ECONOMETRIC ANALYSES OF PUBLIC WATER DEMAND

IN THE UNITED STATES

A Dissertation

by

DAVID RAY BELL

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2011

Major Subject: Agricultural Economics

Econometric Analyses of Public Water Demand in the United States

ECONOMETRIC ANALYSES OF PUBLIC WATER DEMAND

IN THE UNITED STATES

A Dissertation

by

DAVID RAY BELL

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Ronald Griffin
Committee Members,	Kelly Brumbelow
	Oral Capps Jr.
	Bruce McCarl
Head of Department,	John Nichols

December 2011

Major Subject: Agricultural Economics

ABSTRACT

Econometric Analyses of Public Water Demand in the United States. (December 2011)

David Ray Bell, B.A., University of Texas at Austin

Chair of Advisory Committee: Dr. Ronald C. Griffin

Two broad surveys of community-level water consumption and pricing behavior are used to answer questions about water demand in a more flexible and dynamic context than is provided in the literature. Central themes of price representation, aggregation, and dynamic adjustment tie together three econometric demand analyses. The centerpiece of each analysis is an exogenous weighted price representation.

A model in first-differences is estimated by ordinary least squares using data from a personally-conducted survey of Texas urban water suppliers. Annual price elasticity is found to vary with weather and income, with a value of -0.127 at the data mean. The dynamic model becomes a periodic error correction model when the residuals of 12 static monthly models are inserted into the difference model. Distinct residential, commercial, and industrial variables and historical climatic conditions are added to the integrated model, using new national data. Quantity demanded is found to be periodically integrated with a common stochastic root. Because of this, the structural monthly models must be cointegrated to be consistent, which they appear to be. The error correction coefficient is estimated at -0.187. Demand is found to be seasonal and slow to adjust to shocks, with little or no adjustment in a single year and 90% adjustment taking a decade or more. Residential and commercial demand parameters are found to be indistinguishable.

The sources of price endogeneity and historical fixes are reviewed. Ideal properties of a weighted price index are identified. For schedules containing exactly two rates, weighting is equivalent to a distribution function in consumption. This property is exploited to derive empirical weights from the national data, using values from a nonparametric generalization of the structural demand model and a nonparametric cumulative density function. The result is a generalization of the price difference metric to a weighted level-price index. The validity of a uniform weighting is not rejected.

The new approach of weighted price indexing is data intensive, but the payoff is increased depth and precision for the economist and accessibility for the practitioner.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Griffin, and my committee members, Dr. Brumbelow, Dr. Capps, and Dr. McCarl, for their guidance and support throughout the course of this research. I would like to thank the Texas Water Development Board (Contract 2004-483-023) and the United States Geological Survey (Grant 06HQGR0188) for supporting my research. I would like to thank the journals *Water Resources Research* and *Land Economics* for allowing copyrighted material to be reproduced here. I would like to thank all of those who helped to collect and process the survey data and all who responded to our queries.

And Georgia.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

Freshwater withdrawals of 44.2 billion gallons are supplied daily to the public of the United States (Barber 2009). Public supply includes treated and pressurized water delivered to households, businesses, and institutions. How does a change in water quantity resulting from a new policy or project affect the value received by these users? What are the conservation and revenue implications of pricing changes? Developing answers to these and other questions requires a demand function, yet the nature of water demand is quirky, posing unique modeling challenges.

The existence of a demand function for water, where price influences consumption, was not widely accepted before 1967, when Howe and Linaweaver published the first national study of residential water demand. Their finding of significant price effects challenged the use of one-size-fits-all coefficients to approximate daily per capita water use. In that era of industrial expansion, Howe and Linaweaver focused on the implications of their results for efficient plant design. If consumption is invariant to price, optimal sizing of a water facility is simply a matter of providing a quantity of water equivalent to the estimated use coefficient times the estimated service population. Cost sufficiency can then be achieved by billing consumption on the basis of average cost. If, on the other hand, consumption is a function of price, then plants must be designed to deliver the quantity demanded by the service population at a price level that

---

This dissertation follows the style of *Land Economics*.

covers costs (Howe and Linaweaver 1967). The recognition of feedback between price and quantity introduces an economic dimension to a problem traditionally dominated by engineering solutions.

Hundreds of water demand studies have followed Howe and Linaweaver, providing persuasive evidence of the existence of a price effect on consumption (Dalhuisen et al. 2003). Meanwhile, the landscape of water management has changed. In 2011, water production infrastructure is more abundant than it was in 1967, and undeveloped water supplies are scarcer. The benefits of supply enhancement have been largely exploited in many parts of the country (Ward et al. 2006), and benefits of demand management may hold more promise for the future (Renzetti 1992).

Demand analysis for infrastructure planning does not emphasize timing because dams and reservoirs take years to build and operate for decades without rescaling. Demand management with pricing tools, in contrast, is dynamic with annual (or interannual, Hausman et al. 1979) adjustment opportunities. During a drought, for instance, water managers may wish to shrink excess demand by raising rates. Supply-side solutions are inadequate for this purpose because even a serious drought will have abated before new supplies could be developed. The quicker pace of demand management calls for a dynamic element that is largely absent from existing demand models. The models introduced in this dissertation ask not only what the ultimate price effect on consumption will be, but also when price effects will occur. The parameters of a water demand model with dynamic adjustment can inform rate-setters of the effects their policies will have over a given interval, such as a fiscal year. This information can contribute to efficiency, budgetary, and social gains.

Demand at its most basic is the relationship between quantity consumed and price. Most texts and articles on the subject of demand take price to be a primitive and self-evident signal known to all parties. No measure exists in the arena of water demand, though, that can be clearly and unambiguously called a price. Water is transacted at many rates, none of which is known to every consumer. This ambiguity is mostly due to the low value that water managers place on transparency. Complicated multi-block rate schedules and poor communication ensure that consumers perceive the true marginal price of water dimly at best. There is a disconnect between the signal issued by the producer and that received by the consumer. The producer's signal is observable and public, but unwieldy. A quantity-independent consumer price signal is assumed by demand theory but unobserved. The concept of demand is only useful and believable if the two signals somehow coincide. Researchers have struggled and continue to struggle with a definition of water price that represents observable values while maintaining statistical independence. A central objective of the present research is to advance this endeavor.

Most intuitive representations of price under block rate schedules and/or dim price perception result in inconsistent econometric estimation. Some mitigation techniques are reviewed in Chapter II, and a new quasidifference representation of price is introduced. The quasidifference price is statistically independent of quantity consumed, but it only supports estimation of a short-run model. The first of two major datasets developed for this dissertation is introduced and used to estimate a demand model in first differences. The data consist of price schedules obtained for 734 utility systems surveyed by the author, augmented with quantity data from the Texas Water Development Board and economic and climatic data from national data sources. Flexibility is provided in the empirical model to allow for interactions between pairs of covariates. Price elasticity is shown to be variable when the functional form of demand allows for variation.

The quasidifference model is expanded into an error correction (EC) model in Chapter III. Here, the model from Chapter II is augmented with a regressor representing the lagged residual of steady-state demand. By incorporating both dynamic and structural components, the EC model allows consistent simultaneous estimation of demand effects at multiple time scales. A seasonal component is incorporated as well, by estimating a different steady-state demand equation for each calendar month and including only the month-specific lagged residual in the EC model. Seasonal unit-root tests reject seasonal nonstationarity in the residual series, assuring statistical consistency in the EC model.

The second original and major dataset is introduced in Chapter III and used to estimate the empirical demand model. The data consist of over 16,000 observations on 167 municipal utility systems. Price data are obtained from a survey by the author of over 1000 systems nationwide. Quantity data are obtained from eight state agencies and a municipality. National datasets provide economic and climatic data to round out the model. Residential and commercial demand effects are identified separately, although commercial demand dynamics are not shown to be conclusively different from residential demand dynamics. When a long-run structural demand relationship is taken into account, short-run demand response is not shown to be statistically significant.

The quasidifference price representation used in Chapters II and III is the difference between two weighted rate indices. The weighting function is essentially one of convenience. An empirical test of weighting functions on water rate schedules is developed in Chapter IV. Desirable properties of weighted indices are discussed in Chapter II and detailed in Chapter IV. A method is demonstrated for deriving a consistent empirical weighting, if one exists, by estimating a nonparametric cumulative distribution function from a dataset of two-rate schedules. The empirical weighting function is neither normally nor lognormally distributed, but it is statistically similar to a uniform distribution. Like the empirical weighting, uniform weighting is independent of quantity consumed, but it is easier to apply in practice and performs approximately as well as the empirical weighting in both parametric and nonparametric demand models.

The dissertation concludes with Chapter V, a rephrasing of the central problem of price with a synopsis of procedures and findings and some observations relevant to the application of the findings.

CHAPTER II*

AN ANNUAL QUASIDIFFERENCE APPROACH TO WATER PRICE ELASTICITY

Economic views on water demand continue to gain attention as a result of the scarcity sensitivity that is intrinsic to a value-dependent vision of demand. The almost worldwide phenomenon of rising water scarcity makes the economic perspective useful in multiple ways. Among these is the policy significance of signaling scarcity to all water users through more informed rate-making, so as to motivate efficient consumption behavior and conservation activities. Another key advantage of understanding how water usage depends on water value is being able to perform ex ante appraisals of water projects' prospective benefits. Other policy-relevant advantages are also attributable to the economic view of demand or do not become tractable until demands have been estimated.

To firm up these achievements and turn concepts into practice, economists have conducted many empirical investigations of water demand (Renzetti 2002). The study area of greatest concentration pertains to household demand for water in urbanized areas, which is also the subject of this study (Arbués et al. 2003; Dalhuisen et al. 2003). Undoubtedly, a strong contributing factor for this disciplinary emphasis is the comparative availability of reasonably reliable data. As compared to agricultural, industrial, and heavy commercial water usage, residential/urban water use is more likely

---

* This chapter reprinted with permission from, "An Annual Quasidifference Approach to Water Price Elasticity" by Bell, D. R., and R. C. Griffin, 2008, *Water Resources Research*. Copyright 2008 by American Geophysical Union.

to occur in settings where many water users are active, water use is reasonably well metered and not self-reported, and a variety of consumption circumstances can be observed. The latter factor is important for producing an acceptable degree of variation in statistically exogenous variables, so as to permit analysis of potentially influential factors. In all such studies, fundamental requirements are that consumers have the freedom to determine their water use, and that researchers can observe water use and water price(s), as well as other demand-driving factors.

Utility-maximizing consumer behavior is straightforward to model when price and quantity demanded are well known to the consumer, and standard modeling practice is that consumers are presumed to be rationally advancing their own welfares in the data they generate for us. These assumptions are seldom met, however, when the good in question is retail water service. Unlike most goods households buy, water costs are fully revealed to the consumer well after the consumption decision is made, when the monthly or bimonthly bill arrives. When this bill does come, it typically does not transparently communicate water price to consumers.

Even water quantity information is elusive from the consumer's vantage, since water-consuming taps and appliances hardly ever provide volumetric usage information. Nor does a water bill provide the consumer with a fully satisfactory alternative. A water bill does not itemize the array of water use activities conducted by the consumer; instead they are lumped into a single water usage quantity. On top of the quantity-side confusion, discerning the prospective expenditure effects of behavioral modifications can be a challenge for consumers, given that bills are functionally dependent on some or all of the following: a flat fee per period billing, uniform or block rates, seasonal rates, metered-water-dependent sewerage fees, and often fees for the provision of nonwater services such as garbage disposal and energy.

Consumer perception of water's marginal price is especially dim. Evidence suggests that fewer than 10% of customers invest in marginal price knowledge (Carter and Milon 2005). In a recent survey of water utility systems, only 2.9% provided customers with the price schedule on their water bills (Gaudin 2006). Cognizant of the bounds of consumer rationality under costly information, water (and electricity) demand modelers have turned their attention from the price to which consumers allegedly should respond, to ask which price do consumers respond to (Shin 1985). In econometric terms, this requires formal testing of alternate price specifications.

Unfortunately, the gathering of evidence to settle this empirical question has been confounded by the difficulty of producing any price index that conforms to the OLS assumption of a random error term uncorrelated with the independent variable. Competing specifications cannot be fairly compared unless they are measured accurately. Some previous attempts to construct an exogenous price index are reviewed in this article. None has been entirely satisfactory. An alternative index is proposed that incorporates rate information in a hypothetical price difference between rate regimes. The new index is a quasidifference operation: the difference between the observed lagged price and the unobserved contemporaneous price net of demand-side influences. Since it is based on the published (deterministic) supply decisions of the water provider, this price quasidifference does not vary simultaneously with demand and therefore provides a theoretically unbiased estimate of supply price change. From this basis, the relative behavioral influence of alternate theoretical specifications can be compared. It is hoped that this procedure will open the door to a more active generation of behaviorally based price hypotheses. We limit ourselves here to consideration of marginal price and average price specifications only.

Once an unbiased and behaviorally descriptive price index is selected, an equation of annual demand elasticity is calculated. Community-level rate and usage data are obtained for a sample of 385 utility systems in Texas. The breadth of the data may be unprecedented among studies of this kind. The wide range of observed prices in this data may provide a wider applicability for the estimated parameters than previous research. The aggregate character of the data is respected by weighting the quasidifference estimators by the presumed standard lognormal distribution of households across total quantity demanded. A semi-flexible functional form is employed that allows price elasticity to vary linearly with the climatic parameters, resulting in a rejection of the hypothesis of constant elasticity. Unlike the preponderance of demand analyses which are static, this elasticity in differences provides a time-rate of adjustment (one year) rather than an assumed reequilibrium adjustment. This distinction makes the results especially useful for projecting the repercussions of a change in pricing policy over the near future, or in planning successive rate changes.

Sources of and Responses to Price Endogeneity

Charges for residential water service are set administratively, typically only at the beginning of the fiscal year. Consumers experience the rate schedule as they would a market supply correspondence, except that the household supply function is nonconstant when the marginal price of water varies with household usage. In contemporary rate structures the most common form of water price discrimination is the increasing block rate (IBR) structure, which is found in 47% of the present data (with less than 1% exhibiting decreasing block rates). Under IBR or any other rate regime where price is determined simultaneously with the quantity decision, identification issues analogous to those familiar to market demand analysts must be addressed (Working 1927). It is also possible that the choice to adopt IBR is itself endogenous (Hewitt 2000b; Reynaud et al. 2005).

In choosing a consumption quantity, consumers subjected to block rates implicitly select a marginal price, even if they are unaware of the choice. If an entire community is modeled as a single representative consumer, this price endogeneity can be exaggerated, spuriously influencing elasticity estimates (Shin 1985). The low-information average price specification is further biased by the algebraic simultaneity of division by the dependent variable when a flat fee is included (Taylor et al. 2004). This problem exists for uniform rates (constant marginal price) as well as for variable block rates. Given these inconvenient properties of observed price measures, research has tried to derive a price variable that more adequately captures the ceteris paribus effect of changing fee schedules. Previous strategies to properly identify the price signal may be generally grouped into reduced form, instrumental variable (IV), and maximum likelihood (ML) techniques (Herriges and King 1994).

Reduced Form Price

The reduced form strategy involves creating a price index of known fee schedule parameters that is independent of observed volume. An early example is provided by Taylor (1975), who proposed regressing on each block of multi-block rates. Since nonlinear fee schedules are multidimensional, this technique incorporates more price information, while eliminating quantity consumed as an argument of price charged. The disadvantages of the approach are the lack of theoretical support, additional complexity (Herriges and King 1994), and misspecification bias. The latter arises from the inaccurate assumption that any given price index will be equally representative across the range of observed consumption quantities.

Instrumental Price

The IV approach (Nieswiadomy 1991) allows price to vary across the observed range, at the cost of additional complexity, by identifying a linear proxy to the theoretical supply curve. Although widespread in studies of competitive markets, IV applied to public utilities suffers a number of disadvantages. One is the problem of estimating a censored variable as a line. IV price estimates evaluated at the extrema are not necessarily a combination of experienced prices or even necessarily greater than zero. The result is a correlation between the IV price and the regression error (Terza 1986). When this problem is addressed with the use of limited dependent variable techniques, the method is equivalent to ML price estimation.

The demand price, that ideal scalar employed by the model consumer's decision process, is ultimately unknown. Demand modeling depends upon parameterizing the demand price in terms of the supply price, i.e. the rate schedule. An IV price is therefore an instrumental estimate of an instrument. Problematically, the IV price correspondence predicts intra-annual price changes that are neither observed nor institutionally feasible. Nevertheless, IV may be necessary if the data used are spot prices at arbitrary consumption levels. If the timing and magnitude of fee schedule changes are known, however, the IV approach is a distant second-best solution, as it is inefficient to reconstruct perfectly known price policies into a stochastic estimate of pricing policy. In Texas as elsewhere, price schedules are available data, so an instrumental estimation of price is unnecessary. Even though household perception of price remains mysterious, the supplier's signal is known to researchers.

Maximum Likelihood Price

ML estimation can be used to probabilistically assign a marginal price to a representative consumer either based on an IV inverse supply function or as a two-stage procedure simultaneously estimating price and quantity demanded (Burtless and Hausman 1978; Herriges and King 1994). The "discrete/continuous" (Hanemann 1984) or "endogenous sorting" (Reiss and White 2005) model is a ML model brought to the arena of water demand by Hewitt and Hanemann (1995). The story behind endogenous sorting is that consumers select the price region (block) in which their consumption will lie, then an exact quantity within the block (Hewitt and Hanemann 1995). The method adds a degree of rationality to the price specification dilemma, but perhaps too much.

The information demand on the consumer under this model is intense (Martínez-Espiñeira 2003), and some studies have experienced difficulty deriving a probability estimate that is positive in the neighborhood of price kinks (Cavanagh et al. 2002). That is, observed aggregate quantity decisions may be assigned negative probabilities, implying that the representative consumer is irrational. Ensuring the existence of a cumulative distribution function under such circumstances is nontrivial. Furthermore, repeated application of ML in a dynamic model is computationally demanding (Reiss and White 2005).

More fundamentally, when aggregate data are modeled with a ML price, the distinction is lost between the representative consumer and representative consumption. If the individual makes a ML consumption decision, the community consumes across the whole probability distribution. Whether the average consumer enjoys average consumption depends heavily upon the normality assumption (Hewitt 2000a). Although the ML approach is unsupportably utility theoretic under incomplete information, it does offer a helpful framework that will be exploited in a forthcoming section addressing aggregation issues.

Quasidifference Price

Assuming that demand for water service is functionally related to price and other exogenous variables, the typical under-identified demand function is

$$w = \omega(p(w), z), \qquad [2.1]$$

where *w*, quantity demanded in a given period, is functionally related to *p(w)*, the price index calculated at *w*, and other variables, z. Net price changes may fruitfully be seen as composed of a policy price change and possibly an endogenous change resultant from a change in quantity demanded. Although the price faced by any particular consumer varies with the level of consumption, the nominal price schedule varies exogenously only when the water provider decides to vary it. Changes in price due to the purchasing power of money are also exogenous, but are withheld from the following discussion for clarity.

Though a price level net of demand-side influences has proven elusive, price change is separable in the derivative,

$$\frac{dp}{dt} = \frac{\partial p}{\partial t} + \frac{\partial p}{\partial w}\frac{\partial w}{\partial t}, \qquad [2.2]$$

$$\text{where } \frac{\partial p}{\partial t} = \Delta p(\bar{w}) \qquad [2.3]$$

is the exogenous price difference evaluated at some consumption level $\bar{w}$.

The form of equation [2.1] to be estimated is the annual difference in demand:

$$\Delta w = w_t - w_{t-12} = \Delta\omega(\Delta p(\bar{w}), \Delta z). \qquad [2.4]$$

Choice of the point $\bar{w}$ depends on the price change that is to be measured, because price changes are not generally uniform. A simple reduced form approach is to evaluate Δp at a single consumption level for all observations. The reduced form assignment of

$$\bar{w}_t = W^*, \text{ for all } t, \qquad [2.5]$$

is too rigid, though, if consumption is not stationary about W^*. Whenever consumers migrate their consumption out of the rate block containing W^*, $\Delta p(W^*)$ will cease to be a good estimator of Δp. If instead,

$$\bar{w}_t = w_{t-12}\,, \qquad [2.6]$$

Δp may be interpreted as the price change that would have obtained if consumption had remained constant from the same month of the year before. The interpretation conforms to both the behavioral model of households reacting to pricing policy and to the ceteris paribus principle of statistical inference. The quasidifference estimator is defined as

$$\Delta p = p_t(w_{t-12}) - p_{t-12}(w_{t-12}) \qquad [2.7]$$

in the linear model, or

$$\Delta \ln p = \ln\left(\frac{p_t(w_{t-12})}{p_{t-12}(w_{t-12})}\right) \qquad [2.8]$$

in the logarithmic models used in this paper.

The necessity of adopting an annual lag when monthly data are available follows from the dominance of seasonal behavior in water consumption patterns. Seasonality has been modeled with climatic variables (Griffin and Chang 1991) and with Fourier harmonics (Renwick and Green 2000), but neither method has completely captured the persistent demand characteristics unique to each month of the year.

This dynamic form dictates a specific interpretation of estimated parameters. The implied consumption response occurs within a single community over the span of one year. Comparisons across communities are no longer applicable, including the common interpretation of cross-sectional variation as a measure of long run adjustment (Kennedy 2003, p. 211). Because the differential form implies a price elasticity of demand for water to pricing policy (and inflationary) changes within a given community, its implications are more relevant to projecting and evaluating incremental local adjustments than basinwide projects with long horizons, which would benefit from the scope of a static model. The results of this estimation should not be used to prescribe an efficient equilibrium pricing policy because adjustment will commonly take longer than the one-year time step emphasized here.

On the other hand, standard structural estimation is not well suited to applications requiring a finite time horizon. The price response implied by such models may take an indefinitely long time to realize. Knowledge of the time-path of adjustment is necessary to describe optimal policies that achieve period-by-period utility system goals such as revenue sufficiency and stability. For example, in cases of acute capacity constraint such as drought, timing is a factor, and a policy based on a structural elasticity may not achieve the desired demand management goal (i.e. a temporary reallocation) before the drought dissipates. An elasticity derived from the approach introduced here is recommended for such applications.

Aggregation

The ML endogenous sorting model recognizes that different consumers make choices that place them in different rate blocks (Hanemann 1984), but the implications for aggregation have not been well explored. The probability that a consumer consumes within a rate block is analogous to the proportion of consumers in an aggregate who consume within that block. This interpretation allows the usual point estimate of aggregate consumption to be replaced with a consumption distribution in the formulation of price indices. When block rates are present in the data, this alternative can greatly improve the explanatory power of price.

Because all consumers do not simultaneously move from block to block, a point estimate of representative consumption and price exaggerates block effects (Shin 1985). Schefter and David (1985) observe that the price faced by the mean consumer may estimate mean price with bias, especially if the variance of consumption is high. The distributional symmetry assumption that justifies point estimates of marginal price is tenuous and has been empirically rejected (Hewitt 2000a; Schefter and David 1985). Distribution of water consumption over households is asymmetrical (with median < mean) and truncated at zero, conforming to a possible gamma or lognormal distribution, as a small number of households consumes a relatively large amount of water. Martínez-Espiñeira (2003) corrects this bias using additional information about customer types to weight marginal price across the community aggregate.

The least precise representation of aggregate marginal price under block rates is a point-mass centered at the mean of consumption multiplied by the price effective at that consumption level. The most precise is a weighting of prices by the actual proportion of consumers whose marginal consumption falls in each block. Lacking agent-level data, the model employed here uses a distributional assumption in lieu of customer type data. Since the standard lognormal distribution is asymmetrical, truncated at zero, and uniquely determined by a single parameter that is conveniently related to mean consumption, $\bar{w}$, the distribution of individual consumption levels for each community in each period is modeled as standard lognormal. A lognormal distribution of w is consistent with OLS assumptions on

$$\ln w = \beta \ln X + \varepsilon, \qquad [2.9]$$

which is the general form on which the present analysis is based. The aggregate quasidifference price variable is therefore a quasidifference operation on a linear combination of prices weighted by a block consumption probability function assumed to be standard lognormal with a mean at the data point $\bar{w}_t$. The assumption of lognormality is much cruder than the sorting devices proposed in the ML models, suggesting that even more precision could be achieved by refinements of the weighting function.

Let $F(w)$ be the cumulative distribution of a standard lognormal function whose mean is $\bar{w}_t$. Given block rate function $P(w) = \{p_i : x_{i-1} < w < x_i\}$, where w is partitioned into N blocks by $x\left(x_0 = 0, x_N = \infty\right)$, the aggregate price index is defined as

$$p = \sum_{j=1}^{N} p_j [F(x_j) - F(x_{j-1})] \qquad [2.10]$$

The procedure is analogous to probability weighting of time-of-day electricity prices (Hausman et al. 1979). Choice of the partition x is straightforward for the calculation of marginal price. All consumers in the same rate block do not share a common average price, however. In the calculation of aggregate average price, x is defined in increments of 500 gallons up to 50,000 gallons, and a mean average price is calculated for each interval. To construct the aggregate quasidifference price, both the contemporaneous and the lagged price functions are weighted by the same distribution and then differenced.

Empirical Model

Dynamic Specification

Most empirically estimated water demand equations have been static. Unfortunately, dynamic studies that have tested the significance of the contemporaneous price variable have found its effect to be insignificant during nonsummer months (Lyman 1992), inconsistent across models (Agthe and Billings 1980), insignificant or unexpectedly close to zero (Carver and Boland 1980). Investigations of natural gas (Balestra and Nerlove 1966) and electricity (Bushnell and Mansur 2005) have also suffered from weak results. Low significance levels in contemporaneous price are consistent with the hypothesis of incomplete information, which would imply a learning process over time (Carver and Boland 1980).

Nauges and Thomas (2003) provide a more revealing dynamic analysis that estimates a statistically significant short-term elasticity of –0.26. Although their study is focused on cross-sectional heterogeneity, an issue set aside in the current research, it is exemplary in the sense of incorporating additional information on pricing practice unique to the region under study.

The price elasticity actually measured by static equations is typically cross-sectional elasticity (Balestra and Nerlove 1966). It can be argued that this is a measure of the longest adjustment term, over which habits and stocks of water-demanding capital have tended to evolve to near-equilibria. Such a horizon is too long to serve all types of policy analysis, however, as suggested above. In contrast, an annual elasticity is sought here. Although a wider range of results are obtained, the central question is, "What is the percent change in consumption over one year following a uniform 1% rate change or its equivalent?"

Flexibility

Quantity-dependent pricing implies highly informed consumers would perceive a nonlinear budget set, then inferring idiosyncratic hypotheses about the resulting price elasticity. In particular, price elasticity is expected to vary with income level, especially for a subsistence good (Dalhuisen et al. 2003). The limited available evidence suggests that price elasticity is also sensitive to climatic conditions in nontrivial ways (Griffin and Chang 1991). Nevertheless, many empirical estimates use the simple log-linear functional form (Equation [2.9]), which imposes constant elasticity. In this instance, a generalization of Equation [2.9] is employed that allows a second-order interaction among the covariates (X_I), based on the translog functional form (Christensen et al. 1973):

$$\ln w = \sum_{i=1}^{I} \beta_i \ln x_i + \sum_{i=1}^{I} \sum_{j \neq i}^{I} \frac{\gamma_{ij}}{2} \ln x_i \ln x_j + \varepsilon \qquad [2.11]$$

Note that the quadratic terms of the full translog model are excluded. While the model indicated by Equation [2.11] is more flexible than that that of Equation [2.9], it is not a globally or even locally flexible form.

Price Elasticity

To distinguish price effects, Equation [2.11] may be rewritten as

$$\ln w = \beta_p \ln p + \sum_{i=2}^{I} \beta_i \ln x_i + \sum_{i=2}^{I} \gamma_{pi} \ln p \ln x_i + \sum_{i=2}^{I}\sum_{j=2}^{I} \frac{\gamma_{ij}}{2} \ln x_i \ln x_j + \varepsilon \qquad [2.12]$$

Introducing a temporal element and taking the first difference produces

$$\Delta \ln w = \beta_p \Delta \ln p + \sum_{i=2}^{I} \beta_i \Delta \ln x_i + \sum_{i=2}^{I} \gamma_{pi} \Delta(\ln p \ln x_i) + \sum_{i=2}^{I}\sum_{j=2}^{I} \frac{\gamma_{ij}}{2} \Delta(\ln x_i \ln x_j) + \nu \qquad [2.13]$$

Equation [2.13] is the estimating equation for the empirical analysis of the next section. Since an expression for the ceteris paribus price effect on the quantity of water demanded is sought, $\Delta \ln x_i = 0$ is stipulated for each *i*-th covariate when calculating price elasticity. Note that, for each term in the second summation,

$$\gamma_{pi}\Delta(\ln p \ln x_i) = \gamma_{pi}[\ln p_t \ln x_{it} - \ln p_{t-1}(\ln x_{it} - \Delta \ln x_i)] = \gamma_{pi}(\Delta \ln p_t \ln x_{it} + \ln p_{t-1} \Delta \ln x_i) \qquad [2.14]$$

Thus [2.13] reduces to,

$$\frac{\Delta \ln w}{\Delta \ln p} = \beta_p + \sum_{i=2}^{I} \gamma_{pi} \ln x_i + \nu \qquad [2.15]$$

In the limit, for small changes,

$$\eta = \beta_p + \sum_{i=2}^{I} \gamma_{pi} \ln x_i \qquad [2.16]$$

is the own-price elasticity of demand. Constant elasticity is confirmed if $\gamma_{pi} = 0$ for every $i \in I$.

Data

The original data for this application consists of monthly water supply series for 385 Texas communities (serving 5.6 million Texas residents). Water use data are provided by the Texas Water Development Board, and corresponding water and sewer service rates are provided by the communities themselves per request. Of 1406 community water providers considered, 734 responded to mailed inquiries seeking water and sewerage rate structures for a five-year period. Raw data expressed in nominal dollars are corrected for inflation in the analysis.

Due to the lag structure of the proposed model, only those communities for which supply and price series are complete from January 1999 to December 2003 are considered further. The 385x60 panel contains 23,100 elements of which 20% are expended in support of the lag structure. Twelve additional observations are excluded because the marginal price changed from zero to a positive quantity over the year, resulting in an undefined log difference. In other cases (<1% of data) where price remained at zero in both periods, log price difference is redefined to be zero. Based on comparisons of community size and monthly usage, the sample is representative of the targeted population although the high variance of both of these measures reduces their ability to verify sample selection bias.

Personal income statistics from the Bureau of Economic Analysis (www.bea.gov /regional/reis/) and climate data from the National Climatic Data Center (NCDC; www.ncdc.noaa.gov/) augment the data. Personal income is aggregated at the county level, with 156 counties represented, or at the metropolitan level of larger cities for which income data are available. Daily temperature and precipitation data are matched by proximity to the nearest NCDC cooperative weather station, usually in the same county as the system observation. All dollar amounts are normalized to December 2003, using the Urban South CPI measurement (www.bls.gov/data/). Data are summarized in Table 2.1. The log difference transformations of the data are summarized in Table 2.2.

Table 2.1. Summary Statistics, N = 23100

Variable	Units	Mean	Standard Deviation
Volume per capita per day	Liters	540.3645	267.2517
Marginal water price	2003$ / kliter	0.6607	0.2966
Marginal sewer price	2003$ / kliter	0.1466	0.2314
Average water price	2003$ / kliter	1.1242	0.5506
Average sewer price	2003$ / kliter	0.4356	0.4215
Monthly personal Income	2003$	2158.2780	499.6223
Average minimum temperature	ºF	55.3835	14.0675
Average maximum temperature	ºF	78.2141	13.2448
Days in month with no precipitation	Days	27.1997	2.7631

Table 2.2. Summary of Differences in Logs, N = 18468

Variable	Mean	Standard Deviation
*d*ln*w*	-0.0141	0.2658
*d*ln*MP* (water)	0.0066	0.0980
*d*ln*MP* (sewer)	0.0012	0.0691
*d*ln*AP* (water)	0.0041	0.0776
*d*ln*AP* (sewer)	0.0084	0.0936
*d*ln*PI*	0.0010	0.0276
*d*ln*Tmin*	-0.0057	0.0916
*d*ln*Tmax*	-0.0086	0.0775
*d*ln*dry*	-0.0020	0.1862

Figures 2.1 and 2.2 illustrate the variation in water price over the sample. The wide variation in price measures provides an advantage in estimation over more geographically limited studies, in that the estimated expression for price elasticity may be confidently generalized over a wider range of price levels. The degree of variation in the other regressors supports the maintained hypothesis that individual (cross-sectional) effects are random across the sample.

Results

Log-linear Model

The centerpiece of this econometric investigation is estimation of a multidimensional elasticity function. However, simple log-linear regressions are performed ahead of the more flexible central regression to guide the selection of independent variables. Variables included in the preliminary regressions are average

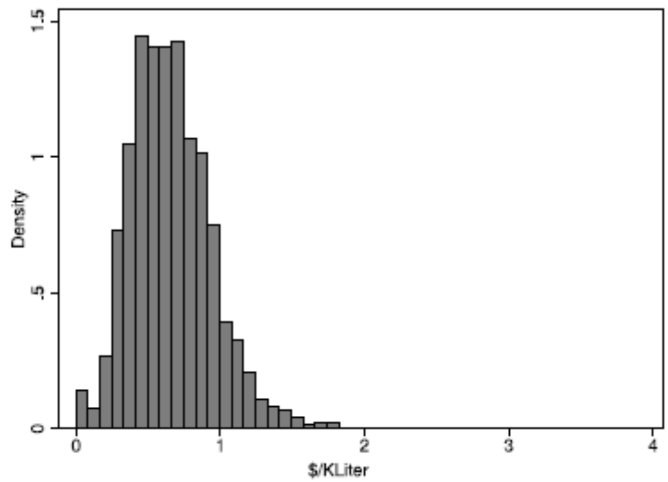


Figure 2.1. Distribution of Observed Marginal Water Prices

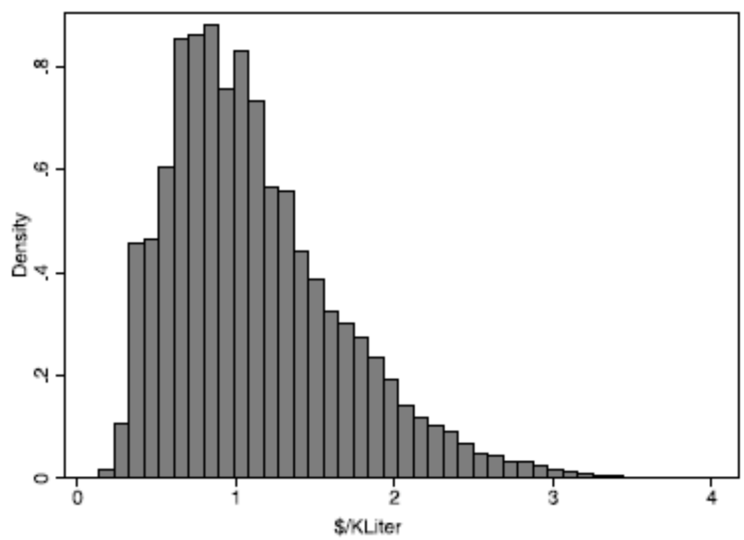


Figure 2.2. Distribution of Observed Average Water Prices

water price, marginal water price, average sewer price, marginal sewer price, monthly income, mean minimum daily temperature, mean maximum daily temperature, and number of days in the month with less than 0.25 inches of precipitation. The regressed values in each case are the annual differences in logarithms of variable levels.

It has been argued that the nonlinear price schedule creates a secondary income effect that ought to be measured (Nordin 1976). The “Nordin difference” (which is not in any way related to the quasidifference introduced here) has not been included, primarily because the assumptive base for identifying such a variable in aggregate data is too tenuous. This choice is also justified ex post by the insignificance of the primary income variable, as will be shown below.

Final selection of price variable is determined by comparing a marginal price model with an average price model using the Akaike Information Criterion or the Schwarz Criterion, both of which are in this case equivalent to finding the specification with the lowest sum of squared errors. Additionally, since parsimony is improved if water and sewer prices can be combined or if minimum and maximum temperatures can be averaged prior to estimation, both of these hypotheses are tested.

Results of the log-linear estimations are shown in Table 2.3, with marginal price variables included in Model 1 and average price variables in Model 2. The lower information criteria corresponding to Model 1 indicate the better fit. On this comparative basis, marginal price is adopted as the price index of the central analysis. The marginal sewer price index, however, is insignificant. For many systems in the sample, marginal price for sewer service is zero in nonwinter months due to many utilities' policies of setting sewer cost ceilings and "winter averages". The insignificant coefficient may indicate that consumers are not aware of these practices. Even considering months where marginal sewer price is strictly positive, the t-statistic for the corresponding coefficient is only –0.35. Consumption is apparently unresponsive to

Table 2.3. Log-linear Regression, N = 18468

Variable	Dependent variable is *d*lnW (t-scores)	
	Model 1	Model 2
*d*ln*MP* (water)	-0.1423 (-7.20)[a]	
*d*ln*MP* (sewer)	0.0077 (0.28)	
*d*ln*AP* (water)		-0.1124 (-4.47)[a]
*d*ln*AP* (sewer)		-0.0462 (-2.21)
*d*ln*PI*	-0.0662 (-0.96)	-0.0738 (-1.07)
*d*ln*Tmin*	-0.3238 (-12.14)[a]	-0.3220 (-12.06)[a]
*d*ln*Tmax*	0.9062 (26.78)[a]	0.9051 (26.73)[a]
*d*ln*dry*	0.0947 (8.49)[a]	0.0952 (8.53)[a]
Constant	-0.0069 (-3.63)[a]	-0.0070 (-3.66)[a]
Akaike criterion	2267.3	2289.3
Schwarz criterion	2322.1	2344.0
F(6, 18461)	208.84	204.94

[a] $p < 0.01$.

marginal sewer pricing. This variable is not included in the central regression.

An alternative, nested test of water price specification is proposed by Shin (1985). Here both average price and marginal price variables are included in the same regression. Parameter estimates for this regression are not reported. Interpretation of the marginal price and average price coefficients as $(1–k)\beta$ and $k\beta$, respectively, allows a measure of the relative influence of marginal and average price on the consumption decision (Shin 1985). For water, we obtain k = –0.76. On the basis of k < 0.5 for water service, marginal price is indicated as more influential than average price. The corresponding value of k = 2.89 for sewer service supports the average price specification for sewerage. Bearing in mind that 0 < k < 1 in a well-specified Shin test, these results are curious.

The hypothesis that the effect of an increase in daily low temperature is equivalent to the effect of an increase in daily high temperature is rejected. Both variables are included in the more flexible regression.

The existence of an income effect is rejected. An unfortunate characteristic of the income data is that the variation is cross-sectional except for the CPI normalization, and is therefore unnoticed by this procedure. It is plausible that income is insignificant because of slow aggregate response to income change, but it is more likely in this case that the income measure is simply too broadly aggregated to accurately identify the spending power of a single community. Perhaps a refined monthly income measure would produce better results. Income is not included in the following regression.

Semi-Flexible Model

The final regressors are differences in logs of marginal water price, average low temperature, average high temperature, and number of days without precipitation, as well as the differences in products of each pair of independent variables' logarithms. The results of this central regression are summarized in Table 2.4. Due to the inclusion of the interactive product regressors, the intuitive value of Table 2.4 is limited, although the strong significance of these interactive terms justifies the use of the more flexible functional form. In particular, the significance of the price interactions allows the rejection of the hypothesis of constant price elasticity across the sample. The Breusch-Pagan statistic of 1.51 for this regression (p = 0.219) fails to reject the null hypothesis of homoskedasticity.

Table 2.4. Log-nonlinear Regression, N = 18468

	Dependent variable is *d*lnW (t-scores)
Variable	Coefficient
*d*ln*MP*	1.2901 (6.15)
*d*ln*Tmin*	-1.8966 (-3.25)
*d*ln*Tmax*	-15.8700 (-19.30)
*d*ln*dry*	-7.7871 (-17.96)
*d*ln*MP*•*d*ln*Tmin*	0.1904 (3.57)
*d*ln*MP*•*d*ln*Tmax*	-0.4392 (-6.13)
*d*ln*MP*•*d*ln*dry*	-0.0809 (-3.05)
*d*ln*Tmin*•*d*ln*Tmax*	1.6405 (23.47)
*d*ln*Tmin*•*d*ln*dry*	-1.6420 (-10.42)
*d*ln*Tmax*•*d*ln*dry*	3.3590 (14.48)
Constant	-0.0053 (-2.85)
Adjusted R^2	0.1172
F(10, 18457)	246.21

p < 0.01 for all estimated coefficients.

Applying the coefficients in Table 2.4 to Equation [2.16] results in the elasticity equation,

$$\eta\square = 1.290 + 0.190 \log t_{min} - 0.439 \log t_{max} - 0.081 \log d \qquad [2.17]$$

Each of the individual coefficients is significant at the 99% level. Demand for water service is more elastic when daily high temperature is higher or when more days of the month pass without precipitation. Demand is less elastic when daily low temperatures are higher. The magnitude of the coefficient on high temperature is higher than that on low temperature, implying that hotter months see an increase in price elasticity (more elastic demand).

Price elasticity evaluated as a linear combination of mean variable levels and regression coefficients is found to have a mean of –0.127. The standard deviation of η, estimated using the regression standard errors and variance-covariance matrix is 0.0188, implying that demand is inelastic but significantly downsloping at the mean. The estimate is consistent with other recent research on short term elasticity (Martínez-Espiñeira 2004; Renwick and Green 2000). It is somewhat lower in absolute value than most cross-sectional static models (Dalhuisen et al. 2003), consistent with the hypothesis of adjustment lags greater than one year.

Summary

Fully informed and rational consumers will use water until the monetized marginal benefit of the next unit is equal to its marginal price. Yet, price and quantity information is dimly available to water customers, and these consumers cannot improve their information conditions without experiencing costs. Imperfectly informed consumption behavior is therefore the norm. Less informed consumers may be expected to optimize with respect to a lower information price index, for example average price. Although Shin (1985) provides a test of relative explanatory power between two proposed price indices, the test results are only meaningful if both indices consistently represent the theoretical quantities they purport to represent. Prices based on observation or on the usage of a representative consumer are endogenous and not necessarily unbiased.

Prices constructed as instrumental variables can be a poor fit because strong instruments are generally lacking. Continuous, linear pricing is characteristic of IV prices but not of the actual price-setting process. Maximum likelihood prices are not guaranteed to be fully defined throughout the range and do not aggregate well to the community level. If a complete rate history is known, an alternative strategy is to calculate the difference in a defined price index for each consumption level before and after a rate change. In the case of aggregate data, these hypothetical differences should be weighted by the probability density of each consumption level. We assume that consumption is distributed standard lognormally and weight the prices corresponding to each block by the probability density of consumption in the block. A tradeoff of operating in differences is that cross-sectional variation in variable levels disappears, limiting application of the results to annual adjustments. This idiosyncrasy can be put to good use, however, given the thinner water demand literature on adjustment over time.

A comparison of information criteria for log-linear regressions on quasidifferenced marginal and average prices indicates that marginal price change is more influential than average price change. Sewer price changes are not shown to be significant, nor are income changes. An equation of marginal price elasticity of demand is derived from a more flexible regression of annual change in monthly water use on changes in marginal price, mean low temperature, mean high temperature, and number of days without significant precipitation.

The data are an original set of system-level price, quantity, income, and climate observations for 385 systems in the state of Texas, USA. The dataset is remarkable due to its volume and the variety of systems polled, water providers for millions of Texans. Own-price elasticity is shown to vary with climatic conditions. The derived mean elasticity of –0.127 in the first year is plausible in relation to previous research. It is less elastic than most structural estimates of long-run elasticity, implying an adjustment period longer than one year.

As water demand adjustment behavior remains incompletely understood, further research that demonstrates both shorter and longer demand patterns in an integrated way would contribute significantly to modeling and policy-setting efforts. A fundamentally elusive element is the decision mechanism of the retail water consumer. Since neither marginal price nor average price appears to capture this mechanism fully, developing and testing of new price indices is to be anticipated. In further research on aggregate demand under block pricing, more consistent and representative price indices could be developed by incorporating probabilistic methods from endogenous sorting models previously applied only to micro-level data.

CHAPTER III*

URBAN WATER DEMAND WITH PERIODIC ERROR CORRECTION

Local media have applied the phrase "water crisis" so often in describing the condition of some city that it has become cliché (Russ 2009; Cregan 2009; Evans 2009; Bond 2009). Despite the sensationalizing rhetoric, excess demand for publicly supplied urban water persists in many places and is arising in others. The resulting management issues underscore the troubled and oft-politicized nature of water planning. Urban water supply is naturally monopolistic due to its high capital requirements. Therefore, an assumption of the invisible hand theorem is unmet, and socially efficient allocation is not automatic. Most decision making is conducted by public water authorities which do not have a strong track record of efficient adaptation (Grafton and Ward 2008; Lach, Ingram, and Rayner 2005; Hewitt 2000). Yet, experimentation, along with progress, is slowly occurring. Among the policy mechanisms being tried are alternative rate structures and higher rates. Perhaps rates that include water's opportunity costs will eventually be explored, as recommended by economists. If these approaches are to be successful, planners and regulators require consumer demand information to simultaneously establish rates and anticipate the level of water deliveries.

Traditionally, water utility systems have focused narrowly on adjusting water supply to meet level-price water demand (Dziegielewski 1999). Still, efficient supply enhancement requires knowledge of future aggregate demand, and carelessness over

---

* This chapter reprinted with permission from, "Urban Water Demand with Periodic Error Correction" by Bell, D. R., and R. C. Griffin, 2011, *Land Economics*. Copyright 2011 by University of Wisconsin Press.

either revenue-seeking or efficiency-seeking changes to water and wastewater prices may lead to error in demand projections. Even under perfect information, the costs of supply enhancement continue to rise as the most accessible sources of water are tapped to capacity or depleted, necessitating rate changes that subsequently affect quantity demanded. These increasing costs further advance the value of traditionally underemployed demand management strategies, including efficient pricing.

Pricing, or rate-setting, is complicated by the balancing of multiple objectives. Unlike the textbook monopolist, the typical water utility system does not pursue the objective of profit maximization. In addition to economic efficiency, water utilities seek goals such as revenue sufficiency and fairness (Griffin 2006, p. 251), and they rank these more highly than efficiency. While public authorities commonly infer that their rate-setting efforts pursue multiple goals simultaneously, the Tinbergen principle warns that each goal requires a separate instrument (Young and McColl 2005). Moreover, achieving multiple goals involves the solution of a complex set of system objectives, which requires detailed knowledge of demand behavior. For example, when efficiency and revenue sufficiency are conjunctively sought, an efficient marginal rate may be derived from supply information alone, but the calculation of other rate components requires an estimate of future billed volumes (Griffin 2001; Edwards 2006). For demand management policies to improve, analytical techniques for estimating demand must evolve to support them.

Econometric estimates of residential demand for water abound (Dalhuisen et al. 2003), but existing demand estimates lack the detail to support many rate design applications. A time-path of adjustment to consumer equilibrium is seldom explicitly estimated in prior econometric work; seasonal patterns of demand tend to be underrepresented; and the important commercial and industrial sectors of demand are often set aside or assumed to be proportionally linked to residential use. Water management policies that fail to consider the time-path of adjustment risk outpacing consumers' ability to develop new habits or optimize their stocks of water-associated capital, such as landscaping, plumbing fixtures, and appliances. Policies without seasonal considerations risk incurring excess demand during cyclical demand peaks. Policies that fail to differentiate between household demand and commercial demand risk decreased efficiency relative to sector-tailored policies. The empirical results of this research indicate that slow adjustment, seasonality, and sectoral sensitivity are all characteristic of the sample. Simpler models that are unable to accommodate these characteristics are therefore misspecified to some extent.

The present research incorporates a demand function into a dynamic consumption model using the error correction (EC) technique, thereby merging both short- and long-run demand drivers and possibly improving forecast accuracy (Engle, Granger, and Hallman 1989). Seasonal demand behavior is modeled by identifying and accounting for periodic integration in the associated series (Boswijk and Franses 1995a). Commercial and industrial contributions to aggregate demand are modeled by including sectoral intensity factors in the estimating equation. Unlike dynamic water demand studies considering only a single locale, the present research includes original data from a panel of 167 geographically dispersed cities within the United States, observed monthly from January 1995 through December 2005. The breadth of the sample allows a model to be fitted over a wider range of conditions than previously possible. The multitude of community cross-sections allow a uniquely statistical look at the problem of nonstationarity in quantity demanded.

A periodic error correction model is developed in the next section. The model is applied to the national panel in the section following. The last section offers concluding remarks.

Developing the Theoretical Model

Structural and Dynamic Demand

The object of most econometric water demand research is a demand function for water, which is a mapping of consumption quantities over the range of possible prices and other variables. For convenience, the relationship between price and quantity demanded is often characterized by a price elasticity scalar. Such demand functions rely on mathematical structure implied by microeconomic theory, so they are called structural models as opposed to statistical models which are theoretically unrestricted. The early structural models were well suited to examining the belief held by noneconomists that rates do not affect use (Howe and Linaweaver 1967). The persistent testing and rejection of this hypothesis by economists is a contribution of our literature (Espey, Espey, and Shaw 1997), albeit only haltingly applied to policy. The same models are pivotal to the determination of consumer valuation that is necessary for thorough policy and project appraisals, because it is the demand function that identifies the marginal benefit of price and quantity changes.

It is unlikely, though, that water consumers instantaneously adapt to demand perturbations such as price and weather shocks (Griffin and Mjelde 2000, Carver and Boland 1980). During periods of adjustment to new conditions, consumers are constrained by their learned behaviors (habits) and by their inventories of water-associated durable goods, e.g. appliances and landscaping. While adjusting, customers can experience nonzero excess demand in the sense that the quantities they demand would not be optimal in a longer view during which habits and durable possessions can be refined. A demand correspondence is not a demand function if multiple consumption quantities can be mapped to the same argument values, so the structure supporting the typical static demand model can lead to conflicting results in a dynamic context. Conversely, the dynamic demand correspondence offers little insight into consumer welfare or willingness to pay.

The inclusion or omission of an adjustment process separates the time-independent water demand models from the dynamic models. The former class contributes insights primarily at longer time horizons, whereas the latter may more accurately model short-run behavior. A structural model can utilize slow-moving variables that either do not vary or are not measured monthly, just as a dynamic model can incorporate seasonal changes that are insignificant or average out in the long run (Engle, Granger, and Hallman 1989). In contrast to the literally hundreds of structural water demand studies stand only a handful of dynamic demand studies (Bell and Griffin 2008a; Fullerton and

Elias 2004; Nauges and Thomas 2003), although some essentially structural studies have employed the flow-adjustment hypothesis to deduce an adjustment rate (Lyman 1992; Carver and Boland 1980).

Regardless of the comparative advantages of the two approaches, the possibility that forecasts from a structural model may contradict those from a dynamic model creates a tension between them (Engle, Granger, and Hallman 1989). A dynamic model may be the preferred tool for balancing the objectives of controlling water use and covering production costs year to year, but only a structural model can be interpreted as a demand function. Fortunately, advancements in statistical treatment of time series now allow the simultaneous enjoyment of both sets of advantages. The integrated model proposed below will be used in the next section to estimate the demand for water in United States cities and to predict twelve monthly consumption quantities beyond the estimation sample. The model will also facilitate testing for monthly seasonality and instantaneous adjustment.

Error Correction and Periodic Cointegration

The empirical model developed here is an extension of a model in first differences previously used to project annual changes in quantity demanded (Bell and Griffin 2008a). A shortcoming of the earlier application is its omission of a force summoning consumer equilibrium. Even though excess demand is not expected to be identically zero in any particular time period, its tendency toward zero is as omnipresent as individual self-interest, in the sense that individuals are not content with states of nonzero excess demand. Inclusion of a lagged expression of excess demand turns the difference model into an EC model (Engel, Granger, and Hallman 1989).

Excess demand reflects an imbalance that can be improved upon, an error to be corrected. A condition of excess demand implies that a higher level of aggregate utility could have been achieved at the same expenditure level with a different stock of capital or information; thus excess demand is not stable. Although the mechanisms and information requirements for rational consumers to resolve their ideal consumption bundles are not explicitly identifiable, the assumption that a locally stable structural demand exists implies that all solution paths starting from small levels of excess demand converge to points on the structural demand curve (McKenzie 2002, p. 56). Because utility systems set rates in advance of realized demand, convergence implies adjustments in quantity demanded. In a linear EC model, the speed of convergence is represented by the coefficient of the EC variable. If the EC coefficient is positive, the system is explosive. If the coefficient is equal to -1, full correction takes place in one time period. If the coefficient is in the interval (0, -1), correction takes longer than a single period.

It is necessary for the consistency of the model that the EC term is stationary – that its conditional means are distributed about its sample mean – since the dependent variable is presumed to be stationary. The lagged EC term, which is the lagged residual of a structural demand function, may not in fact be stationary if the dependent variable of the structural model is not stationary. If this is the case, the residual will be consistent only if the structural model is cointegrated (Juselius 2006, p. 86). If the left- and right-hand sides of a regression model are cointegrated, their respective lacks of stationarity have cancelled each other, so they are "super-consistent" with respect to each other, and their residual will qualify as a valid EC regressor. Stationarity is relative, however – "a convenient statistical approximation" (Juselius 2006, p. 20) – so it is important not to assign too much weight to the various tests of cointegration, the unit root tests. Super-consistency is not necessary for consistency of EC parameter estimates, and mere consistency can be tested ex post.

Water consumption patterns may exhibit seasonality, an attribute of many macroeconomic series for which new modeling techniques have been proposed within the cointegration literature. The technique of seasonal cointegration dictates the inclusion of multiple EC terms corresponding to multiple seasonal lags (Kunst 1993). Seasonal cointegration is mathematically appealing when the frequency of the data is quarterly but much less so when the frequency is monthly. As the mathematical and computational requirements increase, economic interpretation of the results becomes more elusive. An alternative is periodic cointegration (Boswijk and Franses 1995a). Periodically cointegrated series are bound by a vector of coefficients that vary from season to season. In the case of monthly data, periodic cointegration is represented by a vector of 12 separate variable combinations, one for each calendar month. It has been established that a pair of series cannot compose a valid cointegrating relation in one season without being cointegrated in every season (Castro and Osborn 2008). In addition to an annual cycle, water demand may exhibit cycles longer than one year, but we will leave the possibility to further research. Some communities may experience seasonality that is only approximately annual in frequency, and this characteristic would be better modeled by seasonal integration than by periodic integration.

Use of the EC model has a single, recent precedent in the water demand literature. Martinez-Espineira (2007) illustrates well the difficulties of seasonal cointegration. Nine years of monthly data on a single community are further collapsed into quarterly observations to circumvent the daunting procedure of simultaneously testing stationarity in 12 monthly frequencies. Results of the battery of diagnostic tests are generally consistent but not definitive due to both the small sample size and the low power of existing unit root tests. Martinez-Espineira's paper is nevertheless a milestone in terms of introducing EC to water demand modeling that is extended here in data scope, variable sophistication, and use of the periodic EC alternative.

The object of the present research is community demand for publicly supplied water. The water demanded in the sample is delivered by a sole or majority supplier. Community demand is not synonymous with residential demand inasmuch as businesses as well as residents demand water within the community. Even though it is recognized that water consumption is not entirely residential, an expedient practice when using aggregate data is to simplify analysis by representing the dependent variable as the ratio of quantity demanded to population (for a recent example, Ruijs, Zimmermann, and van den Berg 2008). Such a practice raises questions about the role of commercial and industrial activity in aggregate demand. The assumption that the extent of commercial and industrial water demands are proportional to population is stronger and less appealing than the assumption that the extent of residential demand is proportional to population, especially with respect to a diverse cross-section of cities.

Nevertheless, a demand-per-capita dependent variable facilitates the planning convention of multiplying per capita demand by projected population. More generally, it separates intensive from extensive community demand growth, allowing a population-free intensive demand comparable to the results of a household-level study. Potential and practiced applications of community water demand are many and varied, so a flexible representation has the advantage of interoperability. In this spirit, the community demand model estimated in this research includes sectoral intensity factors as independent variables. By considering the extent of commercial and industrial activity (in dollars) per capita, the model incorporates more sectoral information without sacrificing the advantages of an intensive dependent variable.

Price Specification

Because of the complexity of typical water rates and the absence of a true market for processed water, the rates of exchange for water cannot be called market prices. For a given utility system and client there is, however, a cost of the marginal unit of water, dw, that influences residual income, m. If that marginal cost is designated p then it is equal to the ratio of the change in income, dm, to dw:

$$p = \frac{dm}{dw} . \qquad [3.1]$$

If p were known, it would enter an individual demand function, ω, with income and other prices, **P**:

$$w = \omega(p, \mathbf{P}, m) \qquad [3.2]$$

under the usual assumption of utility-maximizing water consumers operating in an environment of costless transactions. This is the perfect information rationale for using marginal price as an argument of an empirical demand function.

On the other hand, marginal price is generally unknown to the consumer (Foster and Beattie 1981a). By one estimate, fewer than 10% of utility customers are aware of the marginal price of service they face (Carter and Milon 2005). Most water customers receive total consumption and total expenditure information in their periodic bill, but marginal price information is difficult (costly) for consumers to access. Consumers have been found to respond less to marginal prices that are not included in the bill (Gaudin 2006). Modern water rate schedules can be complex, and they may only be available online or not at all, as opposed to being transparently identified within bills. It is plausible, then, that at least some consumers may attempt to decide consumption based on average price, which is essentially marginal price measured with error due either to fixed charges, a variable marginal rate (as with block rates), or both. Proponents of the average price specification are willing to accept a more complex model of consumer behavior to gain explanatory power.

Choosing a price metric is not a casual decision because marginal price and average price are not generally simultaneously consistent within the boundaries of ordinary least squares. For example, if ω is a utility-maximizing demand function and the linear specification

$$w = \alpha + \beta p + \gamma m + \delta \mathbf{P} \tag{3.3}$$

is a consistent estimator of ω, then $w = \omega'(ap, \mathbf{P}, m)$ is also utility-maximizing as long as

$$w = \lambda + \phi ap + \varphi m + \theta \mathbf{P} \tag{3.4}$$

is consistent. Even in the case that average price, *ap*, is computed from a relatively simple rate schedule with a fixed charge, *k*, such that

$$ap = \frac{pw + k}{w} \tag{3.5}$$

[3.3] and [3.4] cannot coexist, since [3.4] implies

$$w = \lambda + \phi p + \phi\left(\frac{k}{w}\right) + \varphi m + \theta \mathbf{P} \tag{3.6}$$

With *k* and *w* both variable, either ϕ = 0, which implies the spurious result that β = 0, or $\phi \neq 0$, which implies that Equation [3.3] is inconsistent (as well as heteroskedastic). Admittedly, a maximum-likelihood pair of quadratic conjugate solutions to [3.6] could be found; but their error structure would be indeterminate, and estimation would be arduous. Reconciling the two price metrics becomes even more complicated as less linear functional forms and more involved pricing policies are considered. Because of this inconsistency, advocates of marginal price specification do not always accept average price as a legitimate alternative specification (Griffin, Martin, and Wade 1981).

It is tempting to pose the question of price perception as a dichotomy between marginal price and average price. Arguably, only marginal price leads to efficient decision-making, yet how can consumers respond to a marginal cost that is unknown? The juxtaposition need not be a dichotomy, though. One alternative is to explicitly model efficient decision-making as a balance between how much water to consume and how much effort to expend in understanding rates. Empirically, this approach is likely to require additional data pertaining to effort expenditures and approaches to information discovery by consumers.

Another alternative is to take marginal price as the theoretical limit of price perception applicable to equilibrium consumption and average price as the month-to-month price metric of least cost. An implication of this "dim perception" model of marginal price is that price knowledge becomes a learning process. Adaptation to a new marginal rate may take months or years (Bushnell and Mansur 2005). In the meantime, customers may rely on the more accessible average price estimate. This research takes the second alternative, essentially specifying a marginal-price demand function within an average-price dynamic consumption equation. Among other benefits, this tack allows the demand function to be interpreted in a standard way by welfare applications without making strenuous assumptions about perfect information. Our choice of price specifications is thus made on theoretical and practical grounds, rather than on the strictly empirical basis suggested by some (Foster and Beattie 1981b).

Aggregation

Aggregation of the community demand function means treating thousands of individual choices as a single decision. When rates are multi-tiered (block rates), these choices include both quantity and price components. One way to match up quantities and prices is to study microdata on individual households and businesses. Observing household budgeting decisions has theoretical appeal, but it is not as informative for policy or project evaluation as direct observation of the community aggregate, and it magnifies statistical endogeneity (Shin 1985). The latter weakness is a consequence of predominately increasing block rate structures, leading to consumption neighborhoods wherein a small positive change in quantity will accompany a large positive change in price, spuriously diluting negative price effects.

An alternative to surveying every household and business within a community is to treat the mean of consumption as a point-mass serving as the representative consumer. In this research, the representative consumer is actually a distribution of consumption levels mapped onto the price schedule. The procedure originated with Schefter and David (1985) and has been employed with some success recently (Diakite, Semenov, and Thomas 2009; Bell and Griffin 2008a; Martinez-Espineira 2003). The introduced distributional information smoothes abrupt endogenous price changes while including more of the complex price schedule in a scalar price metric and acknowledging differential effects of rate changes on people operating in different blocks. Distributed consumption seems considerably more realistic than point-mass consumption, even though additional and potentially ad hoc distributional assumptions are usually required.

Building the Empirical Model

The Sample

The data consist of originally compiled monthly consumption, price, demographic and weather observations on 167 United States cities, each with population exceeding 25,000. The sample spans nine states (Alaska, California, Florida, Indiana, Kansas, Minnesota, Ohio, Texas, and Wisconsin) and the time horizon 1995 through 2005, for 132 possible monthly observations per city. Although expansive, the scope of the sample is constrained by the availability of historical water deliveries data, which is determined by state reporting protocols. Compared with a balanced panel of 22,044 observations, the data are 76% complete, with 16,804 observations. Summary statistics are given in Table 3.1. A detailed account of data collection and data characteristics may be found in Bell and Griffin (2008b), but a few highlights follow.

Table 3.1. Summary Statistics

Variable	Units	Obs.	Mean	Std. Dev.	Min.	Max.
Daily Use	mGal	16804	27.4	68.1	0.034	1340
Population	thousands	16804	141.7	355.6	27.6	3828.5
Commerce	$million	16804	2447.60	6640.20	142.7	81900
ResPrice	$/kGal	16804	2.65	1.57	0.00	10.09
CommPrice	$/kGal	16804	3.14	1.58	0.00	12.24
ResFixed	$/month	16804	17.39	18.53	0.00	302.00
CommFixed	$/month	15728	70.38	99.38	0.00	1132.38
CPI	rel. 1982	16804	1.746	0.128	1.503	1.992
PPI (BMNR)	rel. 1982	16804	1.395	0.111	1.25	1.736
Income	$000/year	16804	22.85	6.81	9.46	53.23
MinTemp	degrees F	16734	53.13	53.62	-3.74	81.84
MaxTemp	degrees F	16734	73.72	16.40	11.39	111.58
DryPart	proportion	16733	0.797	0.014	0.00	1.000

Price observations were gathered through electronic and personal contact with over 1,000 municipal and state agencies nationwide. Water prices for 37,159 observation-months in 319 communities were obtained, with sewer prices for 23,060 observation-months in 210 communities. Missing sewer price observations in the sample panel are estimated from a univariate regression on water prices. Price and cost variables in the analysis are sums of water and sewer prices and costs. Residential prices are those charged to 0.75-inch connections, and commercial prices correspond to 2-inch connections. As illustrated in Table 3.1, the biggest difference between the two schedules tends to be the magnitude of fixed charges.

Within the price sample, an average of 1200 gallons per month is allowed per residential customer and an average of 2700 gallons per business at no marginal charge. Approximately 85% of utility systems bill monthly, with the rest billing bimonthly or quarterly. Fixed charges are lowest on average in New England, although the region is not as well represented as South, West, and Midwest regions in the price dataset. Marginal prices are lowest in the West, perhaps paradoxical for a region associated with increased water scarcity, but reconcilable given the conventional focus of rate design on cost recovery rather than efficiency. Decreasing block rate structures are most common in the Midwest. Nominal marginal rates grew over the sample horizon faster than inflation, but fixed fees increased more slowly; so the relative proportion of water charges attributed to the volume of consumption increased from 1995 to 2005.

Aggregate delivery volumes (Daily Use) were obtained from state records for 216 utility systems over 25,833 observation-months. The existence of historical volume data is rare in the absence of a state-level reporting program, so a bias is incurred against data in the New England region, where perhaps water availability is less of a concern and data collection efforts appear weaker (based on our contacts with data sources). No observations from New England are included in the regressions due to the missing volume data. Monthly volume supplied per capita averages 6.0 thousand gallons (kGal), with Alaska averaging only 4.0 kGal and Texas averaging 7.2 kGal. Winter average supply (December and January) averages 5.0 kGal, whereas summer average supply (July and August) averages 7.8 kGal. Between 0.177 gallons (Alaska) and 0.54 gallons (Texas) are supplied per dollar earned, with a mean of 0.43 gallons per dollar. The winter average is 0.354 gallons per dollar earned, and the summer average is 0.549.

Population data are taken from the U.S. Census, personal income (Income) and nonfarm income (Commerce) from the Bureau of Economic Analysis, and inflation measures (CPI and PPI) from the Bureau of Labor Statistics. The climate measures, monthly highest and lowest recorded temperatures (Min. Temp. and Max. Temp) and the proportion of days when precipitation was recorded (Frac. Precip.) are taken from the National Climatic Data Center.

The Dependent Variable

The applicability of an EC regressor derived from a distinct structural model depends on the stationarity, or integration level, of the regressor itself and not, per se, on the cointegration status of its components. The EC technique applied to stationary series will be equally consistent and confer many of the same advantages as if applied to nonstationary but cointegrated series. Similarly, periodic EC applied to aperiodic series will produce consistent if redundant estimates. Nevertheless, the academic interest in cointegration is sufficient to merit a preliminary examination of the dependent variable, total daily quantity of water demanded per capita.

A candidate test of periodic integration is that proposed by Boswijk and Franses (1995b). Periodic integration with a single common root process (a common stochastic trend) in periodic data is equivalent to aperiodic integration in the same data stacked as an annual vector. A univariate autoregressive equation on the stacked vector can be performed functionally as an autoregression on the pooled monthly data with monthly dummies. A Dickey-Fuller-style test is performed on the vector product of the estimated autoregressive parameters. Just as in the Dickey-Fuller test, the null hypothesis is that the estimated parameter (in this case the vector product of estimated parameters) equals unity, implying the existence of a unit root, thus that the series is nonstationary. The alternative is theoretically one-sided, although a few observations in practice produce an autoregressive parameter greater than unity.

Boswijk and Franses suggest a likelihood ratio test based on imposing the unity restriction, but only the asymptotic distribution of this test is known, and the time series samples here are small. Therefore, a t-statistic on the distance between the nonlinear combination of estimated parameters and unity is presented alongside an F-test on the imposed restriction, with the understanding that results may be more demonstrative than rigorous.

The quantity demanded panel is unbalanced by missing observations, so the periodic unit-root hypothesis is tested separately for each community rather than in a pooled test.

The t-test median over all communities is 0.005, with 78% of observations greater than -1.30. The F-test median is 1.175, with 68% of observations less than 2.70. The distribution of t-scores is displayed graphically in Figure 3.1. Although critical values for both tests vary with the number of time observations in each panel, -1.30 is higher than the 95% critical value for any one-sided t-test, and 2.70 is lower than the 90% critical value for F(1,120), which is the most restrictive case among the panels tested. Therefore, individual test statistics fail to reject the null in over 70% of cases. The hypothesis that all quantity data are periodically integrated of order one cannot be rejected either. These results sound the alarm that residuals generated from a linear regression on quantity demanded will generally not be stationary in the absence of a periodic cointegrating vector. Periodic cointegration is justified on this basis.

The Structural Model

The long-run structural model employed at this stage is a Cobb-Douglas model. The Cobb-Douglas functional form is still the most popular (Basani, Isham, and Reilly 2008; Olmstead, Hanemann, and Stavins 2007; Musolesi and Nosvelli 2007). Other common forms include the semilog (Kostas and Chrysostomos 2006) and linear (Ruijs, Zimmermann, and van den Berg 2008). The Stone-Geary form also has its adherents (Gaudin et al. 2001, Martinez-Espineira and Nauges 2004).

Interpretation of the model parameters is different in the context of an EC model than it would be as a stand-alone regression. In addition to the usual explanations for nonzero residuals, such as measurement error and random innovation, disequilibrium

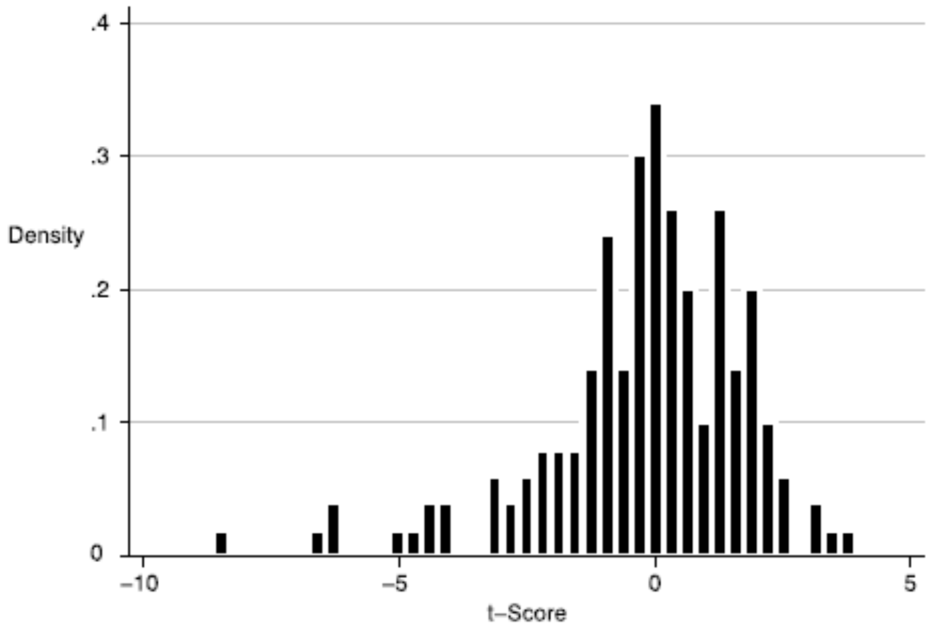


Figure 3.1. Testing for a Periodic Unit Root in Quantity Delivered

due to slow adjustment must also be included. In deference to the periodic integration results of the previous subsection, one structural model per calendar month will be estimated, not as a predictive model but to contribute long-run perspective to the dynamic model.

Covariates include weather and climate measures, residential and commercial marginal price indices, sectoral intensity ratios, and income. The weather measures are average minimum temperature, average maximum temperature, and the proportion of days in the month with less than 0.10 inches precipitation. The climate measures consist of the 30-year averages of each weather measure, by month (Av. Min. Temp., Av. Max. Temp., and Av. Dry Frac.). Personal income is taken directly from the Bureau of

Economic Analysis. Income not only reflects the contemporaneous budget constraint, it proxies the level of capital expenditure, including water-using durables. Unfortunately, the geographic boundaries of the income aggregates do not consistently correspond to the areas of municipal water service coverage. Also, mean personal income may be an insufficient statistic when the distribution of income within communities matters. Finally, personal income data is annual rather than monthly.

Marginal prices (Res. Price and Comm. Price) are adjusted for inflation and weighted across residential and commercial price schedules according to an assumed distribution (Bell and Griffin 2008a). Household consumption and business consumption within a given community are each assumed to be distributed lognormally over quantities demanded. Total, mean, and median consumption are sufficient statistics to describe a lognormal distribution. Medians of consumption for each observation are calculated so that the ratio of mean to median is identical to the ratio observed in the distribution of total consumption across observations (which is 2.48). The unique distributions so defined are mapped onto each residential and commercial rate schedule to produce a weighted marginal price index. The same weighting is applied to average prices in the formulation of the short-run average price indices, although these are summed discretely every 500 gallons from 500 to 100,000, whereas the marginal price indices are integrated continuously. None of the sample communities experiences a price that is strictly zero.

In order to represent community differences in commercial activity, commercial and industrial intensity ratios (Comm. Intensity and Ind. Intensity) are included as covariates by dividing the monetized nonfarm and industrial outputs, respectively, by population. Industry is the subsector of commerce primarily concerned with physically transformative processes, which can in many cases demand high levels of input water. Its inclusion is problematic because its distinction from other forms of commerce is arbitrary, water uses vary widely within the industrial sector, and an unknown portion of the industrial sector obtains water from wholesalers or is self-supplied. The error associated with this measure should therefore be considered underestimated by the regression. Nevertheless, as one of the more significant factors in the regression (Table 3.2), its inclusion is cautiously justified. An industrial price is not included because the combination of measurement error and collinearity with the other price measures would eclipse any reliable explanatory power.

Results of the structural regressions for each calendar month are presented in Table 3.2. The residuals of this regression, lagged one year, will constitute the EC term of the dynamic regression in differences. A hypothesis of this research is that seasonality at the monthly frequency is a consideration in water demand. If seasonality is evident, then the monthly coefficients should be significantly different from the coefficients of a pooled regression of all months. To settle this question, a Chow test is performed comparing each monthly regression to the pooled regression of the other eleven months. The appropriate statistic is $F(12, 14762)$, but for simplicity, results were compared to the 1% critical value of $F(12, \infty)$, which is 2.185. The hypothesis that all monthly parameters are indistinguishable from all pooled parameters is rejected for every month except April (1.920) and October (0.820), which is expected since a pooled average, like a broken

Table 3.2. Results of Cobb-Douglas Regression by Month

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Obs.	1207	1213	1215	1226	1229	1224	1223	1236	1245	1256	1261	1251
Variable												
MinTemp	-0.222	-0.032	-0.223	0.129	-0.554	-0.14	-0.881	0.145	-0.638	-0.478	-0.208	0.009
	0.093	*0.105*	*0.214*	*0.281*	*0.389*	*0.593*	*0.597*	*0.6*	*0.499*	*0.294*	*0.212*	*0.074*
MaxTemp	0.324	-0.136	-0.01	0.691	**2.156**	1.553	**1.866**	1.522	**1.983**	0.578	0.268	-0.306
	0.236	*0.232*	*0.327*	*0.445*	*0.44*	*0.598*	*0.615*	*0.62*	*0.622*	*0.419*	*0.266*	*0.239*
DryPart	0.054	-0.001	0.135	0.038	0.045	0.137	**0.474**	**0.298**	0.187	0.07	-0.024	**0.339**
	0.05	*0.02*	*0.121*	*0.059*	*0.042*	*0.069*	*0.09*	*0.086*	*0.077*	*0.086*	*0.064*	*0.111*
AvMinTemp	**0.312**	-0.042	-0.43	**-1.061**	-0.539	-0.621	0.484	-0.574	-0.095	-0.537	-0.446	**-0.556**
	0.092	*0.149*	*0.272*	*0.321*	*0.413*	*0.639*	*0.651*	*0.646*	*0.586*	*0.35*	*0.266*	*0.173*
AvMaxTemp	-0.45	0.347	**1.121**	1.197	0.383	1.432	1.339	1.567	0.649	**1.518**	0.87	**1.254**
	0.253	*0.278*	*0.399*	*0.503*	*0.516*	*0.666*	*0.652*	*0.681*	*0.738*	*0.515*	*0.353*	*0.327*
AvDryPart	**-0.143**	-0.13	-0.07	0.003	**-0.049**	**-0.091**	**-0.068**	**-0.081**	**-0.099**	-0.059	0.006	**0.224**
	0.053	*0.057*	*0.051*	*0.037*	*0.018*	*0.018*	*0.016*	*0.018*	*0.027*	*0.031*	*0.043*	*0.059*
ResPrice	**-0.123**	**-0.112**	**-0.121**	**-0.151**	**-0.164**	**-0.155**	**-0.175**	**-0.178**	**-0.191**	**-0.196**	**-0.142**	-0.061
	0.042	*0.042*	*0.041*	*0.042*	*0.042*	*0.052*	*0.049*	*0.05*	*0.057*	*0.044*	*0.039*	*0.04*
CommPrice	-0.084	-0.089	-0.105	-0.114	**-0.138**	-0.139	-0.141	-0.136	-0.14	-0.112	**-0.123**	**-0.171**
	0.052	*0.051*	*0.048*	*0.047*	*0.047*	*0.085*	*0.056*	*0.056*	*0.063*	*0.05*	*0.045*	*0.05*
CommIntensity	0.065	0.006	-0.03	0.059	0.172	0.152	**0.236**	**0.277**	0.185	0.157	0.118	0.137
	0.068	*0.068*	*0.069*	*0.07*	*0.07*	*0.087*	*0.085*	*0.085*	*0.095*	*0.078*	*0.067*	*0.068*
IndIntensity	**0.11**	**0.106**	**0.118**	**0.103**	**0.118**	**0.115**	**0.118**	**0.116**	**0.099**	**0.104**	**0.099**	**0.103**
	0.014	*0.014*	*0.014*	*0.014*	*0.014*	*0.017*	*0.017*	*0.017*	*0.019*	*0.015*	*0.013*	*0.014*
Income	**-0.31**	**-0.254**	-0.153	-0.142	-0.222	-0.116	-0.141	-0.159	-0.089	-0.145	**-0.276**	**-0.33**
	0.094	*0.094*	*0.096*	*0.097*	*0.095*	*0.12*	*0.118*	*0.118*	*0.13*	*0.107*	*0.093*	*0.093*
Constant	**7.867**	**6.704**	**4.374**	**2.02**	0.29	-4.173	**-6.58**	**-5.899**	-3.012	1.158	**5.33**	**6.645**
	0.879	*0.939*	*0.972*	*1.076*	*1.174*	*1.609*	*1.619*	*1.548*	*1.585*	*1.142*	*0.888*	*0.904*

clock, should still be right twice per cycle. These results indicate the unlikelihood of a constant demand relationship, and they support the probability of 12-phase (monthly) seasonality, but they do not preclude the possibility of other intra-annual (such as 4-phase) or extra-annual (such as El Niño) cyclic frequencies.

A cursory examination of Table 3.2 reveals that residential price and industrial intensity are the most consistently significant covariates. The price signal is generally stronger in the warmer summer months. Residential demand seems to be more sensitive to seasonality than commercial demand. Although the residential and commercial mean price elasticities of -0.147 and -0.124 are low (in absolute value), their combined mean of -0.272 is consistent with previous research (Dalhuisen et al. 2003). Evidence supporting the hypothesis that water-consuming sectors should be treated separately can be drawn from the sectoral intensity variables and by comparing the effects of the two price indices. Although industrial intensity appears to figure significantly in all months, commercial intensity is only significantly positive in July and August. The residential and commercial price variables (whose pairwise correlation is 0.831) exhibit effects that appear to be generally similar and that are in fact statistically indistinguishable in every period. The null hypothesis that residential and commercial consumption can be adequately described in a single-sector model cannot be rejected, although the evidence supports identification of a separate industrial sector.

Personal income enters negatively, which is an unexpected result. The negative income effect could be essentially spurious, resulting from the disappointingly high level of income aggregation, or it could reveal a higher stock of political capital in more affluent communities. It is possible that such communities could exert a monopsonistic influence on price and quantity supplied; however, the lack of corroborating evidence from prior literature casts doubt on this explanation. Average maximum temperature and average minimum temperature coefficients frequently carry opposite signs, indicating that temperature spread is an important determinant of water demand.

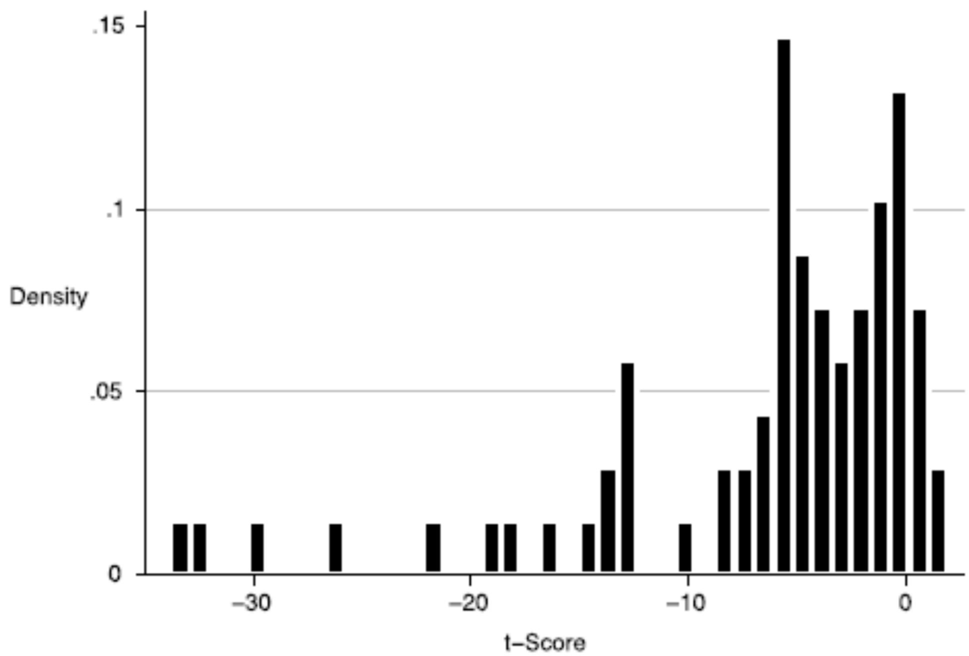


Figure 3.2. Testing for a Periodically Integrated Residual

The distribution of t-statistics testing the null hypothesis of periodic integration in the residuals is illustrated by Figure 3.2. The median t-score is -26.53, with 85% of test statistics lying outside the 95% (one-sided) probability interval of the null, allowing a handy rejection of the hypothesis that the residuals are systematically periodically integrated. The median F-test statistic is 468.0; 83.4% of statistics lie outside the 90% interval. With some caution, it may be said that the left- and right-hand sides of each structural relation are periodically cointegrated. Although the regressions summarized in Table 3.2 are consistent periodic cointegration vectors, they should not be taken as indicative of observed water demand behavior because they omit important short-run components to be addressed in the dynamic model.

The Dynamic Model

The short-run model is a logarithmic model in annual first differences, augmented with pairwise products of covariates. The dependent variable, $\Delta \ln w$, is the annual log difference in daily consumption per capita:

$$\Delta \ln w = \beta_p \Delta \ln \mathbf{p} + \sum_{i=2}^{I} \beta_i \Delta \ln x_i + \sum_{i=2}^{I} \gamma_{pi} \Delta(\ln \mathbf{p} \ln x_i) + \sum_{i=2}^{I} \sum_{j=2}^{I} \frac{\gamma_{ij}}{2} \Delta(\ln x_i \ln x_j) + \delta EC + v \qquad [3.7]$$

Prices, **p**, are annotated separately from the other covariates, **x**, for clarity only. Except for the error correction term, *EC*, the model is taken from Bell and Griffin (2008a). In contrast to the earlier application, both residential and commercial prices are included in the present model. A product of the two prices is not included because its inclusion would obscure the price elasticity calculations. Also, the price metric here is average quasidifference price rather than marginal quasidifference price. The quasidifference price is the difference between two price schedules, weighted by the distribution of lagged consumption. By weighting contemporaneous and lagged price schedules identically, the spurious endogeneity of a quantity-dependent price index is avoided. Only residential and commercial prices, weather realizations, and the EC term are included in this short-run regression since climate, income, and sectoral intensity are assumed to change too slowly to drive an annual model.

Table 3.3. Results of Logarithmic Dynamic Regression

(n = 12547; $p < 0.01$ in bold type)

Variable	Coefficient	Std. Error	t-Score
ResPrice	-0.00194	0.0304	-0.06
CommPrice	0.00471	0.0140	0.34
MinTemp	**-0.579**	0.0789	-7.34
MaxTemp	-0.110	0.144	-0.76
DryPart	-0.350	0.240	-1.46
ResPrice*MinTemp	**-0.0500**	0.00682	-7.28
ResPrice*MaxTemp	**0.0480**	0.00637	7.54
ResPrice*DryPart	0.000375	0.00203	0.18
CommPrice*MinTemp	**0.00703**	0.00137	5.13
CommPrice*MaxTemp	**-0.00733**	0.00137	-5.35
CommPrice*DryPart	0.0000355	0.000594	0.06
MinTemp*MaxTemp	**0.174**	0.0281	6.20
MinTemp*DryPart	0.0753	0.0736	1.02
MaxTemp*DryPart	0.0184	0.110	0.17
EC	**-0.187**	0.00519	-35.99

Estimation results are presented in Table 3.3. The last year of data is withheld from the estimation to test prediction accuracy. The EC term is highly significant, the most significant coefficient in fact, indicating that the pull to equilibrium is a motive force affecting quantity demanded at the annual level. Not only is the coefficient (-0.187) different from zero, it is different from -1, implying a multi-year adjustment path. It is noteworthy that the EC coefficient could reflect behavior other than consumption, such as mitigation of system losses by utilities.

Price covariates are significant only when paired with weather covariates (such as Res. Price*Min. Temp), indicating the effect of the weather on price response. The frequency of precipitation (Dry Frac.) does not appear to be as important in the short run as temperature. Mean high and low temperatures appear to be the motive force behind demand in the short run, not only in themselves but also in governing the effect of price.

Short-run elasticity is composed of two elements for each sector, the immediate response to a price shock and the momentum of adjustment to previous shocks. The shock response is computed with respect to the estimated parameters corresponding to that sector's price index according to the formula

$$\varepsilon = \beta_p + \sum_{i=2}^{I} \gamma_{pi} \ln x_i \qquad [3.8]$$

The adjustment elasticity, embedded in the EC component, can be understood as the proportion of long-run elasticity distributed to each period. Adjustment elasticities are products of the long-run price coefficients reported in Table 3.2 and the EC coefficient reported in Table 3.3.

Table 3.4 shows estimated short-run price elasticities derived by the dynamic model, grouped by month. The theoretical annual elasticity of demand due to a simultaneous price change in both residential and commercial schedules (Annual) is decomposed into sectors, with each sectoral elasticity (Res. Total and Comm. Total) further separated into contemporaneous (Beta) and lagged (EC) components. The

Table 3.4. Annual Average Price Elasticity by Month and Sector (Values are not Significantly Different from Zero)

		Residential			Commercial			
	Obs.	Beta	EC	Total	Beta	EC	Total	Annual
Jan	1227	0.016	-0.023	-0.007	0.000	-0.016	-0.016	-0.023
Feb	1226	0.013	-0.021	-0.007	0.000	-0.017	-0.016	-0.024
Mar	1228	0.011	-0.023	-0.012	0.001	-0.020	-0.019	-0.031
Apr	1239	0.009	-0.028	-0.020	0.001	-0.021	-0.020	-0.040
May	1242	0.006	-0.031	-0.025	0.001	-0.026	-0.025	-0.050
Jun	1238	0.004	-0.029	-0.025	0.001	-0.026	-0.025	-0.050
Jul	1237	0.004	-0.033	-0.029	0.001	-0.026	-0.025	-0.054
Aug	1250	0.004	-0.033	-0.029	0.001	-0.025	-0.024	-0.053
Sep	1259	0.005	-0.036	-0.030	0.001	-0.026	-0.025	-0.055
Oct	1271	0.007	-0.037	-0.029	0.001	-0.021	-0.020	-0.049
Nov	1275	0.009	-0.027	-0.017	0.001	-0.023	-0.022	-0.040
Dec	1265	0.014	-0.011	0.003	0.000	-0.032	-0.032	-0.029

contemporaneous components are slightly positive and generally smaller than the lagged components, contributing little or no price effect. From the EC coefficient reported in Table 3.3, each of the annual disequilibrium elasticity estimates is approximately 18.7% of the total estimated structural elasticity. None of the short-run elasticity means is statistically negative, owing primarily to variation in the data. It appears that communities take a minimum of one year to notice a price change and only begin to react in the second year.

Table 3.5. Comparison of Model Predictions

		(n = 1612)		
Model	Observed	EC	Structural	Dynamic
Mean (gal)	209.9	182.8	163.7	193.8
MAPE	.	21.6	41	21.7
MSE	.	14436	26158	21998

The dynamic EC model is used to project each 2004 observation one annual step forward. The results reported in Table 3.5 include predicted mean daily consumption per capita, mean absolute percent error (MAPE), and mean squared error (MSE) of the preferred model (EC), compared to the observed 2005 data mean and two benchmarks. The benchmarks include predictions from the monthly structural models and from the dynamic model re-estimated without an EC term. The EC model (as well as the non-EC dynamic model) clearly outperforms the structural models on all measures, illustrating the importance of temporal consideration in modeling water demand. The EC model only marginally outperforms the non-EC dynamic model. Inclusion of the EC term is arguably recommended on the grounds of avoiding theoretical misspecification rather than on predictive grounds.

Conclusions

An atypically broad panel of monthly demand data for publicly-supplied water in American urban centers is analyzed using a periodic error correction model. Although the model can be applied to household data as well as community data, the independent variables used in the analysis mitigate the relative weakness of aggregate data. A micro-level approach can be pursued by more conventional means when such data are available.

The EC model allows examination of the time path of demand by integrating shorter and longer perspectives in a single estimation model. The estimated effect of the lagged residual implies that demand adjustments are not instantaneous or even as quick as a single year. The model significance and predictive power of the dynamic model in annual differences allows a rejection of the possibility that a purely structural model is well specified for time-dependent applications. Estimation of distinct structural relations for each calendar month allows a test of seasonality of demand. Rejection of the null hypothesis that structural parameters are equivalent across months suggests that ignoring seasonality can lead to misspecification, even when weather and climatic factors are taken into account. Inclusion of both residential and commercial price indices, as well as commercial and industrial intensity ratios, tests the adequacy of the more common single-sector model. Although some evidence suggests that businesses, especially industrial businesses, demand water differently than households, the single-sector model is not conclusively rejected.

Some new possibilities are suggested by the results. Ignoring the distinction between the sectors may be unwise. Social costs arising from temporary misalignment of supply and demand may be reduced by adjusting residential rates and commercial rates differentially to the same target price. Many rate setters change the entire rate schedule uniformly, either for convenience or out of an interest in intersectoral equity. Recognizing that tensions exist inherently among competing objectives is a necessary step to striking good balances. For example, the dictum of efficiency requires that all sectors face natural water's opportunity costs, and these should be locally equivalent across sectors. Thus, efficiency is best advanced by equal rates except where marginal processing costs differ sectorally.

Demand factors appear to be seasonal in a way that is not entirely captured by climatic conditions. This finding speaks to the use of seasonal management policies, especially if a risk of acutely exceeding peak capacity is present. Short-run demand response may be only marginally significant in summer months, but it is very close to zero in nonsummer months. Also, annual elasticity appears to be much lower than structural elasticity, indicating a very slow adjustment process. Managers cannot expect an immediate readjustment to changing conditions. Applying these findings may require a deeper consideration of the habits and durable possessions of water users, which are both seasonal and slow to evolve. Because of the dim perception consumers have of marginal water cost, information must also be counted as a valuable capital good.

The periodic error correction model produces near-term forecasts with an appealingly low level of error (21.6% MAPE), even though it may not be the best model for projecting future conditions or for testing hypotheses regarding price elasticity. The model and the exercise of developing the model underscore the inadequacy of the term "price elasticity". Many elasticities have been generated in this research alone, varying with time horizon, season, sector, and model. Meaningful comparison and application of these estimates depends on an explicit characterization of which elasticity is to be derived.

Community water consumption series could contain a seasonal stochastic trend, as the series in this sample apparently do. If this is the case, OLS estimates of demand cannot be assumed consistent. Fortunately, the data of the present sample are seasonally cointegrated. Unit root tests are available to assist in the determination of trend stationarity, as are seasonal and periodic integration tests to determine cyclical patterns. When demand relations tend to equilibrate slowly, an integrated structural/dynamic model such as an error correction model will provide improvements in both forecasting and insight over either a static or a dynamic model alone. Finally, as the accuracy of these findings is limited by the quality of the available data, it will be interesting to see if similar findings persist as data recording becomes more widespread, more uniform, and more precise.

CHAPTER IV

EMPIRICAL DERIVATION OF A WEIGHTED BLOCK RATE INDEX

Charges for publicly provided nonmarket goods such as water, natural gas, and electricity can be more complicated than just a single price per unit consumed. The use of multipart rate schemes by public monopolists may be motivated by revenue sufficiency (Boiteux 1971) or equity objectives (Diakite et al. 2008). A prevalent multipart pricing policy examined in this research is the block rate schedule, in which different marginal rates are assigned to different blocks of quantities consumed. Demand modelers struggle with block rates because price uniqueness is fundamental to consumer theory, yet price is not unique in a block rate schedule. The central problem of demand estimation specific to block rates is whether to choose a price among multiple candidate rates (Moffitt 1986; Hausman 1985), or to somehow reduce the rate schedule to a scalar price index (Agthe et al. 1986).

A convenient method of assigning a price to a household consumption observation is to use the rate effective in the block containing that household's consumption quantity (Chicoine and Ramamurthy 1986). Consumption, the dependent variable, is stochastic, so consumption by the same household could fall in a different block in the next period under exactly the same conditions. This would imply that price is also stochastic, as well as dependent on quantity consumed. Inferences drawn from regression demand models are only valid if price is independent of quantity demanded, so the convenient price representation leads to inconsistent demand estimation.

Perplexingly, price exogeneity under block rate schedules depends on how price is defined. The rate of service is functionally tied to the level of consumption and is therefore endogenous with respect to quantity consumed, but the schedule itself is a document that is published prior to consumption and is therefore exogenous. Unfortunately, many demand models allow price to vary when the schedule remains unchanged. If these models are applied to data from a single utility service area, the primary source of observed price variation is consumer behavior. Even though many such models provide econometric techniques to unbias their results as though the changes had been exogenous, the problem is that they model a process that is itself endogenous.

Convenient price representation is even more problematic when applied to aggregate data (Agthe and Billings 1980). Quantity demanded may be modeled on a per-capita basis, but businesses and institutions also contribute to aggregate demand, so average consumption per household may not represent the community's exposure to the rate schedule very well. Furthermore, a community is not a single household with a single budget constraint, so it is unlikely that the whole community will respond to just one segment of the rate schedule. Aggregate demand is more realistically modeled as a distribution of behaviors than as a discrete choice. Since the data modeled in the empirical section of this paper is aggregate in nature, issues of aggregation will be focal in the discussion of alternative price representations.

Evaluation of price representation methods in the literature has been largely theoretical so far. Although price methods have been compared based on the assumption that the most explanatory variable choice is preferred (Shin 1985), none has been empirically tested against an objective standard. The present research provides such a test by asking what continuous price index would generate values to fit an observed inverse demand function. The procedure exploits a large national dataset of community-level water demand featuring an atypically high degree of variability in rate magnitudes, numbers of blocks, and locations of block boundaries (Bell and Griffin 2011).

Some of the rate schedules in the sample are composed of only one volumetric rate. Since these observations have a unique price, they are used to estimate demand and inverse demand functions. Interpreting multiple-block schedules as linear combinations of single-block schedules, a weighting function can be identified from the inverse demand function for any schedule of two blocks. Refinements could theoretically be gained by iterating the procedure for schedules of three blocks and more, but error is magnified for each estimate based on another estimate. Schedules of more than two blocks are used to validate the method instead.

The weighted price index derived in this research is qualitatively different from other price specifications under block rates because it originates from an empirical demand function rather than from theory or convenience. It is necessarily also a product of the auxiliary assumptions used in its derivation (Davis 2004). An example of a restrictive auxiliary assumption common to demand estimation is the constant elasticity assumption, implemented through a Cobb-Douglas functional form of demand. More flexible empirical models have rejected constant elasticity (Bell and Griffin 2008), yet it remains popular in practice (Olmstead et al. 2007). Demand and distribution functions in this research are estimated nonparametrically to reduce the number and severity of auxiliary assumptions. Where parametric estimators fit a member of a family of global functions to the data, nonparametric estimators fit a series of local functions. Values predicted from a nonparametric regression are consequently closer to observed values (error variance is reduced), although computational demands are increased and prediction outside the observed range is sacrificed altogether. The trade-off magnifies the role of observational data in hypothesis testing, but restricts the range of hypotheses that can be tested.

The proposed empirical price index assumes only nonnegativity and continuity of the weighting function and the existence of a continuous demand function with a continuous inverse, whereas other models have implicitly employed discrete or point distributions and restrictive functional forms. The index is invariant to demand shifts on a given rate schedule. This property insures that only exogenous price changes are measured. The primary source of exogenous price change is updating of the rate schedule, although de facto inflationary changes are also present in the data. The model is fitted to aggregate data and requires only commonly available data aggregates.

Notable prior research on demand under block rates is reviewed in the next section, with an emphasis on applicability to aggregate data. A new solution is proposed in the third section and estimated in the fourth. Conclusions are drawn in the final section.

Literature Review

The beginnings of multipart pricing were controversial. Hotelling (1938) argues that public monopolists, even under declining costs, can only maximize social welfare by charging a uniform rate of service equal to the long-run marginal cost of that service. For Hotelling, financing capital costs through general taxation is socially preferable to financing through fixed charges on commodities, and "nothing could be more absurd" than to equate the social value of an enterprise with its repayment of overhead costs. Coase (1946) directly challenges this view, arguing in essence that a price system is as useful in determining which goods to provide as it is in determining the quantities to be provided. Since public monopolies have to answer both questions, the best pricing policies for them involve two parts. The "parts" are conceived as a unit price, equal to marginal product cost, and a fixed charge sufficient to balance the overhead costs necessary to the enterprise. Although fixed charges could be positive or negative (Griffin 2001), negative fixed charges are unknown in practice.

Gabor (1955) shows that a schedule of two unit rates corresponding to two blocks of consumption can achieve the same revenue control as a single rate with a fixed charge. Rate schedules consisting of fixed charges and multiple block rates were already prevalent in electricity provision by the time of Gabor's writing (Houthakker 1951), but Taylor (1975) is credited with the original theoretical treatment of demand under block rates. Taylor defines marginal price conveniently, as the rate applicable to a consumer's marginal consumption, and addresses primarily the problem of budgeting. The possibility of facing more than one price for the same good means that the consumer's budget line is kinked. Consequently, an equilibrium quantity demanded may not be unique or analytically derivable. To triangulate across budget segments, Taylor suggests using both marginal and average price variables in empirical applications.

Nordin (1976) rejects Taylor's use of average price, proposing instead a "difference" variable corresponding to the net loss (or gain) in income between a purchase under a block rate schedule and the same purchase under a constant price policy. This position is based on the inability of an average price specification to produce a unique quantity demanded solution to the classical consumer problem. Average price suffers more than marginal price from endogeneity with quantity and is not well defined on the domain of possible outcomes, i.e. zero consumption. Nordin fails to address the possibility that some consumers, faced with limited information, might use average price to proxy marginal price (Foster and Beattie 1981). If this were true, fixed charges and rate elements encountered before the margin would have a "perception" effect on demand. Indeed, empirical studies have not generated estimates of Nordin's difference variable consistent with a pure income effect. The average-price/marginal-price controversy is interwoven with the question of price representation in the water and utility literature (Bell and Griffin 2011), but it will not be dealt with in detail here.

Agthe et al. (1986) reject the exogeneity of their convenient price representation for aggregate data and introduce a simultaneous-equations approach with instrumental variables (IV). IV estimators are biased but asymptotically consistent (Deller et al. 1986), so they are seen by some researchers as preferable to inconsistent yet convenient estimators. Price instruments can include billed charges and quantities, block rates, and/or miscellaneous regressors such as climate and income. Unfortunately, Agthe et al. use only a binary variable to represent a change in rate schedule, so their research is rather insensitive to truly exogenous price change. IV price estimation can be applied to either household or aggregate data, but does not bypass the fundamental feedback present in the block-rate system. Applied to aggregate data, some IV procedures assume that all households consume on the same block, whereas others weight all block rates equally (Deller et al. 1986). IV prices are linear predictions, so they do not always take on plausible values, especially at data extrema (Terza 1986). They can only perform as well as the instruments selected.

A more sophisticated treatment first identifies a probable rate block using a maximum likelihood (ML) estimator then estimates quantity demanded within the block boundaries. This model has been called endogenous sorting (Reiss and White 2005) or discrete/continuous choice (Hewitt and Hanemann 1995), but will be referenced here as ML (Moffitt 1986). In addition to ML search routines, probit (Terza 1986) and logit (Corral et al. 1998) regressions have been used to model block choice. The weakness of ML is its strong theoretical structure. If limited-dependent-variable estimators are not used, ML is susceptible to intervals of negative probability at block kinks. A two- or three-part error structure is necessary to confine quantity demanded to the chosen block. The error components must be distributed normally, but consistency is hard to test because the components are not observationally identified (Moffitt 1986). Nonnormality of the data has been cited as a source of error in ML models (Blomquist and Newey 2002). Estimating the discrete block choice requires many observations on the same rate schedule, so the method is unsuitable for aggregate data or changing rate schedules (Corral et al. 1998). More intuitively, it is doubtful that real consumers possess the full information and hyper-rationality of the ML demand mechanism.

Schefter and David (1985) propose an average marginal price specification conceived as a weighting of block rates by the "proportion of households in each rate block." Schefter and David's model is motivated by the use of aggregate data, where a distinction must be made between mean marginal price and marginal price at mean consumption. The two statistics are only equivalent under the assumption of symmetrical distribution. An unbounded symmetrical distribution cannot precisely represent consumption, because consumption is bounded by zero. It may be preferable, therefore, to introduce additional information rather than to assume normality. Following this logic, distributional information on household consumption has been used to estimate a weighted price index (Martinez-Espineira 2003), and to aggregate over logit probabilities (Corral 1998). The "proportion of households" interpretation is intrinsically endogenous, though (Diakite et al. 1998). Diakite et al. (2008) aggregate over volume proportions instead of household proportions, but then reestimate their price index in an IV model to produce (asymptotically) consistent estimates. Bell and Griffin (2011, 2008a) approximate the distribution of consumption over blocks as loglinear. This assumption imbues the resulting price index with a number of desirable qualities, but lacks empirical or theoretical justification.

Restrictions are employed implicitly in both ML and IV estimations in the form of distributional assumptions on quantity demanded and the functional form specification of demand. These assumptions may be severe enough to dictate the results of estimation. Nonparametric smoothing is introduced to demand analysis by Varian (1982), as an alternative to reliance on functional form specifications that arise from convenience rather than theory. Varian shows that a finite collection of demand data is rationalizable under general conditions (when the data satisfy the Generalized Axiom of Revealed Preference), so that a smoothing of individual data points easily fits into received demand theory. "Smoothing" means that nearby data points are connected by a series of continuous local functions, instead of being fitted to a global function (such as a line).

Blomquist and Newey (2002) suggest a nonparametric estimation based on local power series that is applicable to the block-rate problem. By imposing continuity rather than a functional form, small local nonconvexities are allowed in the ML estimator without the regions of negative probability that a kinked line would produce. The result is a discrete choice estimate sharing the statistical properties of the data, rather than imposed normality. Unfortunately, the Blomquist and Newey model requires consumption observations on each block segment and cannot adapt to changes in the rate schedule. Residential water demand (Nauges and Blundell 2001) and labor supply (Wu 2002) have been modeled along the same lines. Newey and Powell (2003) present a nonparametric IV model based on series approximation that estimates over two stages to ensure continuity of the reduced inverse function. In the Generalized Additive Model (GAM), the local function is cubic (Hastie and Tibshirani 1990). Smoothness is inversely proportional to the second derivative of the local cubic, so increasing a penalty on the second derivative results in a smoother approximation.

The model introduced in the next section merges the idea of a weighted price index with the technique of nonparametric estimation. Demand data are smoothed in a GAM inverse demand function that is forced into a bijection, so that both demand and inverse demand are continuous functions. In the case where rate schedules consist of two blocks, the inverse demand estimate is interpreted as a linear combination of the two rates. A cumulative density function is derived from these results, the derivative of which becomes the weighting index over quantity blocks.

Standard ML estimation assumes a normal distribution of consumption over the rate schedule. Bell and Griffin (2011, 2008a) impose a lognormal distribution. Both of these distributions will be tested against the empirical findings. Few of the price estimators reviewed in this section are adaptable to aggregate data or rate schedule variation, but an IV estimator using three benchmark volumes as instruments will be compared to the derived estimator in both parametric and nonparametric regressions.

Theoretical Model

The effective marginal price of a single-block rate schedule is not controversial. Given full information and nonzero consumption, the uniform rate satisfies standard definitions of price. Principally, all quantities are bought and sold at the uniform rate. The full information assumption may be problematic inasmuch as fixed charges obscure the marginal price signal, so fixed charges may be considered in this model as a separable covariate. This convenience may be an oversimplification.

Call the uniform rate p, and assume the existence of a general aggregate function for quantity demanded, q, in p and other covariates, x:

$$q = f(p, x). \qquad [4.1]$$

Price is not defined yet for rate vectors of dimension greater than one. The goal of this section is to find a price index, $\tilde{p}(R)$, that satisfies $q = f(\tilde{p}, x)$ for any rate vector R, regardless of dimension. The step function R can be represented as a series of rates, P, and their left-hand termini B, so that $R^1 = \{p, 0\}$ and $R^N = \{(p_1, p_2 \ldots p_N), (0, b_1 \ldots b_{N-1})\}$.

For two-block schedules, $R^2 = \{(p_1, p_2), (0, b)\}$, let $\tilde{p}(R^2) = \alpha_1(b) p_1 + \alpha_2(b) p_2$. It is no loss of generality to represent an arbitrary scalar as a linear combination of two other scalars, but the representation does assume that $\tilde{p}$ is unique for any combination of (p_1, p_2, b). Two-block rate vectors are isolated here because the relation between the rate elements and the price index can be reduced to a function in the single boundary point, b. First, the existence of $\tilde{p}(R)$ is assumed and properties are assigned to it that justify calling it a price index and facilitate the estimation.

The first desirable property of a price index is its identity with respect to unique price: $\tilde{p}(p, 0) = p$. The index should at minimum map a uniform rate schedule to itself, since a uniform rate has every characteristic of a traditional price. Secondly, the negative own-price effect on water demand has been verified by numerous empirical studies (Espey et al. 1997), so a water price index should also satisfy the law of demand. The demand function in Equation [4.1] obeys the law of demand if $\partial f/\partial p < 0$, so the second property stipulated for $\tilde{p}$ is that it mimic the law of demand: $\partial f/\partial \tilde{p} < 0$. Thirdly, it is stipulated that $p_1, p_2 ... p_N \geq 0$ implies $\tilde{p} \geq 0$. Negative rates are not observed in the data, so a negative price effect is implausible. As a more practical concern, the nonparametric estimator is ineffectual outside the range of observations.

Additional properties of an ideal price index follow from the identity, demand, and nonnegativity properties:

(1) $\alpha_1 + \alpha_2 = 1$ for any b. If $\alpha_1 + \alpha_2 \neq 1$, let $p_1 = p_2 = p$. Then $\tilde{p} = (\alpha_1 + \alpha_2)\, p \neq p$, which violates the identity property.

(2) $\tilde{p} = \alpha_b p_1 + (1 - \alpha_b) p_2$, where α_b denotes $\alpha_1(b)$, follows directly from (1).

(3) $\alpha_0 = 0$, since $\{(p_1,p_2),(0,0)\} = \{p_2\}$.

(4) $\alpha_\infty = 1$, since $\lim_{b \to \infty} \{(p_1,p_2),(0,b)\} = \{p_1\}$.

(5) $\tilde{p}, p_1, p_2 \geq 0$ implies $0 \leq \alpha_b \leq 1$. If $\alpha_b < 0$, then any combination of p_1, p_2, such that $\dfrac{p_1}{p_2} > \dfrac{\alpha_b - 1}{\alpha_b}$ would imply $\tilde{p} < 0$, which would violate nonnegativity. A parallel argument holds for $1 - \alpha_b < 0$.

It is clear from Property (5) and the representation of $\tilde{p}(R^2)$ as a linear combination that $\partial \tilde{p}/\partial p_1 \geq 0$ and $\partial \tilde{p}/\partial p_2 \geq 0$. To facilitate the transition from $\tilde{p}(R^2)$ to $\tilde{p}(R^{3+})$, the generalization is made that $\partial\tilde{p}/\partial p_n \geq 0$. Then, comparing R^2 to $R^{2\prime} = \{(p_1,p_2),(0,b+\delta)\}$ as though a small rate block between b and $b+\delta$ held the value p_δ, $\tilde{p}(R^2{}') - \tilde{p}(R^2) = \alpha_\delta(p_1 - p_2)$. Therefore, $\Delta\tilde{p}/\Delta p_\delta = \alpha_\delta \geq 0$. As δ approaches zero,

(6) $\dfrac{\partial\alpha_b}{\partial b} \geq 0$, and so,

(7) α_b is a cumulative density function on $b = [0, \infty)$ (Properties 3, 4, 6).

α_b is estimated in a GAM over the two-block data subset, smoothed only as necessary to impose (6). The dependent variable is defined as

$$\alpha_b = \frac{\hat{p} - p_2}{p_1 - p_2}. \qquad [4.2]$$

$\hat{p}$ is interpolated from the single-block inverse demand function. If $f(p,x)$ is one-to-one, then $\hat{p} = f^{-1}(q;x)$ will be single-valued. $\hat{p} = g(q,x)$ is estimated in a GAM using the minimum smoothing necessary to insure this condition, leading to a unique prediction of $\hat{p}$, and therefore a unique α_b for each observation. In a GAM demand model, the sum of squared differences between local functional values and observed quantities is minimized subject to a "smoothness" penalty, s, on the second derivative of the local function (Hastie and Tibshirani 1990, p. 27):

$$\sum_{i=1}^{n}\left[q_i - f(p_i,x_i)\right]^2 + s\int_{a_0}^{a_1}\left[f''(t)\right]^2 dt \qquad [4.3]$$

where n observations in the neighborhood $[a_0, a_1]$ form the local support. The existence of a functional estimate of either inverse demand or cumulative density depends on underlying monotonic trends in the data, which fortunately exist in both cases. Numerical differentiation produces the weighting function, $\omega(b) = \partial\hat{\alpha}_b / \partial b$, such that

$$\tilde{p} = \int_0^\infty \omega(z)P(z)dz. \qquad [4.4]$$

Empirical Model and Results

Step 1: A Demand Function in One Price

The demand function to be estimated is,

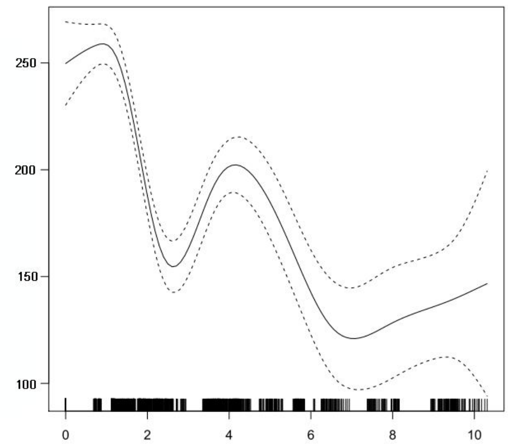
$$q = f(p, k, m, c) \qquad [4.5]$$

where q is quantity demanded in gallons per capita per day, p is uniform price, k is a fixed monthly charge, m is annual personal income, and c is a climate variable composed of mean high temperature plus mean low temperature, times the proportion of days where no precipitation above 0.10 inches was recorded. The data, consisting of 8117 community-months, are summarized in Table 4.1. They are obtained through a campaign of internet searches and follow-up contacts with municipal, district, and state sources conducted by the author, augmented with national weather and economic values. The data were previously employed by Bell and Griffin (2011).

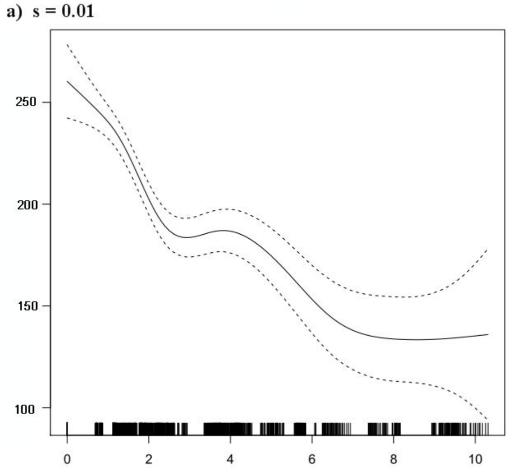
The dependent variable is defined as the quantity supplied per day divided by the population, although it is recognized that this definition discounts the effect of business activity on aggregate demand (Bell and Griffin 2011). Prices and monthly charges are combined for water and sewer service, both of which are based on the same monthly metering. Bell and Griffin (2011) estimate missing sewer rates when water rates are known, but these observations have been omitted in the present research. Prices, charges, and income are normalized to 2005 price levels. It may be seen from Table 4.1 that all variable means for each block subsample are statistically indistinguishable from the data means, although fixed charges are noticeably greater in the one-block subsample than in the others.

Table 4.1. Descriptive Statistics by Numbers of Blocks (*Standard Deviations in Italics*)

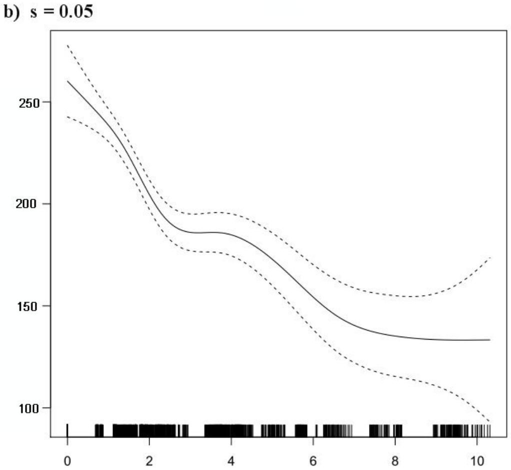
Variable	Units	Single Block	Two Blocks	3 -8 Blocks	All
Water Use	Gallons	211.13	164.64	173.69	197.76
		152.91	*80.25*	*91.36*	*157.59*
Fixed Charges	2005$	34.82	16.96	18.79	20.72
		56.14	*8.55*	*8.70*	*24.96*
Income	2005$	26,968	24,718	25,517	25,903
		8152	*4851*	*7393*	*7283*
Weather Index	Degrees F	105.47	103.42	107.66	106.21
		32.11	*31.78*	*25.67*	*29.14*
Observations		1680	2565	3872	8117



a) s = 0.01



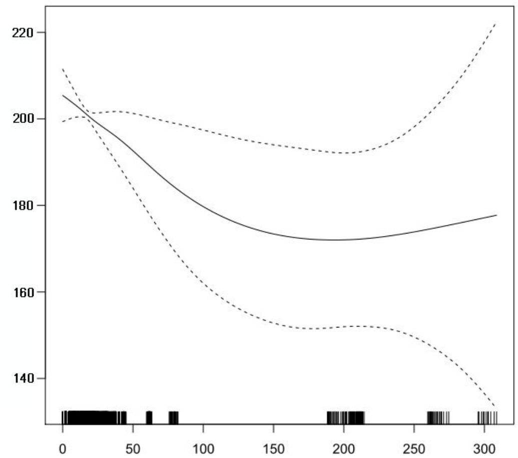
b) s = 0.05



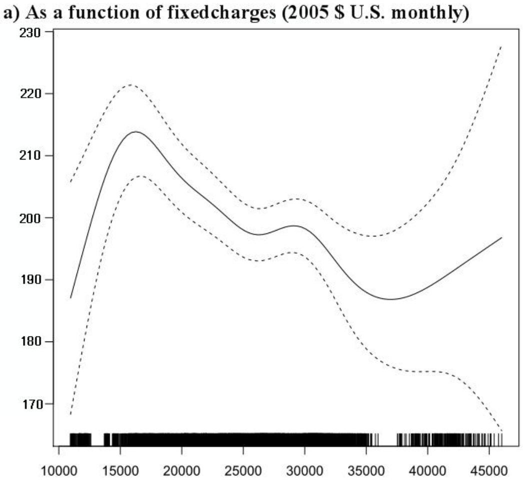
c) s = 0.07

Figure 4.1. Quantity Demanded as a Function of Price, Three Different Smoothing Parameters

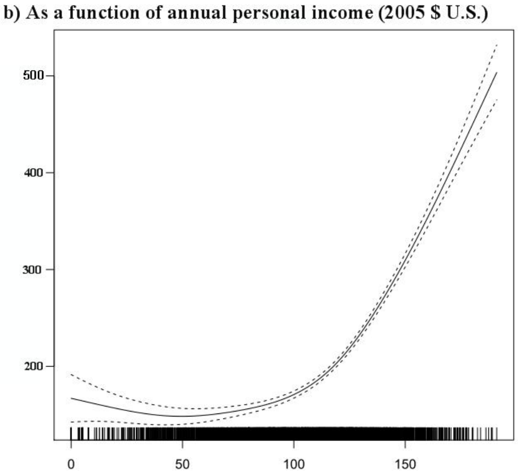
Vertical Axes: Gallons per Capita per Day. Horizontal Axes: 2005 U.S. dollars.



a) As a function of fixedcharges (2005 $ U.S. monthly)



b) As a function of annual personal income (2005 $ U.S.)



c) As a function of climate (hot-dry index score)

Figure 4.2. Quantity as a Univariate Function: Fixed Monthly Charges, Income, and Climate
Vertical Axes: Gallons per Capita per Day

It is important that $f(p)$ is one-to-one and "onto", so that $g = f^{-1}$ produces a unique value for each observation. Figure 4.1 illustrates three different univariate demand functions estimated from the same data, but smoothed to different degrees. The smoothing parameters corresponding to the functions in the figures are s = 0.01, s = 0.05 and s = 0.07. As the smallest value of s such that $\partial q/\partial p \leq 0$ is maintained across the sample, s = 0.07 is chosen as the smoothing parameter of the estimated demand and inverse demand functions. The combined F-statistic for the demand function in 4 regressors over 1680 constant-price observations is 65.65, which indicates a p-score near zero. 45.4% of variation is explained by the model, comparable to an adjusted-R^2 measure of 0.448. Since multivariate nonparametric estimations are difficult to visualize, Figure 4.2 illustrates the partial relationships between the dependent variable and each nonprice regressor. Price alone explains an estimated 7.7% of variations, fixed charges 0.3%, income 1.1%, and weather 24.6%.

All partial p-scores are less than 0.001 except that for fixed charges, for which $p = 0.020$. Given the number of observations and the emphasis of fit over parsimony in GAM, $p > 0.010$ is not a very high statistical bar, and fixed charges fail to reach it nevertheless. From Table 4.1, the mean of fixed charges is only 1.5% of mean income, so it is not surprising that a significant income effect is not observed. Lack of a perception effect may be due to the insensitivity of the model to short-run dynamics. Climate appears to be the dominant driver of demand in this model. From Figure 4.2(c), the positive effect of temperature and absence of precipitation on demand exists across the range of observation. In contrast, the effect of price is generally negative but "wavy" (Figure 4.1), and the effect of income is ambiguous (Figure 4.2(b)).

Step 2: A Cumulative Density Function for Two-Rate Combinations

The inverse demand function is estimated with the same data under the same model and smoothing parameter as the demand function. Predicted values are fitted to each observation in the two-block subsample. These fitted values are entered as $\hat{p}$ in Equation [4.2], where p_1 and p_2 are the two block rates of the observed schedule, to produce a vector of α_b estimates. The estimated α_b of 707 (27.6%) out of 2565 observations is less than zero. For 554 observations (21.6%), the estimated α_b is greater than unity. These extreme values are forced to 0 or 1, respectively.

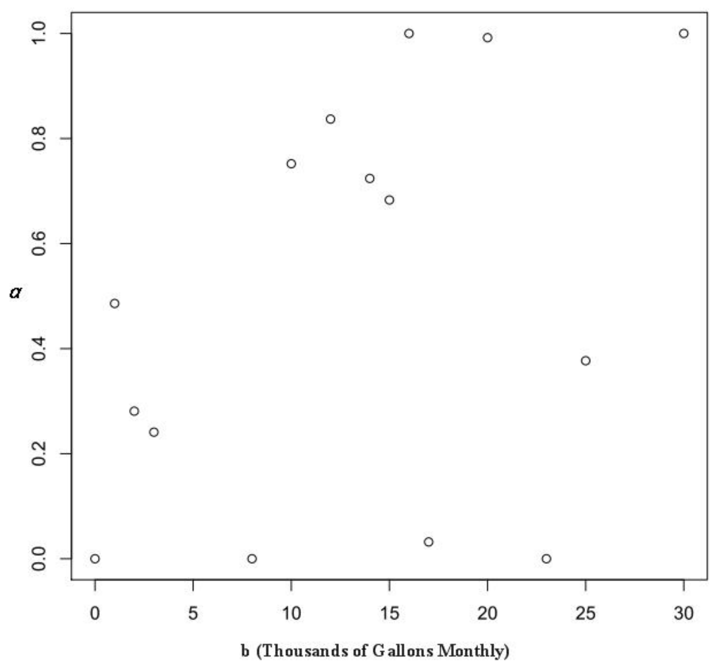
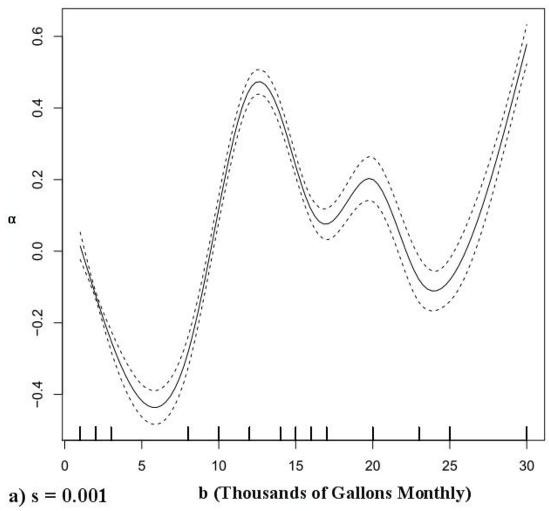
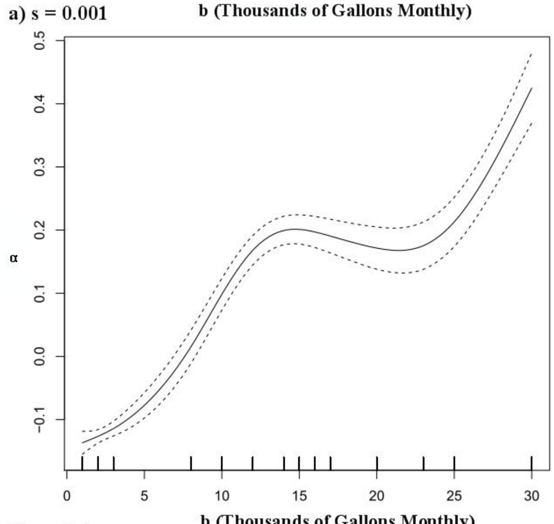


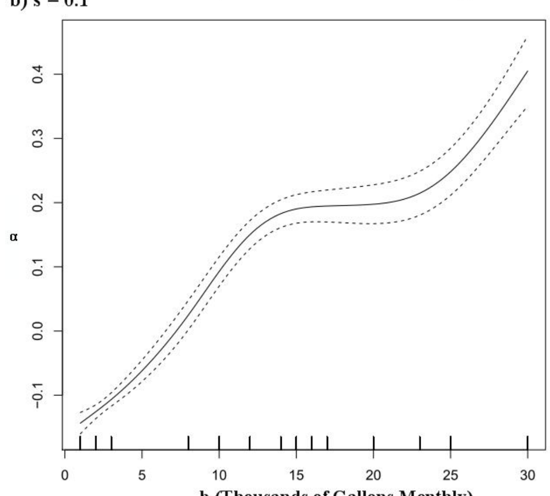
Figure 4.3. Mean Derived Values of α, by Block Boundary b



a) s = 0.001



b) s = 0.1



c) s = 0.18

Figure 4.4. Smoothing from *α(b)* to a Cumulative Density Function

Table 4.2. Estimated Mean and Smoothed Alpha by Block Endpoint

Block Endpoint (Gallons)	Mean Alpha	Standard Deviation	Smoothed Alpha
1,000	0.4860	0.0847	0.2719
2,000	0.2806	0.2722	0.2897
3,000	0.2419	0.1082	0.3087
8,000	0.0000	0.0000	0.4403
10,000	0.7518	0.1592	0.5082
12,000	0.8370	0.2726	0.5659
14,000	0.7237	0.2197	0.5984
15,000	0.6828	0.4414	0.6056
16,000	1.0000	0.0000	0.6089
17,000	0.0316	0.1293	0.6102

The subsample contains 14 different values of b, the break point between the two price blocks. Figure 4.3 plots the means of α_b over these b values, also reported in the second column of Table 4.2. Figure 4.4 shows 3 smooths of these 14 points, at $s = 0.001$, $s = 0.1$, and $s = 0.18$. Figure 4.5 plots fitted values from the $s = 0.18$ model, since it is the least smoothed one-to-one function. The purpose of smoothing is to satisfy Property (6) while doing the least harm to the structure revealed by the raw estimates.

Figure 4.6 is a histogram of densities implied by the model of Figures 4.4 and 4.5, also reported in the third column of Table 4.2. The histogram of this model does not appear to be unimodal. It features two distinguishable masses and a surprising third peak between them at 13,000-14,000 gallons. The histogram depicted in Figure 4.6 does

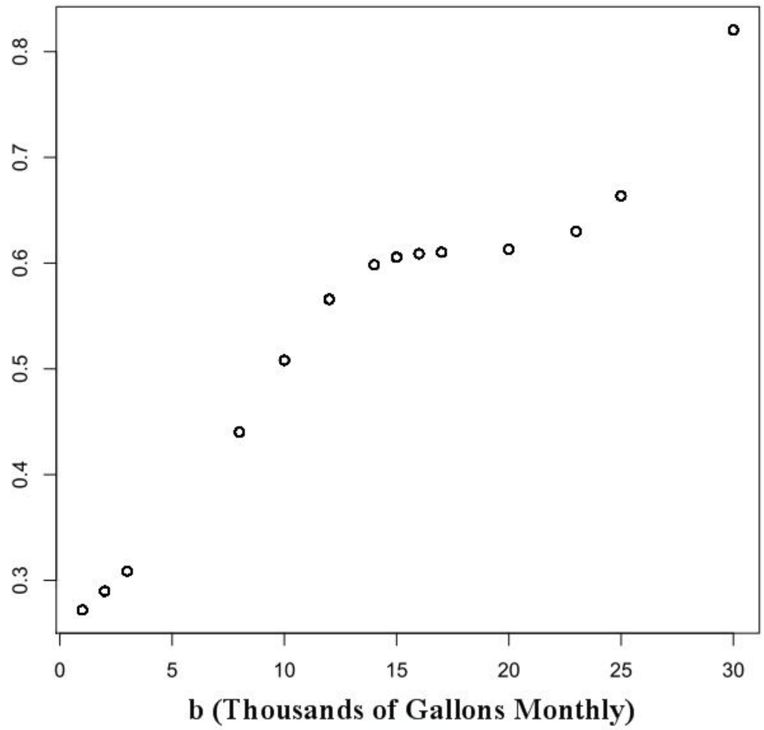


Figure 4.5. Fourteen Points on an Empirical Cumulative Density Function in *b*

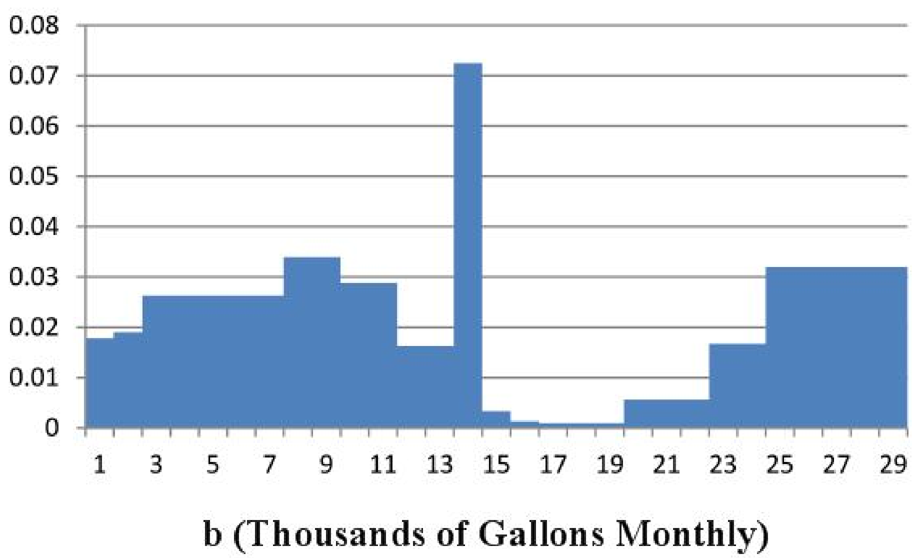


Figure 4.6. Empirically Derived Density Weighting in *b*

not resemble any common distribution by inspection. Whether the empirical distribution approximates either a normal distribution, which would validate parametric ML procedure, or a lognormal distribution, which would validate Bell and Griffin (2011, 2008a), are testable hypotheses. The sudden peak could indicate clustering at kink points (Moffitt 1986), but insufficient data are available to test this hypothesis.

Shapiro-Wilk and Shapiro-Francia tests of normal distribution reject the hypotheses that the derived distribution is normal or lognormal, each with a p-score below 0.01. A linear trend regression, on the other hand, fails to reject the hypothesis that the distribution is uniform about the mean of 0.0213 (for the trend coefficient, $t = -0.50$, $p = 0.62$). Additional distributional hypotheses may be tested similarly, by treating the values of the third column of Table 4.2 as realizations of a random variable. Parametric distributions may also be imposed by fitting the 14 points to the desired functional form.

Although these results are food for thought, the uncensored estimate of α_b is near zero with a standard deviation of 4.036, so distributional advice from even this unusually extensive data is inconclusive. Also, the estimated density of $\alpha_b(30)$ is only 0.821, so the model provides no information on the distribution of the other 0.179 density to blocks over 30,000 gallons per month. Assuming a uniform distribution, the implied support would be 0 to 47,000 gallons.

Step 3: A Demand Function in Weighted Marginal Price

The subsample of observations containing more than two rate blocks is used to test the results of the previous section. A price index for each of these observations is calculated using the derived empirical distribution given in the third column of Table 4.2 and illustrated in Figures 4.4-6. In lieu of better empirical guidance, the full unassigned weight of 0.179 is applied to the rate applicable at 31,000 gallons per month. Another index is calculated as a uniform density from 0 to 47,000 gallons. A third index simply consists of the nonzero rate applicable to the lowest-volume block, i.e. p_1 unless $p_1 = 0$. A fourth index consists of an IV price composed of the predicted values of the convenient (consumption-identified) price regressed on the rates effective at 8,000, 14,000, and 26,000 gallons. The instruments are based on the three peak densities shown in Figure 4.6.

The value of each index is evaluated against the constant-price subsample according to three criteria: 1) Is it distributed similarly to constant price? 2) Does it explain demand as well? 3) Does it provide a similar demand estimate? Addressing these criteria, Table 4.3 presents the moments of each price index, the results of univariate and multivariate GAM demand regressions, and the results of a standard Cobb-Douglas demand regression.

If the distribution of a price index does not resemble the distribution of observed constant price, either the decision to use a multiple-rate schedule is correlated with the magnitude of the monopolist's price signal, or the index does not represent the price signal in some way. The means of all four indices are higher than the mean of constant price (Row 1, Table 4.3), although not significantly so, since the standard deviation of constant price is 2.45 (Row 2). This does not conclusively indicate a relationship

Table 4.3. Comparison of Price Metrics.

Statistic	First Nonzero	Uniform Price	Empirical Price	IV Price	Constant Price
	Distributional				
Mean	5.12	4.38	3.83	4.34	3.09
Standard Dev.	2.30	1.96	1.76	1.73	2.45
Skewness	0.98	0.85	1.02	0.58	1.09
Kurtosis	3.52	3.36	3.59	4.07	3.68
	Multivariate GAM				
F-test	88.34	73.87	75.27	5.58	65.65
Explained %	29.60	26.60	26.40	8.66	45.40
	Univariate GAM				
F-test	37.52	43.62	35.70	15.50	26.11
Explained %	6.08	5.91	5.00	2.04	7.69
	Cobb-Douglas				
Price	-0.2187	-0.3358	-0.4114	-0.1860	-0.2283
	0.0172	*0.0161*	*0.0166*	*0.0160*	*0.0195*
Fixed Charges	-0.2138	-0.2879	-0.3222	-0.2558	-0.0550
	0.0175	*0.0177*	*0.0177*	*0.0183*	*0.0136*
Income	-0.0279	0.0963	0.0714	0.0805	0.0639
	0.0289	*0.0279*	*0.0274*	*0.0295*	*0.0398*
Climate	0.2562	0.2707	0.2825	0.2645	0.5104
	0.0213	*0.0209*	*0.0207*	*0.0215*	*0.0.33*
Constant	5.1297	4.1380	4.4815	4.0318	2.5983
	0.3149	*0.2957*	*0.2927*	*0.3098*	*0.4599*
Adjusted R^2	0.0564	0.0915	0.1135	0.0539	0.2236

between higher price and additional rate blocks. Individually, each of the indices passes the distribution test. Comparisons of skewness and kurtosis (Rows 3 and 4) are meant to be comparative. The empirical distribution best matches the distribution of constant prices in mean, skewness, and kurtosis, but with a lower standard deviation. None of the indices appears to have radically departed from the comparison distribution.

To represent price well, the indices must demonstrate an effect similar to the price effect observed in the constant-price subsample. This requirement includes both goodness of fit and similarity of fit. Similarity is difficult to quantify without parameters, so a Cobb-Douglas regression model is presented in Table 4.3 along with fit measures of the GAM nonparametric regressions from the previous section applied to the new indices. Although F-statistics are presented, the regression in constant price is not comparable to the others because the constant-price subsample is less than half the size of the multiblock subsample (see Table 4.1).

None of the multiprice indices explains as much variation in the tested models as constant price, despite the larger sample size of the former. The indices appear comparable among one another, with the first-nonzero price index somewhat outperforming the empirical and uniform indices. The first-nonzero price also produces the only constant elasticity estimate in the parametric regression that falls within the 95% confidence interval of the constant price estimate. The empirical, uniform, and IV indices produce closer income elasticity estimates, and none of the index models produces an estimate of fixed-charge or climate elasticity that is statistically similar to a comparison estimate. From the visual evidence of Figure 4.2(b), the Cobb-Douglas income variable is likely misspecified. Despite an unimpressive match with the constant-price comparison, the empirical Cobb-Douglas regression does boast the best overall fit of the index regressions, as measured by adjusted R-squared. IV price performs most poorly on most of the criteria. It is possible that the shape of the block schedule has an independent effect on demand (Olmstead et al. 2007). This effect would challenge the similarity criterion and give more credit to the empirically-weighted price index.

Conclusions

The problem of representing price under multiblock rate schedules is recognized, and theoretical progress has been made in defining price when rates are variable. Empirical tests of price representations are difficult to construct. The present research assumes that a linear weighting of block rates can consistently represent an aggregate price signal. A continuous distribution of weights across an arbitrary block schedule is derived by dissecting demand behavior under a two-block schedule and comparing it to demand under constant price. The derived weighted price index is unique among multiblock price representations because it reflects the data rather than just the assumptive base. By construction, the index is independent of quantity demanded, responsive to change in the rate schedules, and applicable to aggregate data.

Despite data spanning 105 communities over 11 years, the results of the research still suffer from a lack of variation in the block variable, b, upon which the procedures heavily rely. Nevertheless, some useful findings are obtained. A series of 14 (α, b) pairs are estimated which can be used to fit parametric distributions, independent of regression functional form (Figure 4.5 and Table 4.2). The uniform distribution is not ruled out, but linear and loglinear distributions are rejected. Furthermore, an empirical basis for a range of price sensitivity is established. Based on the estimated cumulative distribution, 82% of price response occurs at volumes below 30,000 gallons monthly per connection. If community price response is distributed uniformly on the rate schedule, rates effective above 50,000 gallons should be irrelevant to demand.

Researchers may find it redundant to repeat a national survey when a demand function is desired for one community. In this case, they may wish to corroborate the evidence here with their local knowledge of distribution across blocks, or use a uniform distribution (such as from zero to the recommended 47,000 gallons). Even using the first nonzero rate in their schedule is a passable solution, although it is premature to advise discarding rate information.

CHAPTER V

CONCLUSIONS

Urban water demand studies in the period 1967-2002 focused on determining structural own-price elasticity (Arbues 2003). Common research questions in this period included whether an estimated price elasticity agreed with the body of previous studies and whether it was in the interval (-1, 0). Economists were more easily convinced of a price effect on water than were policy makers. One reason may be that substantial price elasticities of -0.3 to -0.6, as found in the literature, do not jibe with practitioners' experience. The present research hints at an explanation: full adjustment to price is too slow to be noticed at the level of a single policy cycle. Better accounting of timing effects can lead to greater acceptance of demand management and more reasonable expectations of its potential.

In the present research, both "price" and "elasticity" are found to be problematic concepts. Many representations of price are reviewed, differing in assumptive base and statistical properties. Many elasticities are found, differing in time horizon and price argument. Results obtained here challenge the traditional simplifications of water demand by applying broad data, emerging econometric techniques, and weighted price indexing. The overall message to researchers is to relax more assumptions and to test the underpinnings of the simple model. The cost of increased complexity is justified by the rewards of increased clarity and relevance to real situations.

We began with the concept of a community as a geographical focus of economic activity and the belief that the community is a meaningful behavioral unit of water demand. Late in 2004, we identified 1400 of the largest providers of water to Texas communities and asked them what their rate schedules had been for the five years preceding, 1999- 2003. The Texas Water Development Board had already requested reports of monthly withdrawals for those years from the same providers. Although 1400 households would have been easier to survey than 1400 communities, estimation results from the community data were strengthened by the diversity of economic and weather conditions and prices observed. Although households in the same community may experience different rates, they all experience the same rate schedule, so any appearance of price variation in such a sample would be illusory and lead to spurious results. Also, community behavior is more directly applicable to regional and national policy dialog than household behavior. Data from the 734 respondent communities form the empirical basis of Chapter II.

Encouraged by our results, we repeated the exercise on a national scale starting in 2006. This time, we pursued a more ambitious time series, the months of 1995 through 2005. Rather than issue a mailed survey, we collected as much rate data as possible from the websites of water utility systems and municipal ordinances before making individual phone and e-mail contacts. This step was facilitated by restricting our universe to only urban communities of population greater than 25,000. There is no national depository of monthly withdrawals, so we contacted the natural resources departments of the states. Fortunately, eight states with communities for which we had rate data also had annual water withdrawal reporting protocols. Communities in these eight states, plus the borough of Anchorage, Alaska, form the empirical basis of Chapters III and IV.

Our surveys revealed a diversity of rate practices among American water providers. Many use schedules of multiple rates that increase and/or decrease with quantity purchased, but some use only one rate, and some do not meter water use at all. Some rate schedules vary with nonvolumetric factors such as elevation, home size, or season. The most prevalent nonvolumetric rate factor is whether the customer is a residence or a business. Among those who offer sewerage service, most estimate sewer use based on water use volume or winter water use volume (winter averaging). Although we regrettably could not capture every aspect of every rate schedule, we could at least represent the volumetric rates themselves.

Rate schedules are like prices in many respects. Most importantly, the area under the rate schedule represents the total variable cost to an individual consumer. Unlike a canonical price, though, a multiple block rate schedule does not uniquely identify an aggregate expenditure because x quantity demanded by a single consumer may not cost the same as x quantity shared among multiple consumers. Therefore the marginal value of water supply depends on an unknown distribution of use at the margin, and a rate-demand function would not be invertible even if it could be cleverly parameterized. The utility of the block rate schedule as an economic tool is limited if it does not correspond to an invertible demand argument.

The power of demand theory would be open to block rate schedules if only they were transitively ordered. Then a family of schedules would correspond to both a unique value of aggregate quantity demanded and a defined scalar price. In other words, any rate schedule would correspond to a unique demand argument. A one-way index from block rates to the real number line would achieve such an ordering. The catch is that contemporaneous quantity demanded cannot be used to identify such an index because quantity is the dependent variable. The goal is exogenous price representation and the solution offered and investigated in this dissertation is weighted price indexing.

Summary

In Chapter II, the shortcomings of prevalent price metrics are discussed. A metric that is weakly representative of experienced price may produce nonsensical or misleading estimates of price effects. A metric that depends on a quantity argument is intrinsically endogenous and susceptible to inconsistent estimation. Even relatively sophisticated instrumental and probabilistic price representations suffer from one or both of these flaws. The proposed quasidifference price is highly representative because it is composed of a weighted combination of observed rates. Endogeneity is eliminated by differencing two prices weighted by the same function.

In the empirical portion of Chapter II, marginal price is found to offer more descriptive power than average price. Weather covariates and quasidifference price are shown to be significant determinants of demand, but mean personal income and availability of central sewerage are not. Price elasticity is shown to vary with respect to all explanatory variables. Price elasticity estimated at the means of each explanatory variable is found to be -0.127. This estimate is lower in absolute value than most static estimates, as is expected from a shorter-run model. An implication for water planning is that the single-year impact of a rate hike will be weaker than predicted by point expansion using a structural elasticity.

The finding of significant variation in price elasticity with respect to other explanatory variables implies that constant elasticity estimates are subject to bias arising from the omission of the other variables. The seriousness of such a bias is unknown but worthy of contemplation since so many constant water price elasticities are estimated in the literature. In addition to bias, constant elasticity is inadequate for precise applications because it does not parameterize relevant considerations such as when, during what season, or for which regions or communities the elasticity estimate is supposed to apply. For instance, it may be relevant to the dialog on climate change that water demand is more elastic at higher mean temperatures.

In Chapter III, dynamic adjustment, seasonality, and a distinct commercial sector are added to the demand model. Adjustment to exogenous shocks is shown to occur slowly, as evinced by the significant EC coefficient of -0.187. This value implies that realizing 90% of a demand adjustment could take 12 years. Seasonality is evident, as structural model coefficients vary significantly by month. Coefficients do not vary across sectors sufficiently to justify the multisectoral approach. Since the dependent variable series is found to exhibit a stochastic seasonal root, consistency of the EC model is initially in question. The finding of residual consistency, however, indicates that the structural model is cointegrated and thus that the results are not only consistent but super-consistent. The rich variation in price elasticity over season and temporal scale testifies to the inadequacy of the concept of "price elasticity" in any but the broadest applications of water demand information. Rather than an elasticity scalar, researchers may be better served to use an elasticity equation, such as Equation 2.17 on Page 29.

The idea of a weighted price index, although used in Chapters II and III, is scrutinized in Chapter IV. Building on the discussion in Chapter II, none of the popular price representations is entirely acceptable as an aggregate price signal. A weighted price index could be a valid representation as long as a few reasonable assumptions are met. These assumptions do seem to be met by the data, although the weighting function derived cannot be considered definitive due to low variation in the locations of block boundaries. Based on the derived price weighting, the most explanatory price index is not distributed normally or lognormally. This is significant because the quasidifference index is weighted lognormally in Chapters II and III. Performance differences among consistent indices are slight, however. The validity of a uniform weighting on the approximate interval 0 to 47,000 gallons monthly is not rejected.

Observations

More than the techniques employed, it is the data that have enabled this research to produce new results that illuminate the mechanisms of aggregate water demand. The techniques demand more variation in rate structure and quantity consumed than a single time series can provide, necessitating a panel approach. Several states collect quantity data from their water providers, but there is no known clearinghouse of rate data. It is imperative that price metrics are built for their ability to independently represent the perceived opportunity cost of consuming water, not for convenient data gathering. Therefore, researchers cannot be satisfied with less than the full rate schedule. The data used in this research are already out of date, and researchers will be forced to repeat the arduous task of data collection until a database of historical rates is established.

Water demand modelers and demand management practitioners alike must be conscious of the temporal aspect of water demand. Aggregate demand is manifestly seasonal, even when weather and climate are taken into account, and aggregate adjustment takes several years. If cross-sectional elasticity is to be interpreted as an adjustment parameter, adjustment time must be assumed to exceed ten years with only a negligible effect in the first year. The role of timing in demand management depends on whether the goal is drought management, peak loading, or facility scaling. In drought management, emergency rates should be sufficiently high to force a timely response. To accommodate or control peak loads, capacity calculations should be based on demand parameters unique to the peak season (usually summer). Timing may be less important to facility designers, although they will want to design capacity based on seasonal loading.

The economist is likely to be more interested in forecasting and welfare measurement than in plant design. The forecasting implications of this research are rather straightforward, with some projections made at the end of Chapter III. The finding of a slow adjustment rate carries a major implication for welfare analysis that may be less obvious. Welfare effects of temporal events, such as drought, should not be derived from calculus on the structural demand function, but on a steeper demand function corresponding to the community's ability to adjust over the horizon of the event. The slope of uncompensated demand in the first year of quantity rationing will be 5-6 times as steep as the long-run uncompensated demand function usually employed. If drought is stochastic, the optimal a priori price level will therefore be higher than the generally accepted static efficient price. In other words, a community may benefit in the long run from exposure to above-equilibrium "conservation" pricing that induces small reductions in demand in wet years, depending on the likelihood and severity of drought.

This dissertation starts from the assumption that water demand can be represented as a function of price, income, and weather variables. Flexible specifications of demand indicate that the relationships among these variables are not simple. Misspecification arising from independence and linearity restrictions is not trivial. On the other hand, data demands become more burdensome as the empirical model gains complexity. The income effect on water demand is ambiguous. The nonparametric income effect estimated in Chapter IV exhibits a U-shape at higher income levels, whereas parametric estimates show an insignificant monotonic income effect. The personal income data used in the analysis reduces the reliability of these particular results, though. Higher moments of aggregate personal income may be needed to clarify the income effect.

Prior, statistically-based approaches to exogenous price representation have not ideally complemented water demand with water's economic idiosyncrasies. Although theoretically elaborate and computationally intense, IV and ML price metrics have not contributed much to model fit or residual independence. A weighted index is an easy-to-apply alternative that short-circuits the feedback between price and quantity consumed.

Because price representation is a persistent problem for water demand estimation, future research to provide either theoretical or empirical support for the shape of a standard weighting function would improve the accessibility of high-level demand analysis to both water management professionals and economists.

REFERENCES

Agthe, D., and R. Billings. 1980. "Dynamic Models of Residential Water Demand." *Water Resources Research*, 16(3): 476-480.

Agthe, D.E., R.B. Billings, J.L. Dobra, and K. Raffiee. 1986. "A Simultaneous Equation Demand Model for Block Rates." *Water Resources Research* 22(1): 1-4.

Arbués, F., M. Á. García-Valiñas, and R. Martínez-Espiñeira. 2003. "Estimation of Residential Water Demand: A State-of-the-Art Review." *Journal of Socio-Economics 32*(1): 81-102.

Balestra, P., and M. Nerlove. 1966. "Pooling Cross Section and Time Series Data in the Estimation of a Dynamic Model: The Demand for Natural Gas." *Econometrica* 34(3): 585-612.

Barber, N. L. 2009. "Summary of Estimated Water Use in the United States in 2005." U. S. Geological Survey Fact Sheet 2009-3098: 1-2. pubs.usgs.gov/fs/2009/3098/2009-3098.pdf.

Basani, M., J. Isham, and B. Reilly. 2008. "The Determinants of Water Connection and Water Consumption: Empirical Evidence from a Cambodian Household Survey." *World Development* 36(5): 953-968.

Bell, D. R., and R. C. Griffin. 2008a. "An Annual Quasidifference Approach to Water Price Elasticity." *Water Resources Research* 44: W08420, doi:10.1029/2007WR006233.

Bell, D. R., and R. C. Griffin. 2008b. *An Economic Investigation of Urban Water*

Demand in The U. S. Texas Water Resources Institute Technical Report TR-331. College Station, TX: Texas A&M University. http://ron-griffin.tamu.edu/reprints/US_UrbanDemand2008.pdf.

Bell, D.R., and R.C. Griffin. 2011. "Urban Water Demand with Periodic Error Correction." *Land Economics* 87(3): 528-544.

Blomquist, S., and W.K. Newey. 2002. "Nonparametric Estimation with Nonlinear Budget Sets." *Econometrica* 70(6): 2455-2480.

Boiteux, M. 1971. "On the Management of Public Monopolies Subject to Budgetary Constraints." *Journal of Economic Theory* 3(3): 216-240.

Bond, L. 2009. "Panel to Confront Sparta Crisis." (Newspaper article) *The Ruston Daily Leader*. July 26.

Boswijk, H. P., and P. H. Franses. 1995a. "Periodic Cointegration: Representation and Inference." *The Review of Economics and Statistics* 77(3): 436-454.

Boswijk, H. P., and P. H. Franses. 1995b. "Testing for Periodic Integration." *Economics Letters* 48: 241-248.

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

Bushnell, J. B., and E. T. Mansur. 2005. "Consumption under Noisy Price Signals: A Study of Electricity Retail Rate Deregulation in San Diego." *Journal of Industrial Economics* 53(4): 493-514.

Carter, D. W., and J. W. Milon. 2005. "Price Knowledge in Household Demand for

Utility Services." *Land Economics* 81(2): 265-283.

Carver, P. H., and J. J. Boland. 1980. "Short-and Long-Run Effects of Price on Municipal Water Use." *Water Resources Research* 16(4): 609-616.

Castro, T. del B., and D. R. Osborn. 2008. "Cointegration for Periodically Integrated Processes." *Econometric Theory* 24(1): 109-142.

Cavanagh, S. M., W. M. Hanemann, and R. N. Stavin. 2002. "Muffled Price Signals: Household Water Demand Under Increasing-Block Prices." (Conference proceedings) *Fondazione Eni Enrico Mattei*, Trieste, Italy: 1-45.

Chicoine, D.L., and G. Ramamurthy. 1986. "Evidence on the Specification of Price in the Study of Domestic Water Demand." *Land Economics* 62(1): 26-32.

Christensen, L. R., D. W. Jorgenson, and L. J. Lau. 1973. "Transcendental Logarithmic Production Frontiers." *The Review of Economics and Statistics* 55(1): 28-45.

Coase, R.H. 1946. "The Marginal Cost Controversy." *Economica* 13(51): 169-182.

Corral, L., A.C. Fisher, and N.W. Hatch. 1998. "Price and Non-Price Influences on Water Conservation: An Econometric Model of Aggregate Demand under Nonlinear Budget Constraint." Working Paper No. 881. Department of Agricultural and Resource Economics and Policy, University of California at Berkeley.

Cregan, A. 2009. "An 'Inconvenient Truth'." (Newspaper article) *The Intelligencer*. August 03.

Dalhuisen, J. M., R. Florax, H. L. F. de Groot, and P. Nijkamp. 2003. "Price and Income Elasticities of Residential Water Demand: A Meta Analysis." *Land Economics*

79(2): 292-308.

Davis, G.C. 2004. "The Structure of Models: An Application to Theory Reduction and Testing with a Production Example." *Journal of Agricultural and Resource Economics* 29(1): 65-78.

Deller, S.C., D.L.Chicoine, and G. Ramamurthy. 1986. "Instrumental Variables Approach to Rural Water Service Demand." *Southern Economic Journal* 53(2): 333-346.

Diakite, D., A. Semenov, and A. Thomas. 2009. "A Proposal for Social Pricing of Water Supply in Côte d'Iviore." *Journal of Development Economics* 88(2): 258-268.

Dziegielewski, B. 1999. "Management of Water Demand: Unresolved Issues." *Water Resources Update* 114: 1-7.

Edwards, G. 2006. "Whose Values Count? Demand Management for Melbourne's Water." *The Economic Record* 82(S1): S54-S63.

Engle, R. F., C. W. J. Granger, and J. J. Hallman. 1989. "Merging Short- and Long-Run Forecasts." *Journal of Econometrics* 40(1): 45-62.

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

Evans, B. 2009. "Water Dispute Strains States' Governors' Relations." *Miami Herald*. August 1.

Foster, H. S., and B. R. Beattie. 1981a. "Urban Residential Demand for Water in the United States: Reply." *Land Economics* 57(2): 257-265.

Foster, H. S., and B. R. Beattie. 1981b. "On the Specification of Price in Studies of

Consumer Demand under Block Price Scheduling." *Land Economics* 57(4): 624-629.

Fullerton, T. M. Jr., and A. Elias. 2004. "Short-Term Water Consumption Dynamics in El Paso, Texas." *Water Resources Research* 40: W08201, doi:10.1029/2004WR003260.

Gabor, A. 1955. "A Note on Block Tariffs." *The Review of Economic Studies* 23(1): 32-41.

Gaudin, S. 2006. "Effect of Price Information on Residential Water Demand." *Applied Economics* 38(4): 383-393.

Gaudin, S., R. C. Griffin, and R. C. Sickles. 2001. "Demand Specification for Municipal Water Management: Evaluation of the Stone-Geary Form." *Land Economics* 77(3): 399-422.

Grafton, R. Q. and M. B. Ward. 2008. "Prices versus Rationing: Marshallian Surplus and Mandatory Water Restrictions." *Economic Record* 84(S1): S57-S65.

Griffin, R. C. 2001. "Effective Water Pricing." *Journal of the American Water Resources Association* 37(5): 1335-1347.

Griffin, R. C. 2006. *Water Resource Economics: The Analysis of Scarcity, Policies, and Projects*. Cambridge, MA: The MIT Press.

Griffin, R. C., and C. Chang. 1991. "Seasonality in Community Water Demand." *Western Journal of Agricultural Economics* 16(2): 207-217.

Griffin, R. C. and J. W. Mjelde. 2000. "Valuing Water Supply Reliability." *American Journal of Agricultural Economics* 82(2): 414-426.

Hanemann, W. M. 1984. "Discrete/Continuous Models of Consumer Demand." *Econometrica*, 52(3): 541-561.

Hastie, T. and R. Tibshirani. 1990. *Generalized Additive Models*. Boca Raton FL: Chapman & Hall.

Hausman, J.A. 1985. "The Econometrics of Nonlinear Budget Sets." *Econometrica* 53(6): 1255-1282.

Hausman, J. A., M. Kinnucan, and D. McFadden. 1979. "A Two-Level Electricity Demand Model." *Journal of Econometrics* 10(3): 263-289.

Heckman, J.J. 1979. "Sample Selection Bias As a Specification Error." *Econometrica* 47(1): 153-161.

Herriges, J. A., and K. K. King. 1994. "Residential Demand for Electricity under Inverted Block Rates: Evidence from a Controlled Experiment" *Journal of Business and Economic Statistics* 12(4): 419-430.

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

Hewitt, J. A. 2000b. "An Investigation into the Reasons Why Water Utilities Choose Particular Residential Rate Structures." In *The Political Economy of Water Pricing Reforms*, ed. A. Dinar. Oxford University Press, New York, NY: 259-277.

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

Hotelling, H. 1938. "The General Welfare in Relation to Problems of Taxation and of

Railway and Utility Rates." *Econometrica* 6(3): 242-269.

Houthakker, H.S. 1951. "Some Calculations on Electricity Consumption in Great Britain." *Journal of the Royal Statistical Society. Series A* 114(3): 359-371.

Howe, C. W., and F. P. Linaweaver, Jr. 1967. "The Impact of Price on Residential Water Demand and Its Relation to System Design and Price Structure." *Water Resources Research* 3(1): 13-32.

Juselius, K. 2006. *The Cointegrated VAR Model*. Advanced Texts in Econometrics. New York: Oxford University Press.

Kennedy, P. 2003. *A Guide to Econometrics, 5th ed*. The MIT Press, Cambridge, MA.

Kostas, B., and S. Chrysostomos. 2006. "Estimating Urban Residential Water Demand Determinants and Forecasting Water Demand for Athens Metropolitan Area 2000-2010." *South-Eastern Europe Journal of Economics 1: 47-59.*

Kunst, R. M. 1993. "Seasonal Cointegration, Common Seasonals, and Forecasting Seasonal Series." *Empirical Economics* 18(4): 761-776.

Lach, D., H. Ingram, and S. Rayner. 2005. "Maintaining the Status Quo: How Institutional Norms and Practices Create Conservative Water Organizations." *Texas Law Review* 83(2027): 2027-53.

Lyman, R. A. 1992. "Peak and Off-Peak Residential Water Demand." *Water Resources Research* 28(9): 2159-2167.

Martínez-Espiñeira, R. 2003. "Estimating Water Demand under Increasing-Block Tariffs Using Aggregate Data and Proportions of Users per Block." *Environmental and Resource Economics* 26(1): 5-23.

Martinez-Espineira, R. 2007. "An Estimation of Residential Water Demand Using Co-Integration and Error Correction Techniques." *Journal of Applied Economics* 10(1): 161-184.

Martinez-Espineira, R., and C. Nauges. 2004. "Is All Domestic Water Consumption Sensitive to Price Control?" *Applied Economics* 36(15): 1697-1703.

McKenzie, L. W. 2002. *Classical General Equilibrium Theory*. Cambridge, MA: MIT Press.

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

Musolesi, A., and M. Nosvelli. 2007. "Dynamics of Residential Water Consumption in a Panel of Italian Municipalities." *Applied Economic Letters* 14(6): 441-444.

Nauges, C., and R. Blundell. 2001. "Estimating Residential Water Demand under Block Rate Pricing: A Nonparametric Approach." www.inra.fr/esr/vie/animations/ jeuneschercheurs/pdf/2001/nauges.pdf.

Nauges, C., and A. Thomas. 2003. "Long Run Study of Residential Water Consumption." *Environmental and Resource Economics* 26(1): 25-43.

Newey, W.K., and Powell. 2003. "Instrumental Variable Estimation of Nonparametric Models." *Econometrica* 71(5): 1565-1578.

Nieswiadomy, M. L. 1991. "A Note on Price Perception in Water Demand Models." *Land Economics*, 67(3): 352-359.

Nordin, J. A. 1976. "A proposed modification of Taylor's demand analysis: comment."

The Bell Journal of Economics 7(2): 719-721.

Olmstead, S. M., W. M. Hanemann, and R. N. Stavins. 2007. "Water Demand under Alternative Price Structures." *Journal of Environmental Economics and Management* 54(2): 181-198.

Reiss, P. C., and M. W. White. 2005. "Household Electricity Demand Revisited." *Review of Economic Studies* 72(3): 853-883.

Renwick, M. E., and R. D. Green. 2000. "Do Residential Water Demand Side Management Policies Measure Up? An Analysis of Eight California Water Agencies." *Journal of Environmental Economics and Management* 40(1): 37-55.

Renzetti, S. 1992. "Evaluating the Welfare Effects of Reforming Municipal Water Prices." *Journal of Environmental Economics and Management* 22(2): 147-163.

Renzetti, S. 2002. *The Economics of Water Demands*. Kluwer Academic Publishers, Boston, MA.

Reynaud, A., S. Renzetti, and M. Villeneuve. 2005. "Residential Water Demand with Endogenous Pricing: The Canadian Case." *Water Resources Research 41*: W11409, doi: 11410.11029/12005WR004195.

Ruijs, A., A. Zimmermann, and M. van den Berg. 2008. "Demand and Distributional Effects of Water Pricing Policies." *Ecological Economics* 66(2-3): 506-516.

Russ, K. 2009. "Water Crisis: In-Fighting Stalls Solutions." (Newspaper article) *CityWatch*. August 4.

Schefter, J. E., and E. L. David. 1985. "Estimating Residential Water Demand Under Multi-Part Tarrifs Using Aggregate Data." *Land Economics* 6(3): 272-280.

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

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

Taylor, R. G., J. R. McKean, and R. A. Young. 2004. "Alternate Price Specifications for Estimating Residential Water Demand with Fixed Fees." *Land Economics* 80(3): 463-475.

Terza, J. V. 1986. "Determinants of household Electricity Demand: A Two-Stage Probit Approach." *Southern Economic Journal* 52(4): 1131-1140.

Varian, H. R. 1982. "The Nonparametric Approach to Demand Analysis." *Econometrica* 50(4): 945-973.

Ward, F. A., J. F. Booker, and A. M. Michelsen. 2006. "Integrated Economic, Hydrologic, and Institutional Analysis of Policy Responses to Mitigate Drought Impacts in Rio Grande Basin." *Journal of Water Resources Planning and Management* 132(6): 488-502.

Working, E. J. 1927. "What Do Statistical 'Demand Curves' Show?" *Quarterly Journal of Economics* 41(2): 212-235.

Wu, X. 2003. "Nonparametric Labor Supply Estimation under the Earned Income Tax Credit and Welfare Programs." socserv.socsci.mcmaster.ca/cesg2003/ ximingpaper2.pdf.

Young, M. and J. McColl. 2005. "Defining Tradable Water Entitlements and

Allocations: A Robust System." *Canadian Water Resources Journal* 30(1): 76-72.

VITA

Name: David Ray Bell

Address: 1407B Cometa St.

Austin, TX 78721

Email Address: belldr@tamu.edu

Education: Ph.D., Agricultural Economics, Texas A&M University, 2011.

B.A., Liberal Arts, The University of Texas at Austin, 1994

Economics major added, 2003

Published Works:

Bell, D., and R.C. Griffin. 2011. "Urban Water Demand with Periodic Error Correction." *Land Economics* 87(3): 528-544.

Baker, J., W.D. Shaw, D. Bell, S. Brody, M. Riddel, R.T. Woodward, and W. Neilson. 2009. "Explaining Subjective Risks of Hurricanes and the Role of Risks in Intended Moving and Location Choice Models." *Natural Hazards Review* 10(3): 102-121.

Bell, D., and R.C. Griffin. 2008. "An Econometric Investigation of Urban Water Demand in the U.S.," Texas Water Resources Institute, Texas A&M University.

Bell, D., and R.C. Griffin. 2008. "An Annual Quasidifference Approach to Water Price Elasticity." *Water Resources Research*, doi:10.1029/2007/WR006233.

Bell, D., and R.C. Griffin. 2006. "Community Water Demand in Texas as a Century is Turned." Report to the Texas Water Development Board.