



**Miguel Bachrach and William J. Vaughan**

# **Household Water Demand Estimation**



Working Paper ENP106

**Inter-American Development Bank**  
Productive Sectors and Environment Subdepartment  
Environment Protection Division

March 1994



*Household Water Demand Estimation*

*Miguel Bachrach and William J. Vaughan*

*March 1994*

*Miguel Bachrach, formerly an Operations Evaluation Officer at the Inter-American Development Bank, is now in private business. William J. Vaughan is a Senior Economist in the Project Analysis Department of the IDB.*

*This paper was originally developed for the IDB's Operations Evaluation Office as background for the ex-post evaluation of rural potable water projects. The authors thank Jorge Ducci, Christian Gómez, and Glenn Westley for their comments and suggestions. The views and opinions expressed herein are those of the authors and do not represent the official position of the Inter-American Development Bank. This paper has undergone an informal review process but has not been submitted to a formal peer review. Its purpose is to elicit reader comments and suggestions, and is primarily intended for internal use. Comments should be forwarded directly to the authors.*

**Inter-American Development Bank  
1300 New York Avenue, N.W.  
Washington, D.C. 20577**

## **Household Water Demand Estimation**

### **C O N T E N T S**

#### ***General Demand Estimation Issues***

#### ***Simultaneity, Alternative Price Measures, and Multiple Equilibria***

*Simultaneity*  
*Choice of Variables*  
*Properties of the Demand Function*

#### ***Sample Selection Bias***

*Problems with Ordinary Least Squares, even in the Absence of Simultaneity*  
*The Heckman Approach: Sample Selection Bias Correction with no Simultaneous Equations*  
*Extension of the Sample Selection Model to the Simultaneous Equations Situation*

#### ***Model Specifications and Data***

*Supply Variables*  
*Demand Variables*  
*Water Demand Equation Specification*

#### ***Estimation Results***

#### ***Conclusions***

*Simultaneous Equation Bias Exists*  
*Average Price May not be Relevant*  
*Sample Selection Bias May Matter*  
*What to Do?*

#### ***References***

## ***Household Water Demand Estimation***

To estimate the benefits of a potable water supply project, some idea of the parameters of the demand function is needed to calculate a Marshallian consumer's surplus welfare measure. Often, given a baseline estimate of pre-project consumption and price, an elasticity estimate can be used to reconstitute the underlying demand function and the consumer surplus integral (see Powers and Valencia 1980), but the elasticity has to be obtained from the econometric estimation of a demand relation. In an early review of twenty-six household water demand elasticity estimates published before 1978, Hanke (1978) reported a wide range, the lowest estimate being -0.01 and the highest -6.71. Gómez (1987) reports a narrower range from the Bank's own work in Latin America and the Caribbean, with all estimates below -0.65.

Any estimate of the benefits of a reduction in the price of water due to a water supply project arrived at using an elasticity-based calculation will be sensitive to the magnitude of that elasticity. So will the benefits based on a method that uses estimated demand functions directly. This paper discusses the relative merits of the statistical procedures that can be used to obtain elasticity estimates or their underlying water demand functions.

### ***General Demand Estimation Issues***

A review of the water demand literature indicates that most problems encountered in demand estimation and the techniques available to address them are derived from or have been dealt with in the literature on electricity demand. In both cases, the issues discussed most frequently have been:

- 1) Whether there is a simultaneous equations problem when estimating demand for water with multipart rate schedules and, if there is one, what techniques should be used to correct for it?
- 2) Whether average or marginal price is the relevant measure in estimating the demand function.<sup>1</sup>

In addition to these problems, another issue that has historically not received as much attention in the water demand literature is how to deal with sample selection bias in the face of a rate schedule combining a fixed charge for consumption below the level where a block rate tariff per unit consumed takes effect. Accounting for this phenomenon in statistical estimation, possibly in combination with (1) and (2) above, essentially involves deciding how to handle observations in the fixed charge category, where no unit consumption price is observed.<sup>2</sup>

---

<sup>1</sup> Another issue is whether the presence of declining block rates invalidates the estimation of individual demand curves because of the possibility of multiple equilibria. With micro data on individuals facing decreasing block rate tariffs, a continuous, unique demand schedule may not exist.

<sup>2</sup> In the application to rural localities in Argentina (see below), estimation with this data is complicated by the fact that a large number of customers fall in the fixed charge, minimum consumption block. It would not be uncommon to expect this in localities serving relatively poor households having few

The following two sections review each of these problems. The next section explains the data setup, after which the results of estimating water demand functions using a sample of 685 families from 34 localities in rural Argentina taken in 1987 are presented and discussed. The results obtained with the different estimation techniques designed to address the issues identified above are compared, and some conclusions are drawn in a final section.

### ***Simultaneity, Alternative Price Measures, and Multiple Equilibria***

#### *Simultaneity*

Until recently, the water demand literature has paid limited attention to simultaneity problems. It is worthwhile discussing this in some detail since the presence of, and the different solutions to, simultaneity have direct implications for the method of demand estimation.

A necessary condition for unbiased and consistent parameter estimation under Ordinary Least Squares (OLS) is that there be no correlation between the error term and any of the explanatory variables. With multipart block rates this condition is not fulfilled, since prices are endogenously determined by the consumer along with quantity demanded. Even though the rate schedule (the supply function) is knowable a-priori by the investigator, the consumer chooses the preferred price-quantity pair, implying that both price and quantity are endogenous, not just quantity subject to a fixed price. Or, said otherwise, unless price is infinitely elastic, a simultaneous equations problem theoretically must exist in the face of block rate schedules even if the parameters of the latter are known a-priori (Westley 1984). To illustrate this problem let us start with a simple demand function of the form:

$$Q = \alpha + \beta_1 Y + \beta_2 P + u \quad (1)$$

where:

Q	=	quantity demanded
$\alpha, \beta$	=	parameters
Y	=	income
P	=	price
u	=	random error term

If we were estimating the demand for domestic appliances by individual households, for example, it is unlikely that purchases made by any household would affect the market price. At this micro level, market prices are not influenced by (are independent of) the quantity demanded by a single household, so the error term (u) in Eq. (1) is uncorrelated with P.

In contrast, the price of water chosen by any household will be correlated with the random error term in Eq. (1) if the supplier's water tariff schedule relates the price charged to the quantity a household purchases.

---

expensive water-using appliances. If marginal price for this group is arbitrarily set to zero in estimation the slope of the demand curve will be biased. Removing these observations from the estimation sample introduces sample selection bias.

Suppose the relevant price for the consumer is the average price<sup>3</sup> of water, and the price depends in stepwise block fashion on the amount consumed:

$$P = [(P_f + r_1(q_1 - q_f) + r_i(q_{i+1} - q_i) + \dots + r_j(Q - q_{j-1}))]/Q \quad (2)$$

where:

- $q_f$  = quantity available for consumption at a fixed charge
- $P_f$  = fixed charge for water consumption at or below  $q_f$
- $r_i$  = block rate for block  $i$  (where  $j$  is the last consumption block)
- $q_i$  = block interval for block  $i$ .
- $P$  = average price
- $Q$  = total quantity consumed.

By substituting Eq. (1) in Eq. (2) it is easy to see how the errors from Eq. (1) are correlated with  $P$ . With three blocks, Eq. (2) can be written as:

$$P = \frac{P_f + r_1(q_1 - q_f) + r_2(q_2 - q_1) - r_3q_2}{Q} + r_3 \quad (3)$$

Substituting the expression for  $Q$  from Eq.(1) in the denominator of (3) we have:

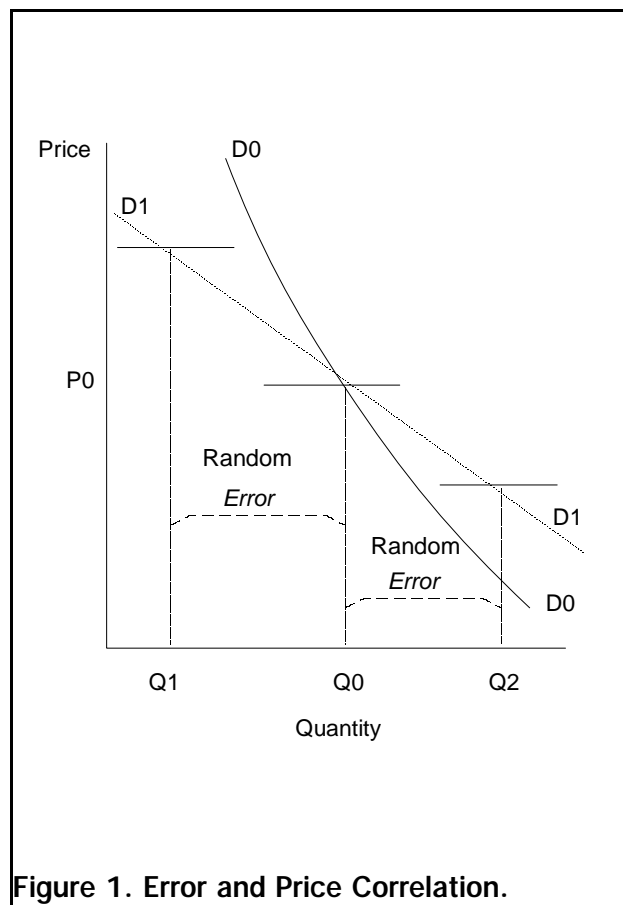
$$P = \frac{P_f + r_1(q_1 - q_f) + r_2(q_2 - q_1) - r_3q_2}{\beta_0 + \beta_1 Y + \beta_2 P + u} + r_3 \quad (4)$$

The error term ( $u$ ) in the denominator of Eq. (4) is correlated with the average price, so bias in all parameter estimates of  $\beta_i$  will be introduced if  $P$  is used as an independent variable to estimate Eq. (1). The larger the error term ( $u$ ) the smaller the average price, since the consumption level will be in one of the blocks with lower marginal rates.

---

<sup>3</sup> This is an intuitively appealing assumption. The paradigm of the ideal consumer who is fully aware of all the rates in the water tariff schedule and exhibits optimizing behavior by equating marginal willingness to pay to marginal supply price in a particular block ignores the cost of acquiring information about tariff rates for a commodity that usually commands a small share of total household expenditure. Although the approximating behavior embodied in the concept of average rather than marginal price makes intuitive sense, it cannot be accepted with our Argentinean data set (see the Results section). The water demand literature has debated this issue at length. For a recent and particularly good explanation, see Nieswiadomy and Molina (1991).

Even if the quantity demanded is a function of the marginal price of water instead of the average price, the same simultaneity bias problem can exist. In Figure 1 the real (and unobserved) household demand schedule  $D_0$  intersects the supply schedule in the block where the price is  $P_0$ . With income constant, suppose an exogenous random event like a change in weather causes a change in the quantity demanded. A random increase in water consumption would be observed as, say,  $Q_2$  and a random decrease would be observed as  $Q_1$ . The error term ( $u$ ) is negatively correlated with price ( $P$ ), since a large and positive random error ( $u$ ) reduces the observed  $P$  and a large negative error increases the observed  $P$ . The OLS assumption of independence between the error term and the explanatory variables is violated because the water price is not a constant--it depends on the quantity chosen. As a consequence OLS estimation will erroneously yield the parameters of  $D_1$  instead of  $D_0$ . In other words, using OLS with a declining block schedule will underestimate or overestimate demand elasticity depending on whether the supply schedule is steeper than the demand schedule. With a rising block schedule the price response would be overestimated.



**Figure 1. Error and Price Correlation.**

In the face of simultaneity, instrumental variables (IV) estimators such as two stage least squares (2SLS) are preferable to ordinary least squares to produce consistent parameter estimates. Curiously there is no consensus in the water demand literature on this issue. While the presence of simultaneous equation bias and the need for instrumental variables estimation could be resolved via a Hausman test on a case by case basis (Hausman 1978, Nakamura and Nakamura 1981) empirical tests for simultaneity bias are uncommon in the water demand literature.<sup>4</sup> The use of OLS has often been justified a-priori on the basis of intuition, custom, or (to us) apparently specious reasoning, without bothering to test for the validity of the OLS error assumptions.<sup>5</sup>

Another way to understand demand function identification is illustrated in Figure 2. There the three demand schedules represent the demand of three families with identical characteristics except for income (which rises from  $D_0$  to  $D_2$ ). The decreasing block rate supply schedule

<sup>4</sup> For an exception, see Nieswiadomy and Molina (1991) who reject the OLS assumption of price exogeneity using the Hausman test.

<sup>5</sup> A review of water demand analyses performed for IDB projects shows that OLS rather than IV estimation has always been used. Chicone and Ramamurthy (1986) make an argument for OLS based on the evidence from previous studies.



represents the tariff regime in the locality ( $S_0$ ). With this information it is not possible to identify the demand function since the only data available are the three points of intersection between the three separate demand schedules and the three block rates representing the segments of the single public utility rate (supply) schedule. With the information in Figure 2, OLS can only produce a linear approximation of the discontinuous step function for supply; the parameters of the demand function are not identifiable.

In order to distinguish supply from demand we need some variation in the supply schedule that can be accounted for by an independent variable that does not affect the demand schedule. This condition is usually obtained by using more than one rate schedule (as in Figure 3). These can be constructed using time series information on rates for a single community, a cross-section of different communities, or both.

An identifiable demand function is depicted in Figure 3. The several supply schedules  $S_0$  to  $S_2$  clearly trace out the several demand schedules.<sup>6</sup>

The problem of simultaneity in the presence of a block rate structure has been dealt with in three ways. The first solution is to apply two-stage least squares, as in Halvorsen (1975) or Westley (1984). For example, in a system like:

$$Q = \alpha + \beta_1 Y + \beta_2 P + u \quad (5)$$

$$P = f(R(l,t),Q) \quad (6)$$

where: R = rate schedule  
l = locality  
t = time

one can estimate the structural coefficients of (5) by regressing marginal price, P, against all the independent variables of the system, and then using the prediction of P ( $\hat{P}$ ) in equation (5). An important requirement for the applicability of this method is to have a known set of rate schedules with sufficient variation either in time period or by locality.<sup>7</sup>

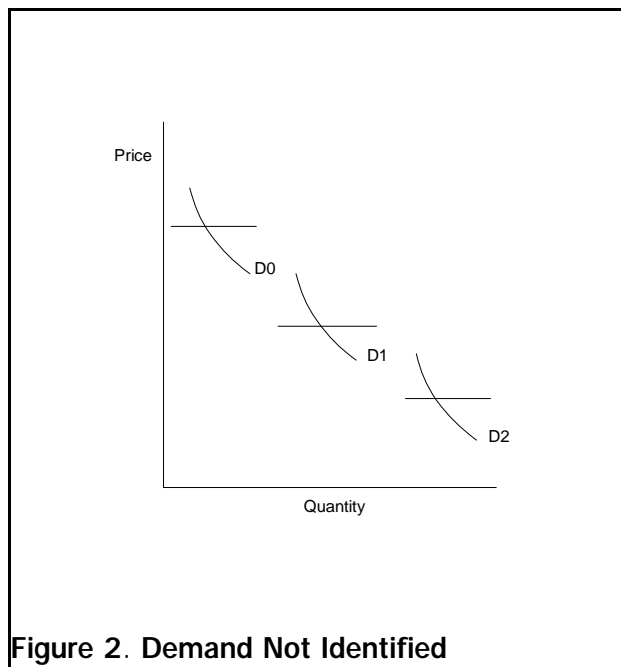


Figure 2. Demand Not Identified

<sup>6</sup> This assumes that the price slope of the demand function is constant over time or across localities, respectively. Otherwise the demand function cannot be identified.

<sup>7</sup> Having sufficient variation may be more complicated than it seems. In our sample we had some multicollinearity problems when constructing representations of the rate schedule using Westley's method.

The second way to solve the simultaneity problem is by applying fixed point or iterative least squares methods (see Martin et.al., 1984). These techniques are similar to 2SLS, and basically consist of estimating (5) and (6) independently and using the predicted values of Q and P ( $\hat{Q}$  and  $\hat{P}$ ) to reestimate (5) and (6) in a second iteration. This process is repeated until the values of  $\hat{Q}$  and  $\hat{P}$  converge.<sup>8</sup> This technique was not applied to our survey because its statistical properties are less well known than 2SLS.

The third solution actually argues that there is no real simultaneity if the problem is properly specified (Blattenberger 1977, and Taylor, Blattenberger and Rennhack 1981 and 1982). It is argued that demand and supply relations can be derived and estimated from a reduced form. In this formulation the consumer makes a decision based on the entire rate schedule; viz:

$$Q = \alpha + \beta_0 * Y + \beta_1 * [R(l,t)] + u \quad (7)$$

The schedule  $R(l,t)$  is exogenous to the amount purchased, and in a formulation like Eq. (7) there is no apparent simultaneity problem. While the rate schedule can be represented in many ways to estimate (7), its economic meaning is not entirely clear. Moreover, most proxies that could represent the rate schedule will be correlated with the amount of water consumed. In particular, the two proxies normally chosen are the average price of water ( $P_{avg}$ ) or the marginal price of water ( $P_{marg}$ ), which will be correlated to the quantity of water consumed. A solution to this problem is to create an artificial approximation to the rate schedule and derive average or marginal prices from it. Two approximations proposed by Taylor, Blattenberger and Rennhack 1981 (TBR) in the case of electricity can be directly applied to water demand estimation.

The first creates a linear approximation to the electricity bill ( $B_i$ ) and derives marginal prices from it. In the case of TBR, who work with aggregate data, the first step was to take the electricity bill in each locality at discrete consumption intervals and aggregate each interval by state.<sup>9</sup> For example, if the first interval is from 0 to 100 Kwh/month, and there are 50 localities in each state, the first observation,  $B_{(1)}$ , would represent the weighted average total cost of water in the 100 Kwh/month interval for the 50 localities. Assuming there are K different intervals, the second step is to regress the  $B_k$  "observations" against consumption, or rather, against each particular interval limit ( $Q_k$ ).<sup>10</sup> That is, if  $B_{(k)}$  is the average bill for a particular block they regress:

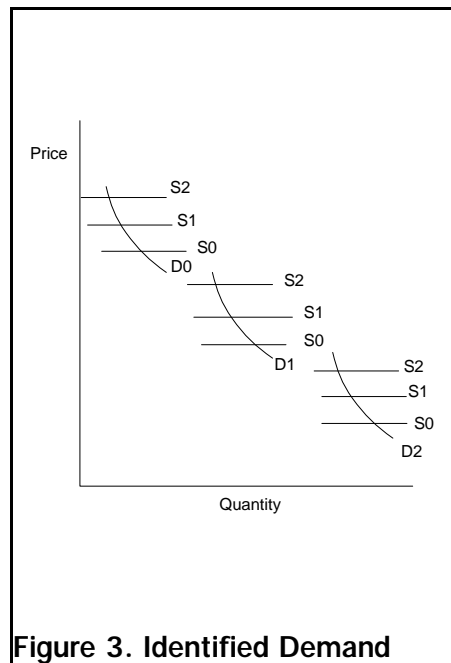


Figure 3. Identified Demand

<sup>8</sup> A technical presentation of the properties of this technique may be found in Dutta (1975).

<sup>9</sup> The weights in each interval given by the number of people in the locality that consume in that particular interval with respect to the total number of people in the interval for the whole state.

<sup>10</sup> Each interval can be defined at the lower level to avoid indeterminacy of the last block.

$$B(k) = \alpha + \beta Q_k + u \tag{8}$$

From (8) the estimate of  $\beta$  is interpreted as the marginal price, while the intercept  $\alpha$  is interpreted as the intramarginal cost.<sup>11</sup> Applying this method to the sample of rural communities in Argentina we linearized the rate schedule of each locality around discrete blocks common to all localities in the sample (from 1 to 65 m<sup>3</sup>/month in unitary increments). In this way, the information from all localities can be used jointly for demand estimation while each individual faces a marginal price that corresponds to his or her own locality.

The second instrument is similar to (8) above, except that instead of using an approximation to the bill schedule, TBR use an approximation to the average price schedule. For example, let:

$$b(k) = B(k)/Q_k \tag{9}$$

where  $b(k)$  represents the average price of interval  $k$ . The average price schedule can be approximated by assuming that:

$$b(k) = \alpha * Q^{\beta} \tag{10}$$

so, taking logarithms:

$$\ln(b(k)) = \ln(\alpha) + \beta \ln(Q_k) + u \tag{11}$$

where  $\ln(\alpha)$  and  $\beta$  represent the estimated parameters of the logarithmic transformation of the average price schedule, and  $u$  is an error term.<sup>12</sup>

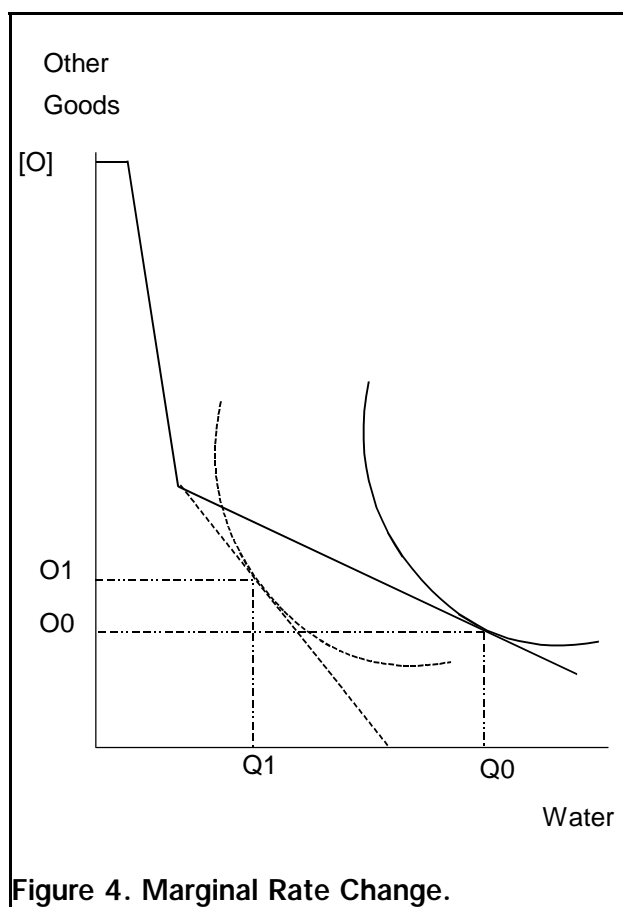


Figure 4. Marginal Rate Change.

<sup>11</sup> The significance of intramarginal cost will be explained below.

<sup>12</sup> Of TBR's two instruments, in our empirical application we found the rate schedule approximation was more satisfactory, but used neither in our final runs. See Section E for a discussion of the alternative

### Choice of Variables

Two of the classic problems with multipart rate schedules are: how to take into account the effect of changes in intramarginal rates, that is, changes in rates that do not correspond to the current level of consumption; and whether marginal or average prices should be used to estimate the demand function.

The first problem is neatly illustrated in Taylor's (1975) survey of electricity demand and can be seen in Figures 4 and 5. The figures are plotted in commodity space, where "O" represents a composite commodity (Other Goods) and "Q" represents water, with the usual indifference curve-budget line representation. In both figures the rate schedule consists of a fixed rate (Pf) represented by the horizontal line segment and two block rates (r1 and r2) represented by the negatively sloped line segments. The original position of the rates is represented by the solid linear line segments, and dashed lines represent the post change rate schedule.

In Figure 4 the marginal rate, or the rate at which consumption is currently (originally) obtained, is increased. Figure 4 shows how consumption declines from  $Q_0$  to  $Q_1$  due to the price change. Figure 5 illustrates the effect of a change in the intramarginal rate, and shows the effect on consumption (which goes from  $Q_0$  to  $Q_1$ ). The difference between the two cases is that the first rate increase is at the margin inducing the conventional income and substitution effects obvious in Figure 4. However, the second rate increase is intramarginal — the rate schedule shifts inward leaving the marginal rate in the last block unchanged, so the last segments of both budget lines in Figure 5 are parallel. This means that an intramarginal rate increase or decrease only has an income effect.

Awareness of this problem initially led to several alternative ways to take into account the size and changes in the intramarginal premium. The consensus reached appears to be that the approach proposed by Nordin (1976), and further explained by Griffin and Martin (1976), is the correct way of handling the problem. The solution is relatively simple; the effect of the intramarginal rate structure can be accounted for by creating a variable which represents the difference between what the consumer pays and what he would pay if all

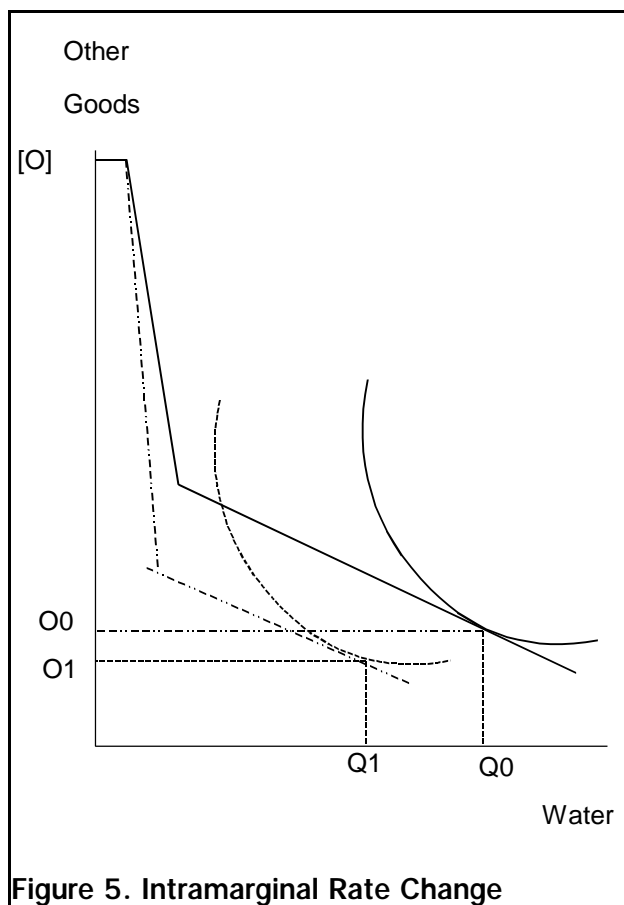


Figure 5. Intramarginal Rate Change

water demanded was charged at the marginal rate.

Since the effect of changes in intramarginal rates is essentially an income effect, it follows that the coefficients of income and the intramarginal premium should be roughly the same in absolute value but with opposite sign. In empirical applications, however, few cases report similar coefficients for income and the intramarginal premium. In fact, the overwhelming majority of reported results show coefficients that are significantly different.

This finding led to three types of reactions. Some have argued that water (or electricity) bills are so small relative to income that consumers will not look into the structure or changes in intramarginal rates. Consequently, consumers will not have a systematic response to changes in intramarginal rates. Others argue that in most studies the intramarginal effect is not estimated correctly, and this introduces a bias in the results. It is then concluded that incorrect estimation explains why the theoretical prediction (equal coefficients) is not empirically verified.<sup>13</sup> Another argument is that the intramarginal effect is so small that it is lost in the random influences that affect consumer behavior. Hence, it is not surprising to find that the value of both coefficients differs (Westley 1981). Finally, other authors note that most demand estimates are based on aggregate data, so the prediction that the income and intramarginal premium effect will have the same size is no longer necessarily valid (e.g. Blattenberger 1977: 115-120).

Agnostic researchers who are skeptical about the premises of standard demand theory argue that while a perfectly informed utility maximizing consumer should react to marginal, and not average price, few consumers may have a sufficiently detailed knowledge of the rate schedule or care to acquire it because of information costs. They may exhibit behavior which is not in strict accordance with economic theory. Instead, consumers may use some average price approximation as a guiding rule of thumb for consumption decisions for what is, after all, a good that usually has a relatively small share in total family expenditure.

Demand estimation is a problem in applied economics, and theory can take us just so far. For practitioners, whether the consumer reacts to the size of (or changes in) the intramarginal premium eventually becomes an empirical question. A refreshing aspect of some of the water demand literature is the notion that the potentially relevant variables for econometric estimation ought to include things the consumer can easily understand and decide on. Meaningful applied models have to sort out what is important from what is not. Water consumers may react to average prices, marginal prices, or even just their total monthly water bill, and data can be used to help resolve the issue.

In this vein, Oppaluch (1984) developed a model (elaborated by Chicoine and Ramamurthy 1986) that can be used to statistically test which price reaction alternative is supported by the data. The starting point is a demand model like:

---

<sup>13</sup> For instance, Schefter and David (1985) argue that most studies use average consumption to calculate the intramarginal premium. This assumes that all consumers would be consuming at the block corresponding to the average consumption, which is unlikely to be the case. The correct variable is the average intramarginal premium rather than the intramarginal premium calculated at the average level of consumption.

$$Q = " + \$_1 P_x + \$_2 P_n + \$_3 \{ [\sum_{i=1}^{n-1} (P_i - P_n) Q_i] / Q \} + \$_4 \{ Y - [\sum_{i=1}^{n-1} (P_i - P_n) Q_i] \} \quad (12)$$

where:

- $P_x$  = prices of other goods
- $P_n$  = marginal price of water at block n, where consumption occurs
- $P_i$  = price of water at intramarginal block  $i < n$
- $Q_i$  = quantity of water at intramarginal block  $i$

The second term in (8) represents the marginal price of water, or the price of water at the block where consumption occurs. The third term is the difference between average price and marginal price and the last term is income minus the intramarginal premium. The average price for consumption up to block n is:

$$P = \frac{\sum_{i=1}^n (P_i * Q_i)}{Q} \quad (13)$$

Average price in Eq. (13) can also be expressed as:

$$P = \frac{[\sum_{i=1}^n P_i * Q_i] + P_n (Q - \sum_{i=1}^{n-1} Q_i)}{Q} \quad (13.1)$$

which can be transformed into an expression that includes a marginal price term and an average price term that is net of the marginal price:

$$P = P_n + \frac{\sum_{i=1}^{n-1} (P_i - P_n) Q_i}{Q} \quad (13.2)$$

which are the second and third terms in Eq. (12). Restricted and unrestricted estimation of Eq. (12) allows the following statistical tests to be made (Chicoine and Ramamurthy 1986):

	<u>Test 1</u>	<u>Test 2</u>
Null Hypothesis H0:	$\$_3 = 0$	$\$_2 = \$_3$
Alternative Hypothesis H1:	$\$_3 \dots 0$	$\$_2 \dots \$_3$

Tests on Eq. (12) can yield the five different outcomes summarized in Table 1.

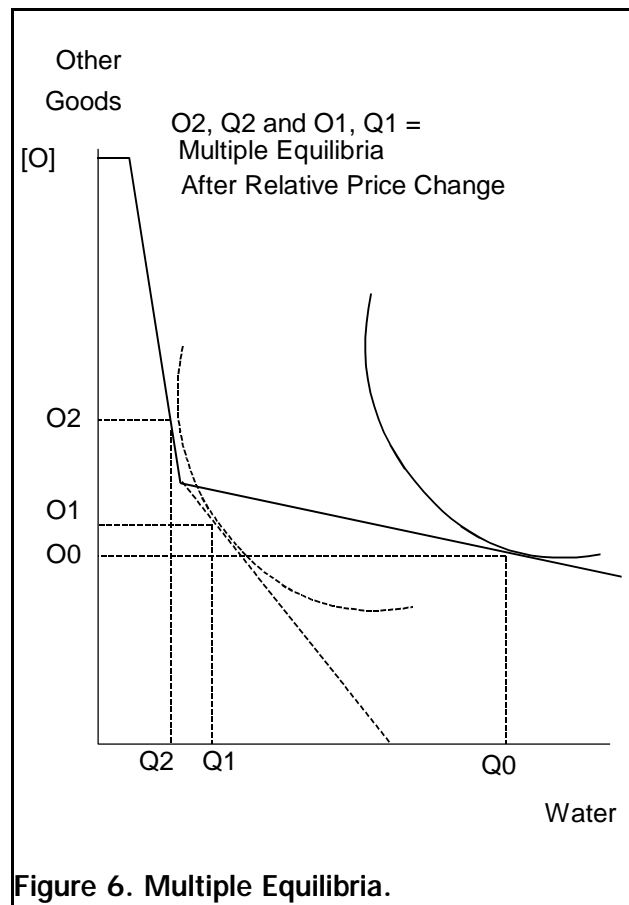
Table 1  
Possible Test Results

Test Results	Conclusion
1. $\beta_3 = 0$ and $\beta_2 \neq \beta_3$	====> Marginal Price Model
2. $\beta_3 \neq 0$ and $\beta_2 = \beta_3$	====> Average Price Model
3. $\beta_3 = 0$ and $\beta_2 = \beta_3 = 0$	====> No Price Response Model
4. $\beta_3 = 0$ and $\beta_3 = \beta_2; \beta_2 \neq 0$	====> Indeterminate Model
5. $\beta_3 \neq 0$ and $\beta_2 \neq \beta_3$	====> General "Decomposed" Price Model.

In other words, if the coefficient of  $\beta_2$  is significantly different from zero (and presumably negative) while  $\beta_3$  is not significantly different than zero (i.e. test result 1) the marginal price hypothesis cannot not be rejected. If both coefficients are insignificantly different from each other (i.e test result 2), then the average price hypothesis cannot be rejected.

*Properties of the Demand Function*

One of the most troublesome aspects of demand estimation when prices are set in multipart block rates is that individual demand functions may be discontinuous. This problem was first highlighted by Taylor (1975), and can be easily understood with the following example. In Figure 6 we have a tariff schedule with a fixed charge and two unit tariff blocks. Initially, an indifference curve corresponding to a single household has tangency point at Q0. If the price of the second block increases the consumer would be indifferent between Q1 and Q2, which means that the solution to this problem is indeterminate. If we assume that the choice is Q1, the demand curve would have a "hole" corresponding to the points between Q1 and Q2. If this is the case, it is not clear what sort of demand curve can be estimated, or more fundamentally, which (if any) meaningful properties of demand systems would be valid.



In the applied literature, two explicit approaches have been proposed to circumvent this problem. The first one argues that aggregate demand functions are not discontinuous to the extent that different individuals have different tastes. By taking an average of a sufficiently large number of individuals, indifference curves can be presumed to be located along the rate schedules.

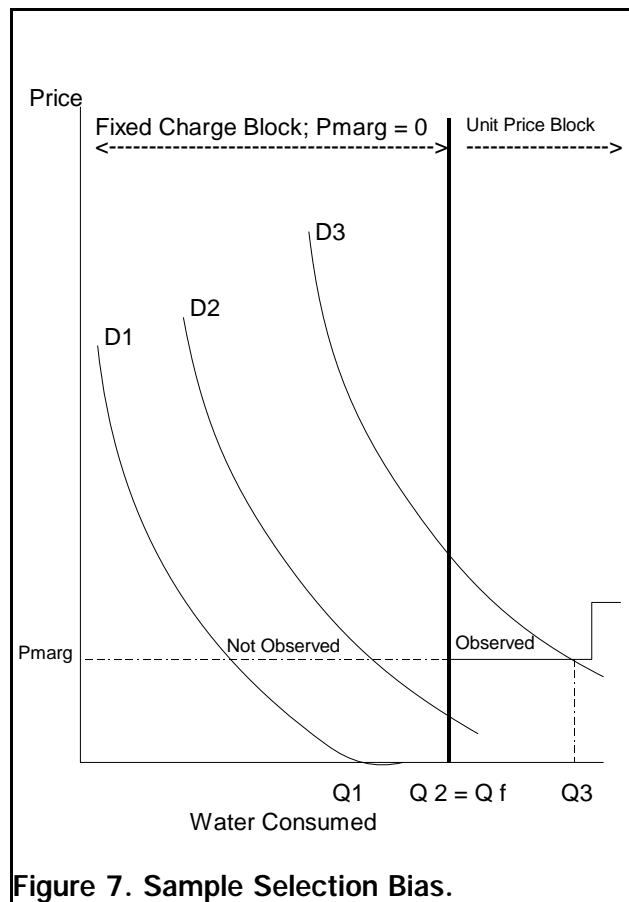
The main problem of this approach is that the aggregate demand curves estimated on the basis of averages for groups of individuals no longer have the standard properties of individual demand functions derived from consistent utility functions. The only properties that remain after aggregation are homogeneity and budget exhaustion (Sonnenschein 1973). Interpretation of results in terms of elasticities is therefore unclear. In addition, it is frequently found that estimates based on averages tend to overstate demand price elasticities.<sup>14</sup>

The second solution to the problem of discontinuity is somewhat roundabout, and has been proposed by Wade (1980). This approach consists of estimating demand indirectly through estimation of the parameters of a Stone-Geary utility function. Once the parameters of the utility function are identified, demand can be calculated using any appropriate optimization algorithm. In this approach the existence of "holes" is still possible, although the quantities of water demanded can always be estimated in a manner consistent with utility theory. In this approach, the concept of price elasticity is transformed into a weighted average of elasticities on different blocks.

In our Argentina application the data are from a cross-section of communities, most of which have increasing block rate schedules. Therefore the kinked demand problem is not serious, and does not affect the estimation techniques for this application, or the interpretation of results. We therefore do not dwell on the issue further.

### **Sample Selection Bias**

Sample selection bias may be important when some sampled consumers are not in the block rate schedule, but instead only pay a fixed charge for quantities of water below some statutory minimum that triggers block rate pricing on a per unit consumed basis (Wade 1980). (Chicoine and Ramamurthy (1986) note the possibility of sample



**Figure 7. Sample Selection Bias.**

<sup>14</sup> See Bohi (1981: 10-12 and 149-151) for an interpretation of elasticities estimated from aggregate data and bias in coefficients.



selection bias, but do not correct for it.) In this case it appears on the surface that marginal price can be set to zero and the fixed charge customers included in the estimation sample. However, more properly, marginal price is not necessarily zero — indeed it is not observed at all for a potentially large number of fixed charge customers with demands characterized by schedule D2 in Figure 7. Their observed behavior contains no useful information on the relationship between marginal willingness to pay (equals marginal price) for alternative quantities of water below the trigger quantity in equilibrium. Figure 7 shows three representative consumers who differ by level of income. Consumers with demand curves D1 and D2 are not in the rate schedule and do not pay an observed marginal price, while the consumer with demand curve D3 can equate marginal willingness to pay with the unit price by consuming at Q3.

If the fixed charge customers who are not in the block rate schedule are deleted from the estimation sample and the demand relation is estimated using only the observations falling within the block rate schedule where unit prices and their corresponding quantities are observed, OLS or IV estimators may produce biased and inconsistent parameter estimates. The simple reason is that the expected value of the error term in the demand equation for observed consumers falling in the block rate schedule is likely to be non-zero, violating the usual assumption.

Because sample selection bias is the most difficult issue to understand and deal with econometrically, its origins and the methods that can be used to handle it are explained in some detail. Only then is the additional complication of simultaneous equations in a sample selection framework introduced, and a consistent technique for estimation of the parameters of the structural water demand relation proposed.

### *Problems with Ordinary Least Squares, even in the Absence of Simultaneity*

For simplicity, suppose we set aside the issue of simultaneity by assuming that rate schedules offer water at a constant per-unit price ( $P_{\text{marg}}$ ) once some consumption threshold ( $Q_f$ ) has been crossed, instead of being step functions beyond the threshold. Also assume two types of consumers — households with a strong demand for water whose demand schedule is, for reasons of income, wealth, or taste, located above (to the northeast of) the demand schedule of a second category of users with a weaker demand for water. In this example, the high income households pay the lump sum customer charge and a marginal rate as well because their demand schedule pulls them into the region of the rate schedule. But low income households pay only the lump sum charge and consume whatever the household requires below the  $Q_f$  threshold based on the role of exogenous socioeconomic variables other than price,  $P_{\text{marg}}$ .

The analyst only observes the price-quantity relationship for the first, wealthier class of consumer falling in the block rate schedule. The less well-off customers array themselves along the  $P_{\text{marg}} = 0$  axis below the  $Q_f$  threshold. While their willingness to pay for the quantities they are known to consume cannot be observed, there is a hypothetical marginal price that would validate or call forth these quantities. Some of the families below  $Q_f$  will use water until its marginal value is truly zero, where their demand schedule cuts the horizontal axis (D1 in the Figure). But other families may have demand schedules that are everywhere below the prevailing rate schedule, and will consequently be clustered around the threshold

$Q_f$  (shown as D2 in the Figure). With random error in the data, there is no way to easily distinguish between these two cases.<sup>15</sup> The true demand schedule for these poorer families is unobserved because of the artifact of the two-part (lump sum cum unit rate) tariff schedule.<sup>16</sup>

While the analyst would like to have price observations for the consumer group in the fixed charge category below  $Q_f$ , these are not observed.  $P_{marg}$  could be coded as equal to zero for each and every consumer in that category, but unfortunately zero has no economic content or meaning for a subset of the observations below  $Q_f$ . It is merely a convention that indicates a missing value for  $P_{marg}$ .

Of course, in estimation one could proceed following the above convention, replace the unobserved prices with zero, and estimate the demand function over the full sample of customers in both classes as if an unavailable marginal willingness to pay value was a true zero. The consequence of so doing, given any set of exogenous variables, would be to pull the estimated price intercept down from the true intercept and to decrease the estimated price slope relative to the true slope. The result would be a false hybrid demand function.

Another approximation would be to construct a marginal price proxy for observations below  $Q_f$  by dividing the fixed charge paid by the quantity actually consumed. Again, just like setting  $P_{marg}$  to zero, there is no guarantee that this average price proxy will correspond closely with the true marginal willingness to pay measure for price.

A final, and more reasonable approach would be to estimate the demand relation only from observations falling in the rate schedule (i.e.  $Q > Q_f$ ) discarding all fixed-charge-only observations where marginal price is not a decision variable and is not observed. Unfortunately, the parameters estimated from this censored sample by OLS may be inconsistent. This arises because the expected value of the error in such an OLS regression is non-zero, violating full ideal conditions of the classical normal linear regression model (see Greene 1990). As discussed next, a correction technique exists to produce consistent demand function parameter estimates.

#### *The Heckman Approach: Sample Selection Bias Correction with no Simultaneous Equations*

For the moment treating marginal price as exogenous, the full sample of water consumers can be separated into those not in the rate schedule for whom marginal price is not observed and those in the rate schedule paying an observed marginal price. Consider the following two regime three-equations model (Heckman, 1979, Lee and Trost 1980, Mullahy 1986):

---

<sup>15</sup> See Wade 1980 for a numerical example and an attempt to correct the problem using the Tobit estimator.

<sup>16</sup> For an innovative approach that can be used to augment the sample of public water utility customers in the rate schedule with additional information on unconnected customers paying high prices to private suppliers (tank trucks) and consuming low quantities relative to public utility customers, see Gómez (1987). In essence, Gómez's method provides information on legitimate price/quantity pairs for families like those in the public utility's fixed charge part of the schedule (e.g. segments of D1 and D2 in Fig. 7 above  $P_{marg}=0$ ) where such information is not observed. However, this method does not solve the simultaneity bias or sample selection problems.

Regime 1. Rate Schedule Demand.  $Q_i \geq Q_f$

$$Q_{1i} = X_{1i}'\beta_1 + P_{\text{marg}}\beta_2 + u_{1i} \quad \text{iff } Z_i(\beta_3 + u_{3i}) \geq Q_f \quad (14)$$

Regime 2: Fixed Charge Demand.  $Q_i < Q_f$

$$Q_{2i} = X_{2i}'\beta_3 + u_{2i} \quad \text{iff } Z_i(\beta_3 + u_{3i}) < Q_f \quad (15)$$

Regime 3: Selection/Latent Demand

$$Q_{3i}^* = Z_i(\beta_3 + u_{3i}) ; \quad u_i = N(0, 1) \quad (16)$$

where  $Q_{1i}$  is the amount of water consumed,  $X_{1i}$ ,  $X_{2i}$ , and  $Z_i$  are row vectors of exogenous explanatory variables, and  $Q_{3i}^*$  represents the difference between the unobserved quantity of water desired and the minimum threshold  $Q_f$ . If this difference is positive, the consumer crosses the threshold into the rate schedule; if it is negative the consumer pays only a fixed charge to consume an amount below  $Q_f$ .

Here the quantity of water demanded if a user is in the rate schedule,  $Q_{1i}$ , is greater than the quantity demanded if the user pays no marginal price,  $Q_{2i}$ , and belongs to the fixed charge regime. Given the exogenous variables  $X_{1i}$ ,  $X_{2i}$ ,  $Z_i$  (and for now  $P_{\text{marg}}$ ) the sample of rate schedule customers is observed only if  $Q_{3i}^*$  is positive; that is if the unobserved intensity of desire to consume water passes beyond the  $Q_f$  threshold defining the point of entry into the rate schedule. Otherwise, the unobserved intensity of desire measure falls below  $Q_f$  and a marginal price/quantity consumed equilibrium pair are not observed. Instead in the second regime only fixed charge demand  $Q_{2i}$  is observed.

In this model the latent measure  $Q_{3i}^*$  is unobserved. But since the variables it relates to (the  $Z_i$ s) are, a discrete indicator can be constructed from the data by observing which regime each observation belongs to:

$$\begin{aligned} I_i &= 1 \text{ for } Q_{3i}^* > 0 \text{ (Regime 1)} \\ I_i &= 0 \text{ otherwise (Regime 2)} \end{aligned} \quad (17)$$

In other words, usable price-quantity pairs are available to the analyst only when  $I_i = 1$  because latent indicator exceeds the  $Q_f$  defining the threshold of entry into the rate schedule.

If the error terms in the demand equation and the intensity equation are correlated with zero means and covariance  $F_{13}$  Heckman considers the expectation  $E(Q_{1i} * I_i = 1)$  which can be written:

$$E(Q_{1i} * I_i = 1) = X_{1i}'\beta_1 + P_{\text{marg}}\beta_2 + E(u_{1i} * I_i = 1) \quad (18)$$

where the notation  $(*I_i = 1)$  refers to the sample selection rule, "conditional on being in the rate schedule."

Consider least squares estimation using only those observations which qualify given the rule in Eq. (16). If OLS is performed on (14), will consistent parameter estimates result? The answer from the sample selection literature is in general, no, unless  $u_{1i}$  is independent of  $u_{3i}$

so the conditional expectation of  $u_{1i}$  is zero and  $F_{13}$  equals zero. This is true because it can be shown (Heckman 1979, Lee and Trost 1980, Olsen 1980, Mullahy 1986) that the expected value of the error in the rate schedule demand equation is not zero:

$$E(u_{1i} * I_i = 1) = F_{13} \mathfrak{S} \quad (19)$$

Here,  $F_{13}$  is the covariance between the error terms in the rate schedule demand function (Regime 1) and the selection decision (Regime 3).<sup>17</sup> The term  $\mathfrak{S}$  is the ratio of the standard normal probability density (N) evaluated at  $Z_i$  to the standard normal distribution function (M, or the probability that observation  $i$  belongs in Regime 1) evaluated at the same point ( $\mathfrak{S}$  is called the Inverse Mills Ratio). If an observation has a high probability of containing useful data for Regime 1, its value for  $\mathfrak{S}$  will be small and positive. Therefore, selectivity bias would be unimportant if all the  $\mathfrak{S}$  were negligibly small and positive because the probability of sample inclusion is high for all observations. Also, unless the covariance between the Regime 1 demand equation and the latent indicator equation is zero (i.e.  $F_{13} = 0$ ), OLS on Eq. (18) will yield inconsistent parameter estimates because  $E(u_{1i} * I_i = 1)$  includes the disturbance but OLS ignores it under the false assumption that  $E(u_{1i}) = 0$ .

Heckman's suggested procedure to estimate Eq. (18) consistently proceeds in two steps. In the first step a Probit model is estimated on the binary indicator  $I_i$  defined by Eqs.(16) and (17), producing consistent estimates of the  $\gamma$  parameter vector. Then for each observation an estimate of  $\lambda$ ,  $\hat{\mathfrak{S}}$ , can be constructed and employed as a regressor in Eq. (18) above using the data from the subsample in Regime 1. The second stage regression is:

$$Q_{1i} = X_{1i} \beta_1 + P_{marg_i} \beta_2 + \hat{\mathfrak{S}} \beta_3 + V_i \quad (20)$$

where  $V_i$  is an error term with mean zero and  $\beta_3$  is an estimate of  $F_{13} = DF_{11}$ . Note that OLS estimates on the Regime 1 sample which ignore sample selection would not include  $\hat{\mathfrak{S}}$  as a regressor.

### *Extension of the Sample Selection Model to the Simultaneous Equations Situation*

Relaxing the assumption that the rate schedules (in a cross sectional sample of consumers in different localities) are characterized by constant prices per unit of consumption introduces the complication of block rate pricing. As explained earlier, this raises the distinct possibility of simultaneous equations bias because marginal price (when it can be observed) is no longer exogenous, but is chosen by each consumer simultaneously with consumption quantity (when it can be observed). The solution to this additional complication is the extension of instrumental variable methods to the sample-separation case generally characterized by two simultaneous equations systems corresponding to the two different regimes and a selectivity criterion determining the regime to which the observations belong (Lee, Maddala and Trost 1980).

---

<sup>17</sup> Note that the standard deviation  $F_{33}$  cannot be estimated, so it is conventionally set to one. Therefore the covariance term  $F_{13}$  between  $F_{33}$  and  $F_{11}$  is conventionally expressed as  $DF_{33}F_{11}$  where  $D = \text{corr}[u_1, u_3]$ . Therefore,  $F_{13} = DF_{11}$ .

Under these circumstances, and ignoring the fixed rate regime (because the endogenous price and quantity variables corresponding to the demand function are only observed when  $I_i = 1$ ), the simultaneity appears only in Regime 1. With simultaneity, and a cross-sectional sample, Regime 1 is now expanded by adding a multi-location block rate supply-side equation to the system along with the demand equation and the probit selection criterion function.

I. Regime 1

Demand

$$Q_{1i} = X_{1i} \beta_1 + P_{\text{marg}} \beta_2 + u_{1i} \quad (21)$$

Supply

$$P_{\text{marg}_i} = L [P][B] + u_{2i} \quad (22)$$

II. Probit selection criterion

$$\text{PROB} (I_i = 1) = 1 - F(-Z_i) = 1 - M(-Z_i) = M(Z_i) \quad (23)$$

Here for any individual,  $i$ ,  $X_{1i}$  is a  $1 \times k$  row vector of exogenous explanatory variables as before;  $\beta_1$ , is a  $k \times 1$  column vector of demand parameters;  $\beta_2$  is the parameter on price;  $L$  is a  $1 \times j$  column vector of  $j$  localities;  $P$  is a  $j \times m$  matrix which represents the price per cubic meter of water paid in each location at  $m$  different consumption intervals;  $B$  is an  $m \times 1$  column vector which identifies the block chosen by the consumer; and  $Z_i$  is a  $1 \times j$  row vector of explanatory variables determining whether an observation belongs in Regime 1 (price and quantity pairs observed) or not; and  $F(\cdot)$  is the cumulative normal distribution function evaluated at  $Z_i$  (for the Probit,  $\text{PROB}$  indicating probability).

On the supply side, the entire across-community block rate structure can be written exactly by the analyst with prior knowledge from the rate schedule step functions reporting  $B$  and  $P$ . Nevertheless, the analyst does not know a-priori with any certainty which of the  $P$  rates will be chosen by any individual, so the  $B$  variable is endogenous in this model — e.g. any  $B_m$  must fall somewhere in one of the blocks defined by  $B$ . So the exogenous variable on the supply side is the location dummy variable represented by  $L$ .<sup>18</sup>

To estimate the structural parameters  $\beta_1$ ,  $\beta_2$  of the demand relation in this model the instrumental variable approach outlined in Lee, Maddala, and Trost can be used. It is directly analogous to two stage least squares with the addition of an estimate of the Mills Ratio  $\beta$  (constructed from the Probit equation) as an extra instrument.

The estimation steps are (Lee, Maddala, and Trost 1980):

---

<sup>18</sup> There are  $j$  dummy variables, equal to the number of communities in the sample minus one if an intercept is in the supply equation. Obviously if the sample included only one community and one time period, the demand function could not be identified.

- 1) Estimate the Probit criterion equation and construct an estimate of the Inverse Mills Ratio  $\delta$  for each individual known to be in the rate schedule.
- 2) Estimate a reduced form equation for marginal price,  $P_{\text{marg}}$ , using all exogenous explanatory variables in the Regime 1 system ( $X_{1i}$ ,  $L$ ) as instruments along with the  $\delta$  from step 1.
- 3) Predict the instrument for marginal price using the equation estimated in step 2.
- 4) Estimate the demand equation on the Regime 1 sample using the exogenous explanatory variables  $X_{1i}$ , the constructed instrument for  $P_{\text{marg}}$  from Step 3, and the constructed Inverse Mills Ratio  $\delta$  from Step 1. The former accounts for the endogeneity of price and the latter for sample selection bias, just as before.

### ***Model Specification and Data***

#### *Supply Variables*

The first step in estimating water demand functions for Argentina was to find or create exogenous variables that could represent supply schedules and circumvent simultaneity. For this purpose we considered variables that could be used to estimate marginal prices and subsequently use the prediction to estimate demand. Initially the following candidates were considered:

- a. Average Marginal Price: represented by the slope of a linear approximation to the total revenue schedule of each locality, as proposed by TBR.
- b. Marginal Rate Schedule (P): represented by the rate charged per cubic meter of water in 16 consumption block intervals common to all localities, as proposed by Westley (1984).
- c. Locality Schedule (L): represented by dummy variables identifying each locality.
- d. Block Schedule (B): represented by dummy variables identifying the block in which each household is consuming in the respective rate schedule.

Examples of the last three types of variables are shown in the following tables. Table 2 shows hypothetical rate schedules for three localities, that is, the cost per cubic meter of water at different consumption intervals. Using the information from the top part of Table 2, rate schedules (P) can be calculated as shown in the lower part. Note that in this hypothetical example each rate schedule is defined by four variables, each one corresponding to a consumption interval. In the sample rural localities in Argentina there were sixteen common intervals with at least one different marginal price each (sixteen rate schedule variables). Two of them had to be discarded because they turned out to be linear combinations of other rate schedules, so the final number of variables representing the rate schedule was fourteen.

Table 2  
Localities

Consumption Range	Cost per M3 in:			
	San Juan	Veronica	Anguinan	
0-100 M <sup>3</sup> /month	0.6	0.3	0.0	
101-200 M <sup>3</sup> /month	0.6	0.6	0.0	
201-300 M <sup>3</sup> /month	1.1	0.9	1.0	
More than 300	1.1	1.2	1.0	
Rate Schedules	RS1	RS2	RS3	RS4
San Juan	0.6	0.6	1.1	1.1
Veronica	0.3	0.6	0.9	1.2
Anguinan	0.0	0.0	1.0	1.0

An example of location schedules is shown in Table 3 below. The location schedule is a set of dummy variables that represent locations. For example, all families in, say, San Juan, would show LS1 = 1.0 and zero for all other L. This is also shown in the lower part of Table 4 (last two columns) which illustrates how block schedules are constructed for particular households in a given locality. The block schedules identify the type of rate schedule faced by each household and in what portion of the rate schedule they consume.

Table 3  
Location Schedule

Name of Locality	LS1	LS2	LS3
San Juan	1.0	0.0	0.0
Veronica	0.0	1.0	0.0
Anguinan	0.0	0.0	1.0

Table 4  
Block Schedules for San Juan Families

Number	Rate	BS1	BS2	BS3	BS4	LS1	LS2
1	0.6	1	1	0	0	1	0
2	0.6	1	1	0	0	1	0
3	1.1	0	0	1	1	1	0
4	0.6	1	1	0	0	1	0
5	1.1	0	0	1	1	1	0
6	1.1	0	0	1	1	1	0

For example, the family identified as number 1 consumes in a block that ranges from 0 to 200 m<sup>3</sup>/month. In this block the price is \$0.60 cents per m<sup>3</sup> and the family is located in locality number 1. In contrast, family 3, which lives in the same location, is consuming at a block ranging between 201 and more than 300 m<sup>3</sup>/month at a price of \$1.1 per m<sup>3</sup>.

Of these four variables, Westley's Marginal Rate Schedule (P) and the Location Schedule (L) were by far the best predictors for marginal price. All subsequent tests and IV regressions use one or another of these two variables as alternative instruments.

### *Demand Variables*

Apart from the usual variables that appear in any demand function (price and income), data from the OEO survey contained information for several other indicators that could potentially explain water consumption. These variables ranged from family attributes (number of household members, number of adults, number of children, etc.), family assets (ownership of land, house, radio, television, etc.), type of sewerage facilities, sources and uses of water, water supply reliability, household water storage facilities, etc.

After carefully reviewing the information, some variables (and some individual surveys) were discarded. With the reduced data set we tested the influence of different combinations of variables. Some of the attributes turned out to be highly collinear (e.g. assets and income); and many of the system attributes (e.g. sources and uses of water, type of sewerage system) were too uniform across the sample to be of any value in terms of additional information. These features of the sample reduced the number of usable variables to:

- a. Marginal Price;
- b. Intramarginal Premium;
- c. Income;
- d. Number of adults in the household (Adults);
- e. Electricity Bill as a wealth proxy (Electricity);
- f. Number of water faucets inside the house (Faucets); and
- g. Average number of interruptions in service per month (Outages).

### *Water Demand Equation Specification*

The specifications tested combine three types of variables: marginal price, intramarginal premium and income, and other variables (number of adults, electricity bill, number of water faucets, etc., as listed above).

The marginal price is in all cases the per unit cost of water at the level where the household is consuming. If the household is consuming at the minimum, fixed charge level, the marginal price is set to zero.

The intramarginal premium is generally defined in the literature as the difference between what the consumer actually pays and what would have to be paid if all units consumed were



charged at the marginal rate. In our application the premium was calculated in two ways: as a lump-sum transfer, or as the difference between average and marginal prices.

In the first case the premium is a subtraction (or addition) to income. The theoretical expectation is that the coefficients from this type of premium and household income will be the same but with opposite sign. In this formulation, the lump-sum also includes the fixed water charge since this cost is a deduction from income related to the act of consuming water, but divorced from the decision of how much to consume at the margin. In some applications, the intramarginal premium is directly deducted from income, which means that the equality of coefficients for both variables is implicitly assumed.

The general model can therefore be estimated in two variants, with and without the constraint on the equality of the coefficients of the intramarginal premium and income. That is:

Model I. Unconstrained, Marginal Price, Inequality of Income and Premium Effect

$$Q = \alpha + \beta_1 * P_{\text{marg}} + \beta_2 * (\text{Premium} + \text{Fixed Charge}) + \beta_3 * Y + \beta_4 * [\text{Other}]$$

Model II. Constrained Income Effect Version of I.

$$Q = \alpha + \beta_1 * P_{\text{marg}} + \beta_2 * (Y - \text{Premium} - \text{Fixed Charge}) + \beta_4 * [\text{Other}]$$

A second formulation of the intramarginal premium, tailored to the Oppaluch test, is defined on a per unit of water basis. In this case consistency with the derivation of the model specification requires that we subtract from income all payments related to water consumption (including the fixed charge). Thus the specification would be:

Model III. Oppaluch's Marginal and Average Price Model.

$$Q = \alpha + \beta_1 * P_{\text{marg}} + \beta_2 * Pr + \beta_3 * (Y - \text{Premium}) + \beta_4 * [\text{Other}]$$

where  $Pr = \text{Average Price} - \text{Marginal Price}$ .

Each specification was estimated with two different samples: the full sample of 685 families, and the subset sample of 279 families consuming above the minimum, fixed charge level. Estimated demand functions for both samples using OLS and 2SLS are reported below for all five model specifications.

The purpose of using different samples was to compare the effects of (erroneously) including families consuming at the minimum level without correcting for the lack of an appropriate marginal price measure. Although the marginal price for these families is set to zero, their consumption level does not provide information on their price/quantity demanded choice (unless the marginal utility of consuming water for public system customers at the observed level is zero, or if the sample of public utility customers is augmented with privately supplied low income customers following Gómez 1987). The relative size of the bias introduced by using data from households in the fixed charge block is not known a-priori, so it is interesting

to compare the results between both samples.<sup>19</sup> Estimation using a smaller sample of consumers in the unit price part of the rate schedule beyond the fixed charge block is a way of correcting this problem. In contrast, the comparison between OLS and 2SLS results illustrates the effects of simultaneous equations bias and sample selection bias.

### ***Estimation Results***

The following tables present the estimation results, which all maintain the hypothesis of a linear demand function.<sup>20</sup> We start with a comparison of OLS models using the full sample of 685 observations, shown in Table 5. These OLS results are generally very poor. The signs of the price and income coefficients are in most cases wrong or insignificantly different from zero, which is evidence of the twin effects of simultaneous equation bias and the improper coding of marginal prices in the fixed charge block as zero. In contrast, the signs of other variables are always correct, often significant and do not change much compared to the ones estimated using 2SLS.<sup>21</sup>

---

<sup>19</sup> Our sample only includes public utility customers, so the implications of Gómez's sample augmentation approach could not be empirically explored.

<sup>20</sup> Other functional forms were explored, but our conclusions about simultaneity and sample selection bias remained robust across functional specification.

<sup>21</sup> The only exception for signs is with Outages in the 2SLS results.

Table 5  
OLS Full Sample Results  
(Absolute Value of "t" Statistic in Parentheses)

Model	I	II	III
Constant	12,588.8 (13.5)	9,441.0 (11.4)	9,388.4 (11.4)
Marginal Price	-590.3 (0.5)	4,370.0 (5.9)	6,179.7 (4.5)
Premium	-498.9 (6.5)	-	20,093.6 (1.5)
Income	-0.1 (0.6)	-.4 (1.6)	-0.4 (1.8)
Adults	257.6 (1.5)	387.0 (2.2)	379.5 (2.1)
Faucets	306.2 (2.6)	352.0 (2.9)	357.0 (2.9)
Outages	-729.0 (1.2)	-165.0 (0.3)	-92.9 (0.1)
Electricity	68.9 (5.8)	61.3 (1.6)	60.5 (4.9)
R <sup>2</sup>	0.18	0.12	0.12
F	21.5	16.9	14.9

Note: Marginal price in fixed charge block coded as zero.

Results using 2SLS on the Full Sample with two different instruments (P or L) for the rate schedule are shown in Tables 6 and 7.

Marginal prices show correct coefficients (except in Model III with the P instrument) and are generally significantly different from zero. The elasticity values for prices, premium and income are low, which is a result consistent with the bias introduced by including families in the minimum consumption category. For example, the highest price elasticity value is only -0.12, the highest elasticity of the intramarginal premium is -0.08 (Table 7, Model I) and income coefficients have correct results only in one case (Model I).

As mentioned above, the values of the coefficients for all other variables (with the exception of service interruptions, Outages) are not significantly affected when estimated with 2SLS. The constant term in both cases roughly corresponds to a consumption level of about 330 liters per connection per day. This is reasonable since the average household size is between 3 and 4 people and the average per capita consumption level in the sample is close to 100 liters per persons per day. For the same reason, the coefficient on the number of adults in the household (Adults) is also rather low, because it implies that adding or taking away one adult from the household will change daily consumption by about 10 to 20 liters per day.

Table 6  
2SLS Full Sample Results With Location Dummy Variable Instrument  
*(Absolute Value of "t" Statistic in Parentheses)*

Model	I	II	III
Constant	11,720.0 (11.1)	9,833.7 (11.4)	10,013.5 (11.0)
Marginal Price	-4,690.0 (3.2)	-1,977.0 (1.7)	-8,115.0 (1.5)
Premium	-302.0 (3.0)	- -	-6,789.0 (1.2)
Income	.0 (0.0)	-0.05 (0.2)	-0.05 (0.02)
Adults	303.0 (1.7)	381.1 (2.1)	405.9 (2.1)
Faucets	383.0 (3.1)	414.3 (3.3)	404.7 (3.1)
Outages	-225.3 (0.0)	129.0 (0.2)	-99.7 (0.1)
Electricity	86.4 (6.7)	82.8 (6.3)	86.5 (6.2)
Elasticities:			
Ep	-0.08	-0.03	-0.13
Epr	-0.07	-	-0.08
Ey	n.s.	n.s.	n.s.

n.s.= Result is not significant.

Note: Marginal price in fixed charge block coded as zero.

It is difficult to judge the reasonableness of the size of other coefficients. The coefficient on Faucets indicates a relationship between the number of inside faucets and water consumption similar to the effect of the number of adults on monthly water consumption. Frequency of service interruptions (Outages) in most cases has the anticipated negative sign in OLS and the wrong sign in 2SLS, but is generally insignificant using either estimator. Finally, the value of the electricity bill (as a proxy for wealth) indicates a positive relationship between wealth and water consumption, with an average elasticity of about 0.15. In terms of price and income coefficients the main change from OLS to 2SLS is that the price, premium and income coefficients take on the theoretically expected signs, even though the elasticities are low.

Oppaluch's test (Model V) is inconclusive. In OLS the price and premium coefficients are both significant, and significantly different from each other, but have the wrong sign. In 2SLS both parameter estimates are insignificantly different from zero at the 95% level, implying no price response. An important implication of these results is that if we estimate with the full sample of observations, OLS will not yield correct results, while 2SLS yields low but theoretically consistent elasticities for the models maintaining a marginal price hypothesis.

Table 7  
2SLS on Full Sample With Rate Block Dummy Variable Instrument  
*(Absolute Value of "t" Statistic in Parentheses)*

Model	I	II	III
Constant	12,164.0 (10.5)	9,940.0 (11.1)	9,817.1 (10.8)
Marginal Price	-7,146.0 (4.0)	-3,697.0 (2.7)	348.8 (0.05)
Premium	-353.0 (2.9)	- -	4,363.8 (0.7)
Income	.09 (0.4)	-0.05 (0.2)	-0.16 (0.6)
Adults	288.2 (1.6)	379.6 (2.0)	364.6 (1.9)
Faucets	397.6 (3.1)	431.0 (3.3)	437.6 (3.4)
Outages	-192.7 (0.3)	209.0 (0.3)	348.0 (0.5)
Electricity	93.8 (6.9)	88.6 (6.5)	86.1 (6.1)
Elasticities:			
Ep	-0.12	-0.06	n.s.
Epr	-0.08	-	n.s.
Ey	n.s.	n.s.	n.s.

n.s. = Result is not significant.

Note: Marginal price in fixed charge block coded as zero.

Table 8 presents the results of applying OLS and 2SLS to the subsample of 279 families consuming above the minimum, fixed charge consumption level instead of the entire sample of 685 households. The results are interesting. First, in contrast to the results using the full sample, the price coefficients (and corresponding elasticities) estimated with OLS using the restricted sample are not too different from the ones calculated using 2SLS. Second, elasticities and price coefficients are much higher in this sample. The price elasticity ranges between -0.2 and -0.4 (compared to -0.04 to -0.12 with the full sample). Third, the size and significance of the premium coefficients is erratic. The results of the Oppaluch test (model III) are particularly inconclusive since the price and premium parameters are significant in OLS (price with the wrong sign) and insignificant in 2SLS. Fourth, income coefficients are significant and have the correct sign in model I, although the elasticity value (0.02) still appears to be low.

The most important general result of this group of regressions is that the effects of sample selection bias have a greater impact than the simultaneity problems in this particular case. In other words, within the subsample of consumers that are above the minimum level, the effects of simultaneity are very small. This is most likely because the distribution around the

different rate schedules is more concentrated in one or two blocks per community, such that in several cases the supply schedule is almost perfectly elastic.

Model	I		II		III	
	OLS	2SLS**	OLS	2SLS**	OLS	2SLS*
Constant	24,508.0 (21.8)	23,927.0 (21.2)	19,830.0 (15.8)	20,104.6 (16.0)	19,495.9 (17.5)	20,087.1 (16.5)
Marginal price	-14,262.0 (12.7)	-13,629.0 (11.7)	-8,077.0 (6.9)	-8,648.0 (7.1)	14,804.0 (5.2)	-5,459.0 (1.2)
Premium	-942.4 (11.0)	-789.9 (9.0)	-	-	25,641.5 (8.6)	3,646.8 (0.7)
Income	.39 (1.8)	.33 (1.5)	-.03 (0.1)	-0.01 (0.0)	-0.63 (2.7)	-0.1 (0.4)
Adults	3.3 (0.0)	51.0 (0.2)	260.0 (1.0)	270.0 (1.1)	249.3 (1.1)	269.8 (1.1)
Faucets	92.4 (0.7)	126.0 (1.0)	301.0 (2.0)	300.5 (2.0)	225.1 (1.7)	290.3 (2.0)
Outages	-1,514.9 (2.0)	-1,338.0 (1.7)	-550.0 (0.6)	-518.0 (0.6)	680.8 (0.8)	-340.6 (0.4)
Electricity	48.8 (4.4)	49.2 (4.4)	49.6 (3.7)	50.1 (3.8)	28.0 (2.3)	47.1 (3.5)
Elasticities						
Ep	-0.45	-0.43	-0.26	-0.27	0.47	-0.17
Epr	0.04	0.03	-	-	-0.62	-0.09
Ey	0.02	0.02	-0.0	-0.0	-0.03	n.s.
R2	0.44 30.90		0.20 11.00		0.20 9.60	

\*\* Using the Location Schedule dummy variable as an instrument.

The presence of simultaneous relationships was also tested with the RESET and Hausman tests. The former is a general specification which may capture problems other than simultaneity while the latter is specifically testing for simultaneity. The results in Table 9 generally support the view that the presence of simultaneity cannot be rejected. Most of the RESET tests indicate the presence of specification problems while all but two of the Hausman tests confirm that simultaneity cannot be rejected.<sup>22</sup>

<sup>22</sup> In Table 9, only the unconstrained models (I and II) using marginal price as the consumer's decision variable can rightly be tested since the introduction of incorrect parameter restrictions as maintained hypotheses would distort the test.

Table 9  
Tests for the Presence of Simultaneity: Small Sample

I. Using the Location Dummy Variable Instrument		
1. Without Sample Selection Correction (Without Lambda)		
Model	I	Critical Value
Reset (2)	12.86	3.84
Reset (3)	11.85	3.00
Reset (4)	8.04	2.60
Hausman	34.57	15.50
2. With Sample Selection Correction (With Lambda)		
Reset (2)	14.67	3.84
Reset (3)	12.02	3.00
Reset (4)	8.15	2.60
Hausman	33.93	15.50
II. Using the Rate Schedule Dummy Variable Instrument		
1. Without Sample Selection Correction (Without Lambda)		
Model	I	Critical Value
Reset (2)	12.86	3.84
Reset (3)	11.85	3.00
Reset (4)	8.04	2.60
Hausman	1.06	15.50
2. With Sample Selection Correction (With Lambda)		
Reset (2)	14.63	3.84
Reset (3)	12.01	3.00
Reset (4)	8.15	2.60
Hausman	4.34	15.50
Note: Numbers in parenthesis after Reset indicate powers of predicted independent variable used in test.		

The results of Lee, Maddala and Trost's adaptation of Heckman's estimation procedure to the simultaneous equations case are presented in Table 10.

Table 10  
Simultaneity and Selectivity Correction: Small Sample with  
Heckman's Two-Step Estimation and Instrumental Variables

(Absolute Value of "t" Statistic in Parentheses)

Model	I	II	III
Constant	24,038.0 (18.6)	19,970.0 (14.8)	19,442.7 (14.9)
Marginal Price	-13,611.0 (11.0)	-8,981.0 (7.4)	-6,586.0 (1.4)
Premium	796.0 (9.0)	- -	2,693.0 (0.5)
Income	0.33 (1.5)	-.03 (0.1)	-0.10 (0.5)
Adults	54.7 (0.2)	266.0 (1.0)	227.8 (1.0)
Faucets	121.0 (0.8)	352.0 (2.4)	318.0 (2.1)
Outages	-133.9 (1.6)	-554.0 (0.6)	-426.1 (0.5)
Electricity	48.6 (4.0)	55.9 (4.1)	52.2 (3.7)
Lambda	-194.0 (1.3)	1,500.0 (1.9)	1,431.1 (1.7)
Elasticities:			
Ep	-0.43	-.28	-0.10
Epr	+ 0.03	-	n.s.
Ey	0.02	n.s.	n.s.

n.s. = Result is not significant.

The results in Table 10 are similar to the results obtained with 2SLS in the small sample.<sup>23</sup> Price elasticities range from -0.10 to -0.43; premium elasticities range from -0.05 to 0.03; and income elasticities range from -0.01 to 0.02. All three models estimated yield theoretically consistent elasticities for price. However, Model I yields positive premium elasticities (which is incorrect), while models II and III yield insignificant (and negative) income elasticities.

In sum, better results were obtained using the subsample of individuals above the minimum consumption under both OLS and 2SLS. Instrumental variables estimation with a sample selection correction on this truncated sample is preferable on a-priori grounds, but did not

<sup>23</sup> For brevity, the first stage Probit estimation results used to produce observation-specific values of the Mills Ratio are not presented. The independent variables included in the Probit step were location dummies, the proportion of income allocated to the fixed water charge, the upper limit of the fixed charge block (in m<sup>3</sup> per month), the number of adults in the household, the average hours of system outage per month, and the monthly electricity bill, a proxy for wealth. Marginal prices were not significant, and were not included.



introduce startling changes in the parameter estimates. The advantage of ignoring simultaneity and sample selection bias by OLS estimation on the truncated sample is simplicity. Its disadvantage is that these problems may be non-trivial in some applications. Also, in many rural areas a large percentage of the population is in the minimum consumption level, and OLS (or 2SLS) estimation on the truncated sample discards information on them by only using families above that level in estimation. The Lee, Maddala, and Trost estimation method makes better use of all the sample information.

### ***Conclusions***

The results of estimating water demand functions from a sample of households connected to public water systems in rural communities in Argentina yield some useful answers to the questions posed at the beginning of this paper.

#### *Simultaneous Equation Bias Exists*

It is clear, both from theory and our Hausman tests, that there is a simultaneous equation bias when using OLS in the face of block-varying tariff schedules. The impact of this bias is stronger when the entire sample spanning the fixed and unit price tariff blocks is used than it is when estimating with the smaller sample that excludes observations in the fixed charge group. But even in the latter case simultaneous equation bias could not be rejected. Therefore, at a minimum, any estimation of water demand should explore an instrumental variables approach (e.g. 2SLS) to try to produce consistent estimates of the price coefficients.

Using OLS on a sample of observations from a single locality and time period that contains both public utility customers and unconnected customers supplied through low volume, higher cost alternatives remedies the problem of unobserved information about low income customers in the fixed charge part of the utility's rate schedule.<sup>24</sup> But, it does not solve the simultaneous equation bias problem. As long as we are confined to a single locality and time period, the use of OLS potentially is as likely to identify a demand as a supply curve, or with luck, some hybrid mixture of both that has a negative price parameter.

#### *Average Price May not be Relevant*

With respect to the choice of average versus marginal prices, the results of the Oppaluch specification are not decisive. But, the results in Tables 8 and 10 suggest that marginal, not average, price is the relevant variable, since marginal price is significant at the 90% level using the critical value for a one-tailed test that does not admit the possibility of a positive marginal price response. The intramarginal premium is in most cases not significantly different from zero. Thus, the average price hypothesis cannot be accepted with our data set.

---

<sup>24</sup> This method is used in Gómez (1987).

### *Sample Selection Bias May Matter*

Sample selection bias may be important, as suggested by the borderline statistical significance of the parameter attached to the Inverse Mills Ratio (the Lambda parameter) in Models II and III of Table 10. Again, however, the results are not clear-cut. Several observations are in order here.

The price coefficients obtained using the Lee, Maddala and Trost estimation method reported in Table 10 do not appear too dissimilar from those obtained using 2SLS in the small sample. Since the Inverse Mills Ratio was estimated independent of prices, the conditional price derivative of consumption (i.e. the price derivative of those within the block rate schedule) is unaffected by its presence in the conditional demand function. But, the estimate of the expected value of water consumed, conditional on being in the rate schedule, is a function of the Inverse Mills Ratio, which can be thought of as an adjustment to the intercept that depends on the household's characteristics.<sup>25</sup>

Depending on the data, the sample selection adjustment could have an impact on the calculation of benefits accruing from a supply price reduction across consumers, since the position of the demand schedule is affected by the value of the Inverse Mills Ratio and the parameter estimate attached to it. In our rate schedule sub-sample, the average value of the Inverse Mills Ratio is 0.5, so the adjustment adds up to about  $0.5 * 1,500 = 750$  cubic meters per month to the constant term in Model II. So corrected, these augmented intercepts are close to the constant terms estimated from 2SLS on the small sample where marginal price can be observed. The hypothetical building of demand schedules for households not in the rate schedule "as if" they faced marginal prices and no fixed charge could bring about a wider disparity between the two approaches.

There are other reasons to prefer Lee, Maddala and Trost's instrumental variables version of the Heckman two-step procedure. It has the advantage of using the information more effectively than 2SLS on the small truncated sample. Finally, the parameter estimates of the 2SLS estimator which ignores incidental truncation introduced by the fixed charge group (i.e. doing 2SLS on the unit block rate customers only) are, at least in theory, inconsistent.

### *What To Do?*

In general, to identify the demand function, data on more than one locality in a single time period or more than one time period at a single locality are needed. Individuals in the public utility's fixed charge category should not be included in the estimation sample, since their marginal willingness to pay cannot be accurately observed, and is not necessarily zero.

---

<sup>25</sup> The expected value of any individual's quantity demanded  $q_i$ , given the values of its row vector of regressors  $\mathbf{x}_i$ , and conditional on being in the rate schedule (where  $\mathbf{B}$  is a row vector of parameter estimates on variables other than  $\mathcal{Q}$ ) is:

$$E[q_i | \mathbf{x}_i, q_i \in \mathcal{Q}] = \mathbf{x}_i' \mathbf{B} + \beta_{\mathcal{Q}} \mathcal{Q}$$

For the estimation sample it makes sense to augment the truncated sample of public utility customers in the unit price blocks of the rate schedule with information on unconnected households using private (or self) supply sources. Preferably, their characteristics (particularly income, assets, and water consumption levels) should be similar to those of the families in the public utility's fixed charge category that have to be dropped from the demand estimation sample. In this way, legitimate high price/low quantity observations can be obtained. But, this group of unconnected customers is akin to a separate locality that, at least potentially, faces unit prices that depend on consumption quantity, so the potential for simultaneous equations bias in such an augmented sample combining connected block rate households and unconnected households still exists.

The likely severity of the simultaneity and sample selection problems can be assessed by a close look at the data. For example, if a small proportion of the sample falls into the fixed charge regime (or a fixed charge regime does not exist in the rate structures observed across space or time) the sample selection correction may be unnecessary. If an overwhelming majority of the subsample in the rate schedule falls into the first unit price block, or the place (or time) varying schedules all have constant rates rather than block-declining or block-increasing rates, then for all practical purposes, simultaneity is a non-issue. If both of these conditions are satisfied by the data, OLS may be acceptable.

If there is a fixed charge regime with constant unit pricing rather than block pricing, simultaneity is no problem, but sample selection bias might be. Estimation on the truncated sample should attempt to account for it, either using Heckman's correction method, or an approximation using an adjustment to OLS (e.g. see Greene 1983, Olsen 1980).<sup>26</sup>

If there is no fixed charge regime, but the observations are arrayed across a wide range of different block-defined unit prices, an instrumental variables estimator like 2SLS would be preferable. The need for it can be established via a Hausman test.

If it appears that simultaneity and sample selection both are present, at a minimum we suggest that the analyst should disregard the fixed charge customers and use an instrumental variables estimator on the in-block household observations.

---

<sup>26</sup> Assume we discard all the information pertaining to the subsample of observations in the fixed charge regime (Regime 2) and retain only those in the rate regime (Regime 1), making the sample truncated rather than censored. Also assume that the observed marginal price  $P_{\text{marg}}$  is constant rather than defined as an increasing or decreasing block rate, so simultaneity is not an issue. Suppose some regressors appear in the sample selection equation (Equation 16 above) but not in the demand equation (Equation 14). If these variables are included in the estimation of Equation 18 on the truncated sample, according to Greene 1983, consistent parameter estimates will be obtained for variables which belong only in the demand equation and not in the selection equation. Variables shared by both equations will still have inconsistent parameter estimates. If the demand equation and the selection equation do not have any variables in common, inclusion of the variables from the selection equation in the demand equation estimation will produce consistent parameter estimates for the latter. This result from Greene (1983) is linked to Heckman's intuition (1976) that the effect of sample selection is that variables that do not belong in the regression appear to be statistically significant in equations fit on selected samples. It is of little help, because a-priori knowledge about which variables belong in each stage of the decision process is usually not available.

In sum, the characteristics of the particular case dictate the degree of econometric sophistication required for the estimation of plausible household water demand functions. Nevertheless, potential estimation problems should not be ignored altogether by the indiscriminate use of simple OLS when its assumptions do not apply.

## **References**

- Billings, Bruce E. and Donald E. Agthe. 1980. Price Elasticities for Water: A Case of Increasing Block Rates. *Land Economics*, Vol. 56, No.1.
- Blattenberg, Gail R. 1977. Block Rate Pricing and the Residential Demand for Electricity. Phd. Dissertation. Ann Arbor, Michigan: University Microfilms International.
- Bohi, Douglas R. 1981. *Analyzing Demand Behavior, A Study of Energy Elasticities*. Baltimore: The Johns Hopkins University Press, Resources for the Future.
- Bohi, Douglas and Mary B. Zimmerman. 1984. An Update on Econometric Studies of Energy Demand Behavior. *Annual Reviews Energy* 1984-9.
- Charney, Alberta H. and Gary C. Woodard. 1984. A Test of Consumer Demand Response to Water Prices: Comment. *Land Economics*, Vol. 60, No. 4.
- Chicoine, David L. and Ganapathi Ramamurthy. 1986. Evidence on the Specification of Price in the Study of Domestic Water Demand. *Land Economics*, Vol 62, No. 1.
- Dutta, M. 1975. Econometric Methods. Brighton, England: South-Western Publishing Co.
- Foster, Henry S. and Bruce R. Beattie. 1979. Urban Residential Demand for Water in the United States. *Land Economics*, Vol. 55, No. 1.
- Gómez, Christian. 1987. Experiences in Predicting Willingness to Pay for Water Projects in Latin America, Washington, D.C.: Inter-American Development Bank, Project Analysis Department.
- Greene, William H. 1983. Estimation of Limited Dependent Variable Models by Ordinary Least Squares and the Method of Moments. *Journal of Econometrics*, Vol. 21, pp. 195-212.
- \_\_\_\_\_. 1990. Econometric Analysis. New York: Macmillan.
- Griffin, Adrian H. and William E. Martin. 1981. Price Elasticities for Water: A Case of Increasing Block Rates: Comment. *Land Economics*, Vol. 57, No.2.
- Hanke, Steve H. 1978. A Method for Integrating Engineering and Economics Planning. *Journal of the American Water Works Association*. September, pp.487-91.
- Hausman, J.A. 1978. Specification Tests in Econometrics. *Econometrica*, Vol. 46, No. 6, pp. 1251-1270.
- Heckman, James J. 1976. The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimate for Such Models. *Annals of Economic and Social Measurement*, Vol. 5, No. 4, pp. 475-492.
- Heckman, James J. 1979. Sample Selection Bias as Specification Error. *Journal of the Econometric Society*, Vol. 47, No. 1, pp. 153-161.

- Heckman, J. 1979. Sample Selection Bias as a Specification Error. *Econometrica*, Vol. 97, pp. 155-162.
- Lee, Lung-Fei and Robert P. Trost. 1978. Estimation of Some Limited Dependent Variable Models with Application to Housing Demand. *Journal of Econometrics*, No. 8, pp. 357-382.
- \_\_\_\_\_, G. S. Maddala, and Robert P. Trost. 1980. Asymptotic Covariance Matrices of Two Stage Probit and Two Stage Tobit Models for Simultaneous Models with Selectivity. *Econometrica*, Vol. 48, pp. 491-504.
- Lee, Terence R. 1969. *Residential Water Demand and Economic Development*. Toronto, Canada: University of Toronto Press.
- Maddala, G.S. 1983. *Limited - Dependent and Qualitative Variables in Econometrics*, New York: Cambridge University Press.
- Martin, William E., Helen M. Ingram, Nancy K. Laney, and Adrian H. Griffin. 1984. *Saving Water in a Desert City*. Baltimore: The Johns Hopkins University Press, Resources for the Future.
- Meroz, Avigdor. 1968. A Quantitative Analysis of Water Demand in Developing Countries. International Bank for Reconstruction and Development, Economics Department Working Paper No. 17.
- Mullahy, John. 1987. "Economic Estimation of Microeconomic Recreation Participation Models" in *Economic Perspectives*. N.J. Rau, editor. London: University College, pp. 91-149.
- Nakamura, Alice and Masao Nakamura. 1981. On the Relationships Among Several Specification Error Tests Presented by Durbin, Wu, and Hausman. *Econometrica*, Vol. 49, No. 6., pp. 1583-1588.
- Nieswiadomy, Michael L. and David J. Molina. 1991. A Note on Price Perception in Water Demand Models. *Land Economics*, Vol. 67, No. 3.
- Nording, John. 1976. A Proposed Modification of Taylor's Demand Analysis: Comment. *The Bell Journal of Economics*, 7.
- Olsen, Randall J. 1980. Approximating a Truncated Normal Regression with the Method of Moments. *Econometrica*. Vol. 48, No. 5. July.
- Oppaluch, James J. 1981. Block Pricing and Consumer Behavior. University of Rhode Island, Department of Resource Economics Staff Paper # 81-20.
- Oppaluch, James J. 1984. A Test of Consumer Demand Response to Water Prices: Reply. *Land Economics*, Vol 60, No. 4.

- Powers, Terry A. and Carlos A. Valencia. 1980. SIMOP Urban Water Model. Washington, D.C.: Inter-American Development Bank, Papers on Project Analysis No. 5.
- Sonnenschein, H. 1973. Do Walras' Identity and Continuity Characterize the Class of Community Excess Demand Functions? *Journal of Economic Theory*, 6.
- Schefter, J.E. and E. L. David. Estimating Residential Water Demand under Multi Part Tariffs Using Aggregate Data. *Land Economics*, Vol 61, No. 3.
- Taylor, Lester D. 1975. The Demand for Electricity: A Survey. *The Bell Journal of Economics*. 6, No. 1 (Spring).
- \_\_\_\_\_, G. R. Blattenberger, and R. G. Rennhack. 1981. The Demand for Energy in the United States. Report to the Electric Power Research Institute RP 1098. Lexington, Massachusetts: Data Resources Inc.
- \_\_\_\_\_. 1982. "Some Problems and Issues in Building Econometric Models of Energy Demand", in *Perspectives of Resource Policy Modeling: Energy and Minerals*. Raphael Amit and Mordecai Avriel, editors. Cambridge Mass.: Harper, Row, and Ballinger.
- Vaughan Jones C., John J. Boland, James E. Crews, Frederick C. DeKay, and John R. Morris. 1984. *Municipal Water Demand, Statistical and Management Issues*. Boulder: Westview Press.
- Wade, Steven Howard. 1980. *The Implications of Decreasing Block Pricing for Individual Demand Functions: An Empirical Approach*. Ph.D. Dissertation, University of Arizona.
- Westley, Glenn. 1981. "The Residential and Commercial Demand for Electricity in Paraguay", Papers on Project Analysis No. 19, IDB, Washington D.C.
- Westley, Glenn. 1984. "The Residential and Commercial Demand for Electricity in Costa Rica", Papers on Project Analysis No.24. Washington D.C.: Inter-American Development Bank.
- Williams, Martin. 1985. Estimating Urban Residential Demand for Water under Alternative Price Measures. *Journal of Urban Economics*, 18.
- \_\_\_\_\_ and Byung Suh. 1986. The Demand for Urban Water by Customer Class. *Applied Economics*, 18.