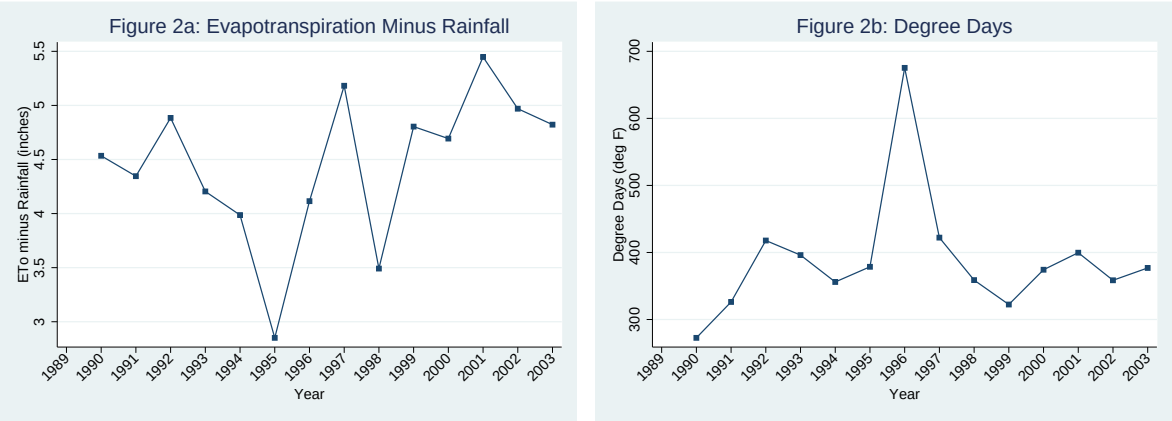


Figure 2a shows ETo minus rainfall, a measure of net outdoor water needs, averaged over the June, July and August billing periods. Figure 2b shows degree days, averaged over the same months.

Estimating the Unconditional Demand Function - A Discrete-Continuous Choice Model

In the conditional demand model, the coefficient on marginal price tells us by how much demand decreases if we increase marginal price by 1%, holding difference constant, for all consumers. However, this is not precisely what happens when we change one or more prices under an IBP system. For example, we might increase the price of the third block without changing the fixed costs or the prices of blocks one and two. Taking choice of block as given, consumers in the third block will face an increase in marginal price with difference unchanged. However, consumers in the first and second blocks will not face a marginal price change. We might attempt to simulate this policy change by only applying the marginal price increase to consumers observed in the third block and calculating the change in expected water demand. Such an estimate would still be incomplete, as it ignores the fact that consumers can switch between blocks.

An unconditional, DCC model can overcome this difficulty by introducing two separate error terms into the demand equation. Burtless and Hausman (1978) first used a two-error, DCC model to estimate the effect of taxes on labor supply. The heterogeneous preferences error term, ε, affects the optimization decision, because it represents household preferences that cannot be observed by the econometrician. It does not affect the conditional demand at the kink points because a variety of preferences support consumption there. To illustrate this point, consider Figure 1b. If the indifference curve that runs through point E were tilted slightly to the left or right, optimal consumption would still be at the kink point. However, if the

indifference curve through point D' were tilted slightly, optimal consumption would no longer occur at that point. The perception error, η, arises after the optimization decision has been made. It can occur for a variety of reasons, including measurement error in the dependent variable, such as misreading the water meter. It therefore affects all of the conditional demands, including those at the kinks (a water meter can be misread just as easily for a consumption of 40 CCF, at a kink point, as for a consumption of 20 CCF, in a block).

The ε term allows a difference between predicted and observed consumption, while the η term allows observed and actual consumption to differ. For example, we might predict, based on demographic and weather characteristics, that a household's optimal consumption is 42 CCF (in block 3). The household may have a positive draw of ε, so that it actually consumes 45 CCF (in block 3), and a negative draw of η, so that we observe it at 39 CCF (in block 2). The OLS model assigns the consumer the lower (block 2) marginal price, and does not allow any probability that the consumer was actually in block 3, and made the optimization decision while facing the higher marginal price. It therefore creates a correlation between the negative draw of $\varepsilon + \eta$ and the lower (observed) marginal price. The DCC model avoids this correlation by assigning any household, no matter what block it is observed in, a positive probability of having actually been in any other block, depending on its draw of error terms.

We add the two error terms to the unconditional demand function (Equation 10), using a 3-block structure:

$$Q = \eta + \begin{cases} \tilde{Q}(P_1, Y + D_1, \varepsilon; \theta) & \tilde{Q}(P_1, Y + D_1, \varepsilon; \theta) < Q_1 \\ Q_1 & \tilde{Q}(P_2, Y + D_2, \varepsilon; \theta) < Q_1 < \tilde{Q}(P_1, Y + D_1, \varepsilon; \theta) \\ \tilde{Q}(P_2, Y + D_2, \varepsilon; \theta) & Q_1 < \tilde{Q}(P_2, Y + D_2, \varepsilon; \theta) < Q_2 \\ Q_2 & \tilde{Q}(P_3, Y + D_3, \varepsilon; \theta) < Q_2 < \tilde{Q}(P_2, Y + D_2, \varepsilon; \theta) \\ \tilde{Q}(P_3, Y + D_3, \varepsilon; \theta) & \tilde{Q}(P_3, Y + D_3, \varepsilon; \theta) > Q_2 \end{cases} \tag{15}$$

We use the same functional form for conditional demand as we used in the 2SLS model and follow the final 2SLS, random effects specification, which excludes the drought years as well as households outside the City, and includes Census controls. Let q be the log of observed demand, q_k be the log of demand at kink point k, and $\tilde{q_k} = \ln(\tilde{Q_k}) = \kappa + \alpha \ln(P)_k + \beta D_k + X'\gamma + W'\delta + \omega_1 d_y + \omega_2 d_s$ be the log of predicted, optimal demand in block k (from now on, i and t subscripts are omitted to avoid clutter).

We also assume that ε and η enter additively, are normally distributed, and are independent of each other, so that $\eta \sim \mathcal{N}(0, \sigma_\eta^2)$ and $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. The normality assumptions preclude us from including simple

fixed effects in the DCC model. However, the results of our conditional demand model give us confidence in ignoring household fixed effects; the coefficients on marginal price for our random and fixed effects models in Specifications (5) and (6) were within 2% of each other, and a Hausman test failed to reject that the fixed effects were uncorrelated with the regressors.

Including these assumptions and re-writing the right-hand-side of Equation 15, we get:

$$q = \begin{cases} \tilde{q_1} + \varepsilon + \eta & -\infty < \varepsilon < q_1 - \tilde{q_1} \\ q_1 + \eta & q_1 - \tilde{q_1} < \varepsilon < q_1 - \tilde{q_2} \\ \tilde{q_2} + \varepsilon + \eta & q_1 - \tilde{q_2} < \varepsilon < q_2 - \tilde{q_2} \\ q_2 + \eta & q_2 - \tilde{q_2} < \varepsilon < q_2 - \tilde{q_3} \\ \tilde{q_3} + \varepsilon + \eta & q_2 - \tilde{q_3} < \varepsilon < \infty \end{cases} \tag{16}$$

The expressions in the first column after the bracket in Equation 16 (e.g., $\tilde{q_1} + \varepsilon + \eta$) represent the continuous choice, while the expressions in the second column (e.g., $-\infty < \varepsilon < q_1 - \tilde{q_1}$) represent the discrete choice. Within a block, $\varepsilon + \eta$ affects the continuous choice, while ε affects the discrete choice (see, for example, the first line in Equation 16). Therefore, the probability of observing a particular demand within that block is the probability that we observe a specific value of $\varepsilon + \eta$ and a range of values for ε. Similarly, at a kink point, η affects the continuous choice, while ε affects the discrete choice (see, for example, the second line in Equation 16), so the probability of observing demand at that point is the probability that we observe a specific value of η and a range of values for ε. We can therefore write the probability of observing a particular demand q as the sum of joint probabilities, which gives us each observation's contribution to the likelihood function:

$$\begin{aligned} \Pr(q) &= \Pr\left(\varepsilon + \eta = q - \tilde{q_1}, -\infty < \varepsilon < q_1 - \tilde{q_1}\right) + \Pr\left(\eta = q - q_1, q_1 - \tilde{q_1} < \varepsilon < q_1 - \tilde{q_2}\right) + ... \\ &\quad ... + \Pr\left(\varepsilon + \eta = q - \tilde{q_3}, q_2 - \tilde{q_3} < \varepsilon < \infty\right) \qquad (17) \\ &= \int_{-\infty}^{q_1 - \tilde{q_1}} f_{(\varepsilon+\eta,\varepsilon)}(q - \tilde{q_1}, \varepsilon) d\varepsilon + \int_{q_1 - \tilde{q_1}}^{q_1 - \tilde{q_2}} f_{(\eta,\varepsilon)}(q - q_1, \varepsilon) d\varepsilon + ... + \int_{q_2 - \tilde{q_3}}^{\infty} f_{(\varepsilon+\eta,\varepsilon)}(q - \tilde{q_3}, \varepsilon) d\varepsilon \qquad (18) \end{aligned}$$

We can use our independence and normality assumptions to simplify the joint distributions of (η, ε) and $(\varepsilon + \eta, \varepsilon)$. Since ε and η are assumed to be independent, $f(\eta, \varepsilon) = f(\eta) f(\varepsilon)$. Given our independence and normality assumptions, the joint distribution of $(\varepsilon + \eta, \varepsilon)$ is also normal, and can be factored into marginal $(\varepsilon + \eta)$ and conditional $(\varepsilon | \varepsilon + \eta)$ distributions. The complete derivation of the log-likelihood function is not

presented here; interested readers are referred to several previous publications (e.g., Hewitt 1993, Waldman 2000, Cavanaugh 2002) for more details. We define $\nu = \varepsilon + \eta$, $\sigma_\nu = \frac{\sigma_\varepsilon}{\sqrt{\sigma_\varepsilon + \sigma_\eta}}$, and $\rho = corr(\nu, \varepsilon)$, and write the log-likelihood function, summed over all households and time periods (as above, omitting i and t subscripts to avoid clutter):

$$\begin{aligned}
\ln L \;=\; & \sum_i \sum_t \ln \Bigg[\frac{1}{\sqrt{2\pi}\sigma_\nu} \exp\left(\frac{-(q-\tilde{q_1})^2}{2\sigma_\nu^2} \right) \Phi \left(\frac{\frac{q_1 - \tilde{q_1}}{\sigma_\varepsilon} - \frac{\rho(q-\tilde{q_1})}{\sigma_\nu}}{\sqrt{1-\rho^2}} \right) + \\
& \frac{1}{\sqrt{2\pi}\sigma_\eta} \exp\left(\frac{-(q-q_1)^2}{2\sigma_\eta^2} \right) \left(\Phi\left(\frac{q_1 - \tilde{q_2}}{\sigma_\varepsilon} \right) - \Phi\left(\frac{q_1 - \tilde{q_1}}{\sigma_\varepsilon} \right) \right) + \ldots \\
& \ldots + \frac{1}{\sqrt{2\pi}\sigma_\nu} \exp\left(\frac{-(q-\tilde{q_3})^2}{2\sigma_\nu^2} \right) \left(1 - \Phi\left(\frac{\frac{q_2 - \tilde{q_3}}{\sigma_\varepsilon} - \frac{\rho(q-\tilde{q_3})}{\sigma_\nu}}{\sqrt{1-\rho^2}} \right) \right) \Bigg]
\end{aligned} \tag{19}$$

The first expression inside the log is the probability of being in block 1 (0-8 CCF); the second expression is the probability of being at the first kink (8 CCF). Similar expressions for the second block and kink are not shown, while the last expression is the probability of being in the third block (40+ CCF). Starting at the 2SLS parameter estimates from Specification (5) in Table 2, we maximize this log-likelihood function using the MAXLIK routine in GAUSS[10], with the Berndt-Hall-Hall-Hausman (BHHH) descent method and the STEPBT (polynomial fitting) algorithm for determining step length. The parameter estimates are presented in Table 6[11].

Table 6 also shows that the variance of the heterogeneous preferences error, σ_ε^2, is approximately 5 times as large as the variance of the perception error, σ_η^2. This is consistent with most previous DCC studies, and implies that the unexplained variance of demand can be attributed more to household heterogeneity than to measurement error (Cavanaugh 2002; Hewitt and Hanemann 1995, Moffitt 1986).

The coefficient on marginal price (-0.92) is the price elasticity conditional on remaining in the same block. We can also estimate the unconditional price elasticity because the DCC model allows the probability of being in each block or at each kink to change when the marginal price changes. We begin by calculating the expected water demand for each household according to Equation 20. Interested readers are referred to Cavanaugh (2002) for the derivation of this equation.

[10] We would like to thank Sheila Olmstead for providing the GAUSS code that served as a starting point for our programming.

[11] Wooldridge (2002, pp. 404-410) notes that the partial maximum likelihood estimator is consistent and asymptotically normal under the standard identification and regularity conditions, even if the errors are correlated within a block group, or for a household over time. However, the standard errors must be corrected to account for the panel nature of the data, as well as the potential clustering due to the Census block-group level data. This can be accomplished by performing a non-parametric block bootstrap, which re-samples across block groups. Due to the time necessary to complete the bootstrapping procedure (the original convergence took approximately 4 hours on a shared server), the standard errors presented in Table 6 have not yet been corrected, and are likely to be too small. Future work will include correcting the errors.

Table 4: DCC Parameter Estimates

Marginal Price	-0.920***
	(0.035)
Difference	-0.014***
	(0.0004)
Temperature	0.0003***
	(0.00001)
Eto	0.117***
	(0.003)
Rainfall	-0.015***
	(0.001)
No. Residents	0.067***
	(0.004)
Resident Age	0.005***
	(0.0003)
Income	0.027***
	(0.002)
House Age	-0.002***
	(0.0001)
No. Rooms	0.041***
	(0.004)
No. Bedrooms	0.01**
	(0.004)
Housing Density	-0.27***
	(0.006)
Pop. Density	0.067***
	(0.003)
Fraction Owned	0.271***
	(0.012)
Constant	1.864***
	(0.022)
σ_η^2	0.141***
	(0.012)
σ_ε^2	0.677***
	(0.003)
Time Dummies	Yes
Drought Excluded	Yes
City Only	Yes
Sample Size	742,241
Mean Log-Likelihood	-0.672

*** and ** represent significance at the 1% and 5% levels, respectively.
Standard errors are in parentheses.

$$
\begin{aligned}
\mathbb{E}(q) \quad = \quad & \tilde{q_1} \exp(\frac{\sigma_\eta^2}{2}) \exp(\frac{\sigma_\varepsilon^2}{2}) \left[1 - \Phi \left(\sigma_\varepsilon - \left(\frac{\ln(\frac{q_1}{\tilde{q_1}})}{\sigma_\varepsilon} \right) \right) \right] + \\
& q_1 \exp(\frac{\sigma_\eta^2}{2}) \left[\Phi \left(\frac{\ln(\frac{q_1}{\tilde{q_2}})}{\sigma_\varepsilon} \right) - \Phi \left(\frac{\ln(\frac{q_1}{\tilde{q_1}})}{\sigma_\varepsilon} \right) \right] + ... \\
& ... + \tilde{q_2} \exp(\frac{\sigma_\eta^2}{2}) \exp(\frac{\sigma_\varepsilon^2}{2}) \Phi \left(\sigma_\varepsilon - \left(\frac{\ln(\frac{q_2}{\tilde{q_3}})}{\sigma_\varepsilon} \right) \right) \qquad (20)
\end{aligned}
$$

We simulate a 1% increase in the price of each block, and calculate expected water demand for each household at the new set of prices. The unconditional price elasticity (averaged across all households), given a 1% increase in the price of each block, is estimated to be -1.04.

VI. Conclusion

Both the 2SLS and DCC models indicate that long-run, conditional demand (i.e., assuming consumers do not switch between blocks) is price-inelastic, with elasticity estimates ranging from approximately -0.76 to -0.92. However, using the DCC model, we can also calculate unconditional price elasticity, allowing consumers to switch blocks when prices change. Simulating a 1% price increase in all three blocks in the DCC model yields an unconditional price elasticity estimate of -1.04.

This paper's contributions to the existing literature center around the unconditional demand model, which has seldom been estimated. The results of the DCC model are important for two main reasons. First, as discussed above, OLS estimates tend to understate price elasticity under IBP systems. Our results add to growing evidence that the usual 2SLS models may also underestimate price elasticity. Demand-side management, through price increases, may therefore be more effective than the current literature suggests.

Second, while accurately estimating price elasticity under an IBP system is important, the most powerful feature of the DCC model is the ability to simulate the effects of different policies. In Section 5, we simulated the effect of increasing the price of each block by 1%. However, this is by no means the only policy change we can consider. We might be concerned that poorer households are paying too much for basic water needs, while richer households have "wasteful" water use patterns such as watering their lawns during the day. To address this issue, we could simulate the effects of a decrease in the block 1 price and an increase in the block 3 price, or the addition of a fourth block with a much higher marginal price. Future work might also include designing optimal block pricing structures. For example, a water manager might seek maximize revenue,

or to minimize use above a certain threshold, subject to an minimum revenue constraint. As more water utilities turn to IBP systems, designing block structures that meet water managers' goals is an important direction for future research.

References

Bureau of Labor Statistics (BLS). Accessed on March 10, 2006. " Consumer Price Index - All Urban Consumers. Area: San Francisco-Oakland-San Jose, CA" <http://data.bls.gov/>.

Burtless, Gary and Jerry Hausman. 1978. "The Effect of Taxation on Labor Supply: Evaluating the Gary Negative Income Tax Experiment." *Journal of Political Economy* 86(6): 1103-1130.

California Irrigation Management Information System (CIMIS). Accessed on March 19, 2006. <http://wwwcimis.water.ca.gov>.

Cameron, Colin and Pravin Trivedi. 2005. *Microeconometrics.* New York: Cambridge University Press.

Cavanaugh, Sheila. 2002. *Essays in Environmental Economics and Policy.* Ph.D. Dissertation, Harvard University.

Dalhuisen, Jasper, Raymond Florax, Henri de Groot, and Peter Nijkamp. 2003. "Price and Income Elasticities of Water Demand: A Meta-Analysis." *Land Economics* 79(2): 292-308.

Deller, Steven, David Chicoine, and Ganapathi Ramamurthy. 1986. "Instrumental Variables Approach to Rural Water Service Demand." *Southern Economic Journal* 53(2): 333-346.

Espey, M., J. Espey and W.D. Shaw. 1997. "Price Elasticity of Residential Demand for Water: A Meta-Analysis." *Water Resources Research* 33(6):1369-1374.

Goddard, Toby. Water Conservation Manager, Santa Cruz Water Department. Phone Conversations on March 13, April 3, and June 16, 2006.

Hanemann, W. Michael. 1984. "Discrete/Continuous Models of Consumer Demand." *Econometrica* 52(3): 541-562.

Hewitt, Julie. 1993. *Watering Households: The Two-Error Discrete-Continuous Choice Model of Residential Water Demand.* Ph.D. Dissertation, U. C. Berkeley.

Hewitt, Julie and W. Michael Hanemann. 1995. "A Discrete/Continuous Choice Approach to Residential Water Demand Under Block Rate Pricing." *Land Economics* 71(2): 173-192.

Hewitt, Julie. 2000. "A Discrete/Continuous Choice Approach to Residential Water Demand Under Block Rate Pricing: Reply." *Land Economics* 76(2): 324-330.

Maddaus Water Management (in association with the Weber Group). March 1998. *City of Santa Cruz Water Demand Investigation.*

Moffitt, Robert. 1986. "The Econometrics of Piecewise-Linear Budget Constraints: A Survey and Exposition of the Maximum Likelihood Method." *Journal of Business and Economic Studies* 4(3): 317-328.

Nieswiedomy, Michael and David Molina. 1989. "Comparing Residential Water Demand Estimates under Decreasing and Increasing Block Rates Using Household Data." *Land Economics* 65(3): 280-289.

Nieswiedomy, Michael and David Molina. 1991. "A Note on Price Perception in Water Demand Models." *Land Economics* 67(3): 352-359.

Nordin, John. 1976. "A Proposed Modification of Taylor's Demand Analysis: Comment." *The Bell Journal of Economics* 7(2): 719-721.

Olmstead, Sheila, W. Michael Hanemann and Robert Stavins. 2005. "Do Consumers React to the Shape of the Supply? Water Demand Under Heterogeneous Price Structures." In Submission. <http://pantheon.yale.edu/ smc77/>

Opaluch, James J. 1982. "Urban Residential Demand for Water in the United States: Further Discussion." *Land Economics* 58(2): 225-227.

Opaluch, James J. 1984. "A Test of Consumer Demand to Water Prices: Reply." *Land Economics* 60(4): 417-421.

Pint, Ellen M. 1999. "Household Responses to Increased Water Rates During the California Drought." *Land Economics* 75(2): 246-266.

Rietveld, Piet, Jan Rouwendal, and Bert Zwart. 1997. "Estimating Water Demand in Urban Indonesia: A Maximum Likelihood Approach to Block Rate Pricing Data." Tinbergen Institute Discussion Paper No. 97-072/3. Tinbergen Institute, Amsterdam.

Shin, John. 1985. "Perception of Price When Price Information is Costly: Evidence from Residential Electricity Demand." *The Review of Economics and Statistics* 67(4): 591-598.

Taylor, Lester. 1975. "The Demand for Electricity: A Survey." *The Bell Journal of Economics* 6(1): 74-110.

U.S. Census Bureau. Census 2000 and Census 1990, Summary Files 1 and 3. Generated by the author using American Factfinder. <http://factfinder.census.gov>.

Waldman, Donald. 2000. "A Discrete/Continuous Choice Approach to Residential Water Demand Under Block Rate Pricing: Comment." *Land Economics* 76(2): 322-323.

Wilder, Ronald and John Willenborg. 1975. "Residential Demand for Electricity: A Consumer Panel Approach." *Southern Economic Journal* 42(2): 212-217.

Wooldridge, Jeffrey. 2002. *Econometric Analysis of Cross-Sectional and Panel Data.* Cambridge: The MIT Press.