Determinants of Demand for Water Used in Texas Communities

by

David R. Bell

and

Ronald C. Griffin

Department of Agricultural Economics
Texas A&M University
College Station, TX 77843-2124

July 2005

for presentation at

Annual Meeting of the
Western Agricultural Economics Association
San Francisco, July 6-8, 2005

Determinants of Demand for Water Used in Texas Communities

Abstract

A panel of monthly water demand data for Texas communities is analysed for sensitivity to price specification, representation of seasonality, functional form, and estimation assumptions. Parameter estimation is improved by allowing for cross-sectional heteroskedasticity and time-series autocorrelation.

Introduction

There is now a wealth of econometric evidence pertaining to urban water demand. Studies have become so numerous that they offer a solid platform for conducting meta-analyses (Espey, Espey, and Shaw 1997; Dalhuisen et al. 2003). Yet, even though prior studies cover a wide geographic territory as well as multiple decades, there is continued interest in methodological refinements and updated information. The primary reason is rising water scarcity. Due to rising scarcity and the important public duties that arise in water management, demand estimation has the power to reveal and retune important parameters. The most important of these parameters may be the price elasticity of demand (Arbués, García-Valiñas, and Martínez-Espiñeira 2003). Though a simple construct in and of itself, such elasticity estimates have become central elements of planning tools and models. Water demand elasticities are now used in highly detailed allocation models. They are used to appraise water rights. They are used in city rate studies. They are key elements of cost benefit analyses of proposed projects, and they are similarly used in policy analyses. Because a great deal is at stake within these applications, there is ample justification for the continued empirical study of water demand. Here, we present some preliminary results from a dataset under construction, including tests of the perennial issues of price specification, functional form, and seasonality, as well as an exploration of gains of precision to be made from heteroskedasticity and nonstationarity assumptions, and discussion of the elasticity estimates of selected models. A thorough analysis of community water demand has not been performed in Texas for nearly 20 years.

Monthly water usage (thousands of gallons) and residential water and wastewater rates (per 1,000 gallons) are reported for 1999-2002 by municipal and autonomous utility systems and water districts. Quantity data is reported to the Texas Water Development Board, and price data is obtained from a questionnaire exclusively for this study. An effort is made to confine the inquiry to suppliers of the residential sector; nevertheless, an assumption is made that the amount of water reported used by the systems is at least proportional to the quantity demanded by households in each service area.

The price data obtained is a typically uniform or increasing block rate schedule applying to users of the smallest or single-residence type meter, and only to those inside city limits where a distinction is made. Price variables under consideration are summary indicies of the price information given. To embrace a wider definition of price, an average marginal price specification is tested alongside marginal price and average price. Most of the sample's rate structures are increasing with volume, so that the average marginal price is typically lower than the marginal price. Previous literature has tended to support average price as the more important determinant of consumption choice (Foster and Beattie 1981; Arbués, García-Valiñas, and Martínez-Espiñeira 2003). Average price is the price signal obtained at least cost by consumers (Gaudin 2005), and it reflects an expenditure effect that is not captured in the marginal price specifications (Arbués, García-Valiñas, and Martínez-Espiñeira 2003). Where consumers are billed for sewerage on the same water volume and water bill, the two prices are combined. An indicator variable is provided to quantify the impact of sewer charges. The variable "SEWER" is unity if wastewater usage is billed on the water bill. Although the price variables represent the actual nominal cost of water purchased, consumers may not experience water and wastewater charges identically.

Size of each system is measured as the number of connections, or meters, served, and included as an independent regressor to gauge the urbanization level of each community.

Frequency of precipitation (over 0.25 of an inch per day) and temperature (average daily high and low ^{o}F) data from local weather stations is obtained from the National Climatic Data Center. These data are transformed into a composite climate index for the empirical models. Personal income data (mean per capita annual income) for each community is obtained from the Bureau of Economic Analysis. Table 1 provides a summary of the sample of 48 communities over 48 months, representing over one million water users.

Least Squares Models

The general model of quantity demanded considered here is Q(P, C, I, n), where exogenous variables include price, climate,income, and the size of the system. An indicator of the presence of volumetric sewer charges on the water bill is included as a adjunct to price in each model. Meter charge is included as a separate variable alongside marginal price specifications but not average price, because it is included in the average price formula. Meter charges are the fixed price of access to utility water, for which water itself is logically a complement, although it is clearly a gross substitute in these regressions. Utility systems may choose between metering all water consumption and including some consumption with the monthly flat fee. Substitution only actually occurs if a household moves from one community to another with a different rate structure, but this positive demand effect of fixed price suggests potential pressure toward substitution. Further research is necessary to the clarification of this issue. Nevertheless, explanatory power seems to be gained from the inclusion of this variable.

Climate is modeled as an index of mean temperature and the likelihood of precipitation. Table 2 is the correlation matrix of independent variables, and Table 3 lists variance inflation factors for each regressor. Belsley, Kuh, and Welch (1980) suggest that a VIF > 30 for any regressor indicates the presence of degrading collinearity. This rule of thumb is not met by any of the included variables, but each price metric will be tested in separate regressions to reduce the effect of covariance.

We present linear, semilog, Cobb-Douglas, Generalized Cobb-Douglas, and Stone-Geary specifications side-by-side. Model selection is driven by two considerations: indications of a good fit to the data, and the theoretical ability of the functional form to provide a characterization of desirable parameters, specifically nonconstant elasticities. The linear form is a benchmark, but is least preferred in fit. The semilog form is useful because of the flexibility of elasticity measurement (Gaudin, Griffin, and Sickles 2001; Arbués, Barberán, and Villanúa 2004), and it fits better than the linear model. The Cobb-Douglas form allows easy comparisons with the literature, scalar estimates of price and income elasticity of demand (Taylor, McKean, and Young 2004), and fits well, but is least flexible in elasticity estimation. This trait is especially worrisome in the modeling of seasonality. The generalized Cobb-Douglas form provides a high-flexibility counterpoint to the other forms examined here (Griffin and Chang 1990). The Stone-Geary form provides a special interest in predicting a subsistence level of demand, embodied in the popular notion of a city's "need for water" (Gaudin, Griffin, and Sickles 2001).

Table 4 shows estimation results for linear, semilog, Cobb-Douglas, Generalized Cobb-Douglas, and Stone-Geary OLS regressions. Student t-scores are given below the parameter estimates. In the Stone-Geary specification, the ratio of income to price is calculated for each potential price metric instead of price and income variables. Explanatory power is largely comparable among the models, as approximated by the adjusted-R^2 statistic. Akaike and Bayesian information criteria are provided as additional, cardinal measures of fit, although neither implies a statistical hypothesis (Akaike 1981). Both ascribe lower values to more efficient models, but only compare models with the same dependent variable. The three measures of fit given agree in favoring the marginal price specification, although in general the signs of the parameter estimates appear to be robust to functional form. Adjusted-R^2 statistics in the range of 0.3 are neither uncommon nor uninformative for cross-sectional studies of this breadth, but the inclusion of additional structural information can greatly elevate this level of certainty, as we will see.

The Stone-Geary demand function is

$$Q = \gamma + \beta \frac{I^* - P\gamma}{P} \tag{1}$$

where γ is a water use threshold (Gaudin, Griffin, and Sickles 2001). This form suffers from a lack of conventional applicability, but it contributes a distinctive hypothesis that is worth discussing briefly. The linearization

$$Q = \gamma + \beta \frac{I^*}{P} \tag{2}$$

approximates a "subsistence level" of demand below which price elasticity is stipulated to be zero. Estimating γ at the sample means from the marginal price regression yields a value of 6,000, meaning that 6,000 gallons per household are completely unresponsive to price signals. This hypothesis remains untested. We set aside the linear and Stone-Geary forms for the remainder of this discussion.

The intuitive draw of the generalized variant of the Cobb-Douglas form (Griffin and Chang 1990),

$$\ln Q = \sum \beta_i \ln X_i + \sum \sum \beta_{ij} \frac{\ln(X_i + X_j)}{2} \ , \tag{3}$$

is its lack of structure. This is borne out as it consistently outperforms the other models on the basis of fit. Elasticity estimates from this form are sensitive to heteroskedasticity, so the independent variables are scaled by their standard deviations prior to inclusion in the model. Parameter estimates in the Generalized Cobb-Douglas model resist simplistic interpretation, so they are not presented in Table 4. Elasticity estimates drawn from a more elaborate treatment of this form are presented later.

Seasonality is modeled as a function of the meteorological statistics average nightly low temperature, average daily high temperature, and proportion of days with rainfall (Griffin and Chang 1990). These factors are combined in a climate index:

$$CI = \left(\frac{\max Temp + \min Temp}{2} \right) * (1 - Frequency\_of\_Rain) \tag{4}$$

The use of monthly indicator variables to model seasonality produces a comparable effect. We conclude that seasonality is adequately captured by our climate statistic and dispense with individual weather indicators. Table 5 presents constant marginal price elasticity estimates by month, illustrating the important point that price elasticity of water demand is not constant with price.

Heteroskedasticity And Autoregression

The sample is heterogeneous, being a broad sample of communities in Texas, and the imposition of homoskedastic residuals may be far-fetched. Intuition also suggests that water consumption decisions might be habitual, therefore not stationary in time. Use of the data as a panel with random effects parameters provides the opportunity to test these hypotheses and to refine coefficient estimates based on the results (Arbués, Barberán, and Villanúa 2004). An iterated generalized least squares procedure is offered against the ordinary least squares regressions with comparative information criteria to explore potential gains in precision.

Cross-sectional samples are susceptible to heteroskedastic residuals across sections. The Breusch-Pagan test for constant variance was conducted on fitted values of the regression. Distributed as a Chi-squared statistic of degree one, the results were 11.83 ($p = 0.0006$) for the Cobb-Douglas specification and 12.89 ($p = 0.0003$) for the semilog, thereby questioning the homoskedasticity assumption of ordinary least squares.

Testing the stationarity of the time series component requires a panel treatment of the dataset. Assuming that heteroskedasticity is confined to variations between communities, and allowing an AR(1) autoregressive process unique to each community, fixed effects parameters are derived iteratively for the marginal price versions of the three top models, the semilog, Cobb-Douglas, and generalized Cobb-Douglas forms. The resultant information criteria and elasticity estimates are summarized in Table 6.

The information criteria indicate that these are powerful estimates. Since the number of regressors in these latter regressions is comparable to that in the previous regressions, the gains come from explanatory power. Although the results of this iterative procedure are superior to those from OLS, the coefficient estimates are not vastly different. Even so, three important insights are gained: the cross-sections of this sample are heteroskedastic; the time series are not all stationary; and the parameters estimated in this way fit significantly better than those obtained through OLS. Potential gains from a similar generalization of the Stone-Geary specification cannot be discounted, but they are not explored here.

Conclusions

Table 6 lists price elasticities of -0.225 to -0.258. Those figures for the Cobb-Douglas form are constant; they are calculated at the sample means for the other forms. It is tempting to say that the price elasticity of water demand in Texas is -0.24. As shown in Table 5, though, price elasticity is not constant from month to month. Lower elasticities in times of greater demand, for example summer months, imply an elasticity that is variant with the independent variables. The semilog form can model this behavior if the increase is linear and simple, but it lacks the interaction necessary to capture subtler relationships. As management policies and projections increase in value, the simplicity of a scalar or linear elasticity estimate is likely to lose its appeal for both modelers and managers, who demand more realistic figures. The technology of water demand analysis has advanced beyond such inflexible forms.

The generalized logrithmic functional form favored by this analysis is flexible, but almost amorphous. Its many elements are computationally cumbersome and difficult to interpret without additional vision into the dynamics of the demand system for water. Each new study raises new questions and new opportunities for research. Meta-analysis offers only a hint of the potential dynamical properties of water demand and is no substitute for ever better, more exacting studies. The rise of the panel as a method of sorting spatial and temporal effects is under-utilized in this field, with the potential to add a revolutionary degree of clarity and confidence to empirical demand models. We will attempt to exploit this potential with a large forthcoming study from which the data discussed here is a small sample.

References

Akaike, Hirotugu. "Likelihood of a Model and Information Criteria." *Journal of Econometrics* 16 (May 1981): 3-14.

Arbués, Fernando, Ramón Barberán, and Inmaculada Villanúa. "Price Impact on Urban Residential Water Demand: A Dynamic Panel Data Approach." *Water Resources Research* 40 (2004): W11402, doi:10.1029/2004WR009092.

Arbués, Fernando, María Ángeles García-Valiñas, and Roberto Martínez-Espiñeira. "Estimation of Residential Water Demand: A State-of-the-Art Review." *Journal of Socio-Economics* 32 (March 2003): 81-102.

Belsley, David A., Edwin Kuh, and Roy E. Welch. *Regression Diagnostics: Identifying Influential Data And Sources Of Collinearity*. New York: John Wiley, 1980.

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

Espey, Molly, James Espey, and W. Douglass Shaw. "Price Elasticity of Residential Demand for Water: A Meta-Analysis." *Water Resources Research* 33 (June 1997): 1369-74.

Foster, Henry S., Jr., and Bruce R. Beattie. "On the Specification of Price in Studies of Consumer Demand under Block Price Scheduling." *Land Economics* 57 (November 1981): 624-29.

Gaudin, Sylvestre. "Water Bills and the Power of Price Signals: Evidence from the U.S." Presented at *IWA International Conference on Water Economics, Statistics, and Finance*, Rethymno, Greece, July 8-10, 2005.

Gaudin, Sylvestre, Ronald C. Griffin, and Robin C. Sickles. "Demand Specification for Municipal Water Management: Evaluation of the Stone-Geary Form." *Land Economics* 77 (August 2001): 399-422.

Taylor, R.G., John R. McKean, and Robert A. Young. "Alternate Price Specifications for Estimating Residential Water Demand with Fixed Fees." *Land Economics* 80 (August 2004): 463-75.

TABLE 1
SUMMARY STATISTICS

N = 2304

VARIABLE	MEAN	STANDARD DEVIATION	MINIMUM	MAXIMUM
Volume	1.480E+08	7.030E+08	1.095E+06	7.310E+09
Population	32169.34	158585.80	812.00	1144646.00
Connections	9073.46	42148.30	360.00	300854.00
Volume per Connection	12948.06	7547.48	1684.00	71431.04
Rainfall Frequency	0.108	.08	0.00	.467
Low Temperature	55.82	13.88	16.34	79.00
High Temperature	78.33	12.92	42.06	104.20
Climate Index	59.88	13.42	28.26	89.63
Personal Income	25961.98	5651.88	18153.00	36047.00
Marginal Price	3.46	1.71	0.90	8.24
Average Marginal Price	3.52	1.73	0.86	8.24
Average Price	2.81	1.36	0.47	7.61
Fixed Price	37.47	25.56	4.65	110.00

TABLE 2
CORRELATION MATRIX

	MARGINAL PRICE	AVERAGE MARGINAL PRICE	AVERAGE PRICE	FIXED PRICE	SEWER BILLED	CLIMATE COMPOSITE	INCOME	CONNECTIONS
Marginal Price	1.00							
Average Marginal Price	0.96	1.00						
Average Price	0.88	0.92	1.00					
Fixed Price	0.31	0.34	0.36	1.00				
Sewer Billed	-0.02	0.002	0.002	-0.43	1.00			
Climate Composite	-0.05	-0.06	0.002	0.06	-0.10	1.00		
Income	0.12	0.14	0.17	0.06	0.09	0.04	1.00	
Connection	-0.11	-0.17	-0.13	0.01	-0.10	0.04	0.08	0.84

TABLE 3

VARIANCE INFLATION FACTORS

Variable	VIF	1/VIF
Average Marginal Price	20.39	0.049
Marginal Price	13.60	0.074
Average Price	6.85	0.146
Fixed Price	1.49	0.671
Sewer Billed	1.33	0.754
Connections	1.10	0.907
Income	1.06	0.943
Climate	1.04	0.961
Mean VIF	5.86	

TABLE 4
ESTIMATED COEFFICIENTS OF LEAST-SQUARES REGRESSIONS WITH t-SCORES

N = 2304		VARIABLE PRICE	FIXED PRICE	SEWER	CLIMATE	INCOME	CONNECTIONS	CONSTANT	ADJUSTED R-SQUARED	AIC	BIC
Linear	Marginal	-1574.01	31.58	1563.50	169.02	0.39	5.5E-03	-3736.01	0.30	46923	46883
		(-18.97)	(5.18)	(5.19)	(17.01)	(16.21)	(1.72)	(-4.2)			
	Average Marginal	-1580.33	37.01	1770.87	167.63	0.40	1.6E-03	-4022.70	0.29	46935	46895
		(-18.61)	(5.95)	(5.84)	(16.82)	(16.49)	(0.5)	(-4.53)			
	Average	-976.57		1054.66	181.35	0.38	7.9E-03	-5550.22	0.21	47167	47133
		(-9.29)		(3.71)	(17.33)	(14.93)	(2.35)	(-6.07)			
Semilog	Marginal	-0.10	2.6E-03	0.10	1.2E-02	2.2E-05	8.5E-07	8.25	0.31	2439	2398
		(-18.88)	(6.75)	(5.24)	(18.91)	(14.29)	(4.14)	(144.5)			
	Average Marginal	-0.10	2.9E-03	0.11	1.2E-02	2.2E-05	6.1E-07	8.22	0.30	2463	2423
		(-18.1)	(7.33)	(5.82)	(18.69)	(14.49)	(2.95)	(143.72)			
	Average	-3.4E-02		0.06	1.3E-02	2.0E-05	1.1E-06	8.10	0.21	2740	2705
		(-4.91)		(3.08)	(18.93)	(12.37)	(5.14)	(136.27)			
Cobb-Douglas	Marginal	-0.34	0.09	0.07	0.65	0.43	0.08	1.78	0.35	2289	2249
		(-19.69)	(5.8)	(3.82)	(18.18)	(10.38)	(11.86)	(4.15)			
	Average Marginal	-0.30	0.07	0.08	0.65	0.44	0.08	1.70	0.32	2375	2334
		(-17.03)	(4.61)	(4.16)	(17.83)	(10.36)	(11.36)	(3.9)			
	Average	-0.09		0.05	0.69	0.35	0.11	2.23	0.25	2617	2583
		(-5.05)		(2.52)	(18.15)	(7.88)	(14.01)	(4.87)			
Stone Geary	Marginal		54.54	1700.73	172.50	0.67	6.6E-03	-6850.73	0.39	46581	46547
			(9.81)	(6.17)	(18.71)	(31.77)	(2.25)	(-10.11)			
	Average Marginal		49.16	2056.30	172.63	0.62	-3.9E-03	-6106.94	0.32	46820	46786
			(8.37)	(7.07)	(17.77)	(26.15)	(-1.23)	(-8.52)			
	Average			1296.97	190.80	0.29	8.9E-03	-2600.46	0.20	47215	47186
				(4.53)	(18.04)	(14.37)	(2.6)	(-3.64)			
Generalized Cobb-Douglas	Marginal	*							0.48	1887	1761
	Average Marginal	*							0.45	1993	1867
	Average	*							0.33	2406	2315

* Individual parameter estimates are not revealing in this functional form

TABLE 5
MARGINAL PRICE ELASTICITY ESTIMATES BY MONTH
FROM THE COBB-DOUGLAS FUNCTIONAL FORM

	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
Elasticity	-0.19	-0.20	-0.25	-0.35	-0.44	-0.44	-0.50	-0.50	-0.43	-0.38	-0.29	-0.23
t-Score	-3.00	-3.33	-4.31	-5.91	-7.7	-7.05	-8.75	-8.25	-7.13	-6.41	-5.02	-4.03

TABLE 6
Results of Iterated Generalized Least Squares Regressions Assuming Heteroskedasticity and AR(1)

	PRICE ELASTICITY	INCOME ELASTICITY	LOGLIKELIHOOD	AIC	BIC
Semilog	-0.232	0.519	846.6	-1679	-1639
Cobb-Douglas	-0.258	0.413	803.5	-1593	-1553
Generalized C-D	-0.225	0.175	947.6	-1851	-1725