10-1-1980

An Analysis of Residential Water Demand Schedules in Arkansas

Richard K. Ford
University of Arkansas, Fayetteville

Joseph A. Zieger
Unversity of Arkansas, Fayetteville

AN ANALYSIS OF RESIDENTIAL WATER DEMAND SCHEDULES IN ARKANSAS

by
Richard K. Ford
and
Joseph A. Zieger
Department of Economics

AWRRC

Arkansas Water Resources Research Center

University of Arkansas
Fayetteville

Publication No. 76

1980

AN ANALYSIS OF RESIDENTIAL WATER DEMAND SCHEDULES IN ARKANSAS

by

Richard K. Ford
Eastern Kentucky University

Joseph A. Ziegler
University of Arkansas

University of Arkansas Water Resources Research Center
Fayetteville, Arkansas

October, 1980

TABLE OF CONTENTS

PAGE NO.

LIST OF EXHIBITS

CHAPTER 1

INTRODUCTION

Nature of the Problem

Water has held man's interest for centuries, and a considerable body of knowledge has been developed about this subject, but it has only been considered only recently by decision makers in an economic sense. Because water is required to sustain life, the thought of its being an economic good to individuals has not been considered applicable or "fair" by those individuals who are concerned with its acquisition and distribution. Moreover, water has been in fairly abundant supply relative to demand for most of the country. However, as this situation is upset by increases in population and water use, a realization of the situation must be confronted. As Wollman states in his 1960 publication, ". . . while the country is in no imminent danger of running out of water, no longer is it in the enviable position of having enough water at all times for conceivable uses."[1] Given the fact that water is an economic good, some thought must be given to its production and distribution. However, as with other economic goods and services, the amount of production will depend upon the demand for that good, among other things.

---

[1] Wollman, Nathaniel, "A Preliminary Report on the Supply of and Demand for Water in the United States as Estimates for 1980 and 2000," Water Resources Activities in the United States, Committee Print No. 7 (Washington: Government Printing Office, 1960), p. 1.

Objectives of the Study

The basic objective of this study is to determine and compare the residential demand for water in each of the five Arkansas Water Resource Planning Areas (AWRPA) as shown in Exhibit 1-1. These areas are selected in order to coincide with other studies currently being investigated by other researchers concerned with water for the state of Arkansas.[2] Although the individual areas have formal names, they will be referred to as simply watershed regions, and differentiated by number (as defined in the list below) in order to facilitate the readability of this study.

ARKANSAS WATER RESOURCE PLANNING AREAS

I. Mississippi-St. Francis and Crittenden County

II. Ouachita and Mississippi-Tensas

III. White

IV. Lower Arkansas and Benton County

V. Lower Red

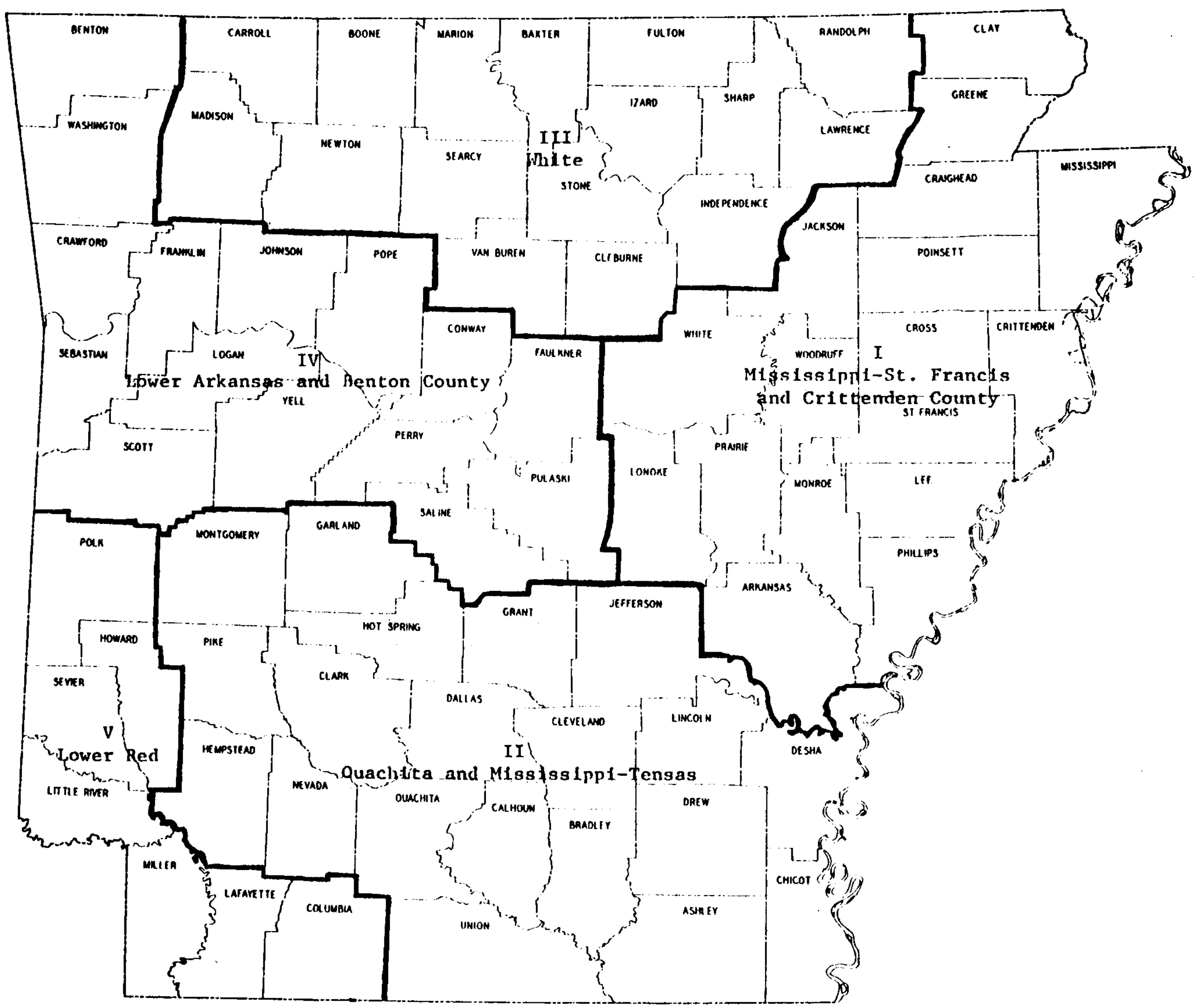
In order to accomplish the main objective of this study, individual models for each of the five Arkansas watershed regions (AWRPA) will be derived theoretically and verified empirically so that comparative static analysis can be employed on the water demand elasticities. The water demand models will be derived initially from a composite of the significant variables found in similar water research projects as well as other theoretically probable factors. The empirical verification of the theoretical models will require the sampling of a large portion of the population of each region. From the resulting data base, regression analysis will be employed to estimate water demand functions for each region and the state

---

[2] Shulstad, Robert N., Joseph A. Ziegler, and Eddie D. Cross, "Projected Water Requirements and Surface Water Availability for Arkansas," University of Arkansas, Special Report 61 (April 1978), p. 3.

EXHIBIT 1-1

Arkansas Water Resource
Planning Areas
(AWRPA)



as a whole. Each of the regional models will be compared with the state model to determine if there are significant differences.

To enhance the usefulness of this study, explanatory models rather than predictive models will be used for comparable analysis of the regions. Although the predictive models can be useful to decision makers who are trying to predict future conditions, they can serve only a limited role since their use requires implicit assumptions about the relationship of the individual factors. In particular, the use of predictive models for forecasting purposes requires the assumption that the price of water will remain proportionally constant through the projection period. This assumption can severely limit the forecasting ability of the predictive models, especially if the reason for forecasting is to inquire into efficient resource allocation. To avoid the limiting assumptions of "all things remaining equal" while projecting future resource allocation, explanatory models will be developed and employed for the comparison of the regions. Not only are the explanatory models more useful in forecasting under various assumptions, they also allow the individual user to study the relationships of the factors to the question of the quantity of water consumed.

Usefulness of the Study

Individuals concerned with the subject of residential water demand can benefit from the research efforts of this study. Both decision makers and researchers concerned with local, state or national residential water demand can find useful and meaningful information within this study to aid them in their endeavors. Although few individuals will find the answer in final form that they seek to particular water demand questions within this volume, most can benefit greatly from the empirical results achieved by this work.

On the local level, this study can serve to aid future urban water system planners within the state of Arkansas. Currently, many misconceptions seem to abound within the field of water system design. Water designers frequently don't acknowledge that the price charged residential users of water will affect its usage. A statement contained within a textbook on sanitary engineering reflects this sentiment: "Rates are obtained by dividing the system costs by the volume of water delivered . . ."[3] When rates are derived in this manner, the designer must use some means other than water price to determine the size of the system that is to be built. However, the size will determine the size of the system that is to be built. However, the size will determine the costs involved, and therefore the rates or price to be charged according to the textbook just mentioned. If the price of water is not a factor in determining the amount of water used by the households, then the households could adjust their usage to the point that the existing water system is not the optimal size. This could result in additional costs to the city in having to expand its water system, when the optimal size could have been built in the first place.

On the state level this study can contribute in the areas of resource planning and welfare policies. As population increases and water becomes increasingly scarce in some parts of the state the need to develop effective and efficient water resource policies will become more apparent. In particular, policies which affect the usage and distribution of current water supplies will have to be developed. In addition, scarce capital improvement funds will have to be allocated on a set of cost-effective priorities which reflect both efficiency and equity. As will be explained in a later section of this study, the water demand models can be used to estimate the effects of various pricing policies on water use and total expen-

---

[3]Fair, Gordon M., John C. Geyer, and Daniel A. Okun, Water and Wastewater Engineering (New York: John Wiley and Sons, 1966), Vol. 1, p. 14.

ditures. They can also be used to help determine the efficient allocation of capital improvement funds. However, the authors do not intend to suggest that this one study would provide sufficient grounds to assign water development fund priorities. Other economic and noneconomic factors must also be considered e.g., water supply schedules, population densities, political situation, and equity.

In addition to assisting the state resource planners in developing policies to allocate water resources, this study can serve also as a valuable aid in predicting future trends and levels of water usage and water requirements for the state. The forecasting of future water use is not new. Most studies use a "water requirements approach" which usually combines projected economic data with estimated water use coefficients. This approach does not consider explicitly the effect of price on the use of water nor does it usually consider the effect of various economic policies. It generally assumes that the trends which influenced water use in the past are stable and can be extended into the future. However, the quantities of water used are highly dependent on the prevailing objectives of society and upon the methods and purposes of water use (e.g., whether for industry or irrigation), upon economic policy (whether fully priced or subsidized), and other variables that can be influenced by public policy. It seems reasonable to question the usefulness of water requirements projections *per se* since these other factors are either so dominant or uncertain in the long run as to be undependable for decision. Because of the uncertainities surrounding future water policies the water requirements approach to forecasting future water use is not likely to be very reliable. Moreover, when incorporated into water planning programs it tends to result in self-fulfilling prophesies, i.e., the projected requirements are often used to generate support for specific water projects.

This study places major emphasis on an alternative approach which analyzes the influences of the factors that control or influence the use of water. This approach deals explicitly with the determinants of water use including price and public policy. Forecasts of water demand must take into account the different values of water used for various purposes. Water does not have an unlimited economic value. Moreover, differences exist in the incremental values of water used for different purposes and water users tend to decrease their water use with increased prices for water. Reallocation of water resources can be expected to occur as the effective price of water increases. Consequently, future forecasts of water use should consider demand schedules which reflect the incremental value of water in alternative uses. These demand schedules should consider both quantity and quality characteristics since these are inseparable.

From a macro or national level, the results of this study can also contribute to the general level of knowledge of water demand developed over the past twenty years. Since the mid-fifties, researchers have attempted to verify empirically various models to explain water demand. While they have improved successively both the models and methodology their results have often differed. These differences can be traced, in large part, to the use of different techniques and sample populations. The incompatibility of data bases, moreover, prevents a meaningful interpretation of these differences. Whether, and to what extent, the previously observed differences are the result of different techniques or data bases is a question which can be addressed in this study, i.e., a single methodology will be applied to a common data base. The results of this study can add meaning to what otherwise is a set of disjointed research studies dealing with a similar topic.

CHAPTER 2

REVIEW OF RELATED LITERATURE

The purpose of this chapter is to review studies related to the estimation of the residential demand for water. The review will proceed chronologically and concentrate on those studies which made a significant contribution. Consequently, it is not comprehensive, nor is it intended to be. A summary of each of the studies will be followed by a summary discussion and observations on the state of the art.

The investigation of the economic relationships that exist between water consumed and water price and income levels is a relatively recent endeavor. Before the 1950's, almost no work had been reported in the professional journals on the subject of domestic water demand. In 1957, however, Harris F. Seidel and Robert E. Bauman published the results of their work which appears to be the first major reported attempt to correlate water usage with the price of water.[1] In this study, the researchers use aggregated water production data and construct a simple linear regression model to estimate the effect that the price of water has upon the amount of water consumed. Although this study remains noteworthy as a major early contribution to this field, it contains several significant shortcomings which include the use of aggregated data, the use of production rather than distribution data, and the use of average price as the only explanatory variable. Moreover, the estimate of average price is made from a sample which in-

---

[1] Harris F. Seidel and Robert E. Bauman, "A Statistical Analysis of Water Works Data for 1955," Journal of American Water Works Association 49 (December 1957): 1531-1566.

cludes a large number of cities which charge flat rates. This estimate tends to bias the results since flat rates remove the motive for an individual customer to vary water use with price, the very relationship that the study tried to determine.

In 1963 Manuel Gottlieb published the results of his study which concentrated on an attempt to determine demand elasticities for both income and prices with respect to water usage.[2] Using two separate time periods (1952 and 1957), he produced a comparative static aggregated multiple regression model which containes both income and price observations as independent variables and the amount of water used as the dependent variable. The water use variable, however, includes total water production and, consequently, overstates residential water use.

His 1952 regression results reveal price and income elasticities of -1.23 and .45 respectively with an R^2 of .67, while his 1957 model shows price and income elasticities of -.65 and .58 respectively with an R^2 of .69. His estimating equations take the form:

$$\log Y = \log a + b \log X + c \log Z$$

where: Y = consumption (in millions of gallons per year)

X = average household income (in dollars)

Z = average price (in cents per millions of gallons)

In November of 1963 J.C. Headley published the results of his inquiries into the relationship between income and water demand.[3] Although he acknowledges that many variables can effect the quantity of water demanded by a household, Headley selects a data set which is nearly homogene-

---

[2]Manuel Gottlieb, "Urban Domestic Demand for Water: A Kansas Case Study," Land Economics 39 (May 1963): 204-210.

[3]J.C. Headley, "The Relation of Family Income and Use of Water for Residential and Commercial Purposes in the San Francisco-Oakland Metropolitan Area," Land Economics 39 (November 1963): 441-449.

ous with respect to all variables except income. This is accomplished by selecting a sample of fourteen communities within the San Francisco Bay area, most of which have identical rate structures and climatological conditions. His raw data is obtained from the U.S. Census and the individual water companies of the communities. He tests his model by using data for two different years; this enables him to cross-check the results of the cross sectional models as well as to formulate a simple time series model. He demonstrates a definite positive relationship between income and the quantity of water consumed by the households, as well as an elastic relationship with respect to income for both 1950 and 1959 (1.49 and 1.29 respectively). However, his time series model does not generate similar results for the income elasticity of demand which ranged from .004 to .406 for the 14 communities over the nine year period. He concludes that the time derived elasticity would be more useful in predicting future demand, while the cross sectional derived coefficient would be more useful for an individual water company contemplating a rate change.

A third noteworthy article appears in the professional literature in 1963, i.e., the work of J.W. Milliman appears to be the first study by an economist dealing with the theoretical aspects of urban water demand and supply.[4] Although he concentrates his arguments on the problems of supply, he observes that the pricing systems as typical in the industry seemed to have been counter productive with respect to resource allocation. The declining block rate pricing systems, the most common form of pricing metered water, "are a vestige from the days when water supply systems had a great deal of excess capacity."[5] The author observes that the rates had

---

[4] J.W. Milliman, "Policy Horizons for Future Urban Water Supply," Land Economics 39 (May 1963): 109-132.

[5] Ibid., p. 119.

become institutionalized and had not been changed to adjust for the alterations in resource availabilities. In particular, water pricing was based on the policy of recovering the original cost of the water system. Milliman advocates the use of marginal cost pricing of the water to urban users, and centers his reasoning around traditional supply and demand analysis, thereby recognizing the importance of determining the urban water demand function and its elasticities.

Charles W. Howe and F.P. Linaweaver use cross sectional data from a sample of thirty-nine areas scattered across the contiguous United States; the sample is selected such that there is substantial homogeneity in the physical and economic characteristics of the dwelling units.[6] The data set contains information on inside and outside water usage (estimated from seasonal differences), identifies metered and flat rate pricing systems, septic tanks and sewerage systems, and identifies apartments from single family dwellings. Using multiple regression analysis they estimate separate demand functions for domestic or internal water use and external or sprinkling water use. The generalized structure of the domestic demand function is:

$$Q_{a,d} = f\ (v,\ a,\ d_p,\ k,\ p_w,\ n_p,\ (rpi))$$

where:

$Q_{a,d}$ = average annual quantity demanded for domestic purposes in gallons per dwelling unit per day

v = market value of the dwelling unit in thousands of dollars

d_p = number of persons per dwelling unit

---

[6]Charles W. Howe and F.P. Linaweaver, "The Impact of Price on Residential Water Demand and its Relation to System Design and Price Structure," Water Resources Research 3 (1st Qtr. 1967): 13-32.

a = age of the dwelling unit in years

k = average water pressure in psi

p_w = the sum of water and sewer charges that vary with water use, evaluated at the block rate applicable to the average domestic use in each study area

n_p = number of billing periods per year

(rpi) = regional price index

From this generalized model, three separate models are developed for subgroups within the sample set. The best fitting reported log-linear equations are:

I Metered dwelling units with public sewer systems;

$$Q_{a,d} = 206 + 3.47v - 1.30p_w \qquad (R^2 = 0.72)$$

II Flat rate pricing to apartments using public sewer systems;

$$Q_{a,d} = 28.9 + 2.39v + 33.6d_p \qquad (R^2 = 0.89)$$

III Metered price system connected to septic tanks;

$$Q_{a,d} = 30.2 + 39.5d_p \qquad (R^2 = 0.96)$$

For the external or sprinkling water demand, factors that affect the lawn are considered, e.g., the amount of rainfall and the evapotranspiration rate. The generalized function is:

$$Q_{s,s} = B_o b^{B_1} (w_s - 0.6r_s)^{B_2} p_s^{B_3} v^{B_4} u$$

where:

$Q_{s,s}$ = average sprinkling demand

b = irrigable area per dwelling unit

w_s = summer potential evapotranspiration

r_s = summer precipitation

$p_s n$ = marginal commodity charge applicable to average summer total rates of use

$b^{B_1}(w_s - 0.6r_s)$ = a physical requirement of the lawn determined from agricultural literature.

The best fitting models for the subgroups within this generalized model form after transferring to natural logs are:

I Metered dwelling units with public sewer systems;

$$Q_{s,s} = 1.09 + 207(w_s - 6r_s) - 1.12p_s + .66v \qquad (R^2 = .73)$$

II Flat rate pricing system and using public sewer systems;

$$Q_{s,s} = 2.0 + 0.783v \qquad (R^2 = .64)$$

The results of this study point to two major findings: (1) domestic demands are relatively inelastic with respect to price and (2) sprinkling demands are elastic with respect to price but less so in the west than in the east.[7] Although Howe and Linaweaver recognize the need to separate the price of water into its marginal components to estimate sprinkling demand, they use average prices for their input estimates for the individual households.

Another water demand study completed in 1967 was conducted by James E. Ware and Ronald M. North whose major objective was to identify the factors which affect demand for water by residential users.[8] Their initial data set is collected from 634 households selected from 14 different communities in northern Georgia, such that there is a wide dispersion of average water prices but a relative homogeneous set of climatological conditions. It is analyzed through the use of stepwise computer routines applied to various multiple regression models. The initial model examined is:

$$Q_d = f\ (X_1, X_2, X_3 . . . X_9)$$

---

[7]Ibid., p. 21.

[8]James E. Ware and Ronald M. North, "Price and Consumption of Water for Residential Use in Georgia," The Atlanta Economic Review (October 1968: 9-13.

where:

X_1 = Number of persons per household

X_2 = Number of bathrooms per household

X_3 = Use of dishwasher

X_4 = Use of clotheswasher

X_5 = Ownership of automatic lawn sprinkler or swimming pool

X_6 = Irrigable lawn area in thousands of square feet

X_7 = Market value of residence, in thousands of dollars

X_8 = Household income in thousands of dollars per year

X_9 = Average price for water and sewerage service in dollars per thousand gallons per month

Both linear and exponential regression models are examined but the exponential form produces the best results. When the model is formulated to include the entire set of independent variables it explains 90 percent of the variance but only the price of water is significant at the .05 level. An exponential model using only the price of water and the income of the households explains 86 percent of the variance and estimates price and income elasticities of -.61 and +.36 respectively. The two main conclusions the authors report from their study are: (1) "consumers are responsive to price differences for water and . . . (2) consumers will increase residential water use when incomes increase."[9]

Steve H. Hanke took advantage of a rare opportunity to study the demand for water within a community under dynamic price conditions.[10] Between the years 1961 and 1963, the water company of Boulder, Colorado converted from a flat rate to a marginal rate pricing system. The water company installed meters on the residential lines leading into each household,

---

[9]Ibid., p. 13.

[10]Steve H. Hanke, "Demand for Water Under Dynamic Conditions," Water Resources Research 6: 1253-1261.

and charged an incremental price of 35 cents for each thousand gallons of water. Hanke compared the water consumed before and after the price system changed, and determined that the amount of water consumed by the individual households is sensitive to price, and that when a higher price is paid for the marginal amount of water consumed, the households are able and willing to use less of it. Not only did Hanke demonstrate that in this one community the amount of water consumed by households followed the law of demand, but that it was more than a temporary reduction. Some observers contended that if the households adjusted their water consumption pattern at all, it would be only a temporary adjustment; and after the novelty of the meters had worn off, the households would return to their original consumption patterns. This was clearly demonstrated not to be the case. However, some doubt still remained after the Hanke Study, i.e., could a price system change result in a reduction of water consumption, or would a price change alone also have the same effect?

R.A. Batchelor attempted to determine the effects of water-using appliances on the demand for water.[11] Batchelor develops his analysis on the basis that a person uses water to achieve some desired outcome but that outcome, and the accompanying use of water, is usually accompanied by the use of some other good or durable. The durables that Batchelor includes in his study are the most common water using appliances and applications of water found within the British household, namely washing machines, dishwashers, showers, garden sprinklers and cars. One of the more original ideas set forth within this study is the distinction and use of both household income and household wealth as determinants of water use. The inclusion of these variables, as well as water using appliances, should increase the

---

[11] R.A. Batchelor, "Household Technology and the Domestic Demand for Water," Land Economics 51: 208-223.

explanatory power of the water demand models, i.e., since the wealth of the household will affect the techonology of the water-using appliances within the home, wealth can be used to determine the level of water demand or shifts in the level, while the rate of growth of water demanded will be affected by the wealth elasticity of water demand.

A more recent contribution to the study of residential water demand was conducted by Donald R. Andrews and Kenneth C. Gibbs.[12] Using 1973 data from Dade County, Florida, these researchers develop two separate water demand models to explain water consumption for residential housing units. A total of 11 water companies are selected based on a dispersion of pricing schemes; price and quantity information concerning water is retrieved from company records and matched with block census data. Quarterly data is collected from a sample of 355 residential units.

The first model constructed by Andrews and Gibbs is based on the assumption that the consumer is aware of only his total water bill, and therefore, the average water price is the relevant variable with respect to price. This idea conformed to the conventional research methods employed up to this time. The generalized model takes the form:

$$Q_d = f\,(AP,\ I,\ RS,\ HWH,\ D_1,\ D_2,\ D_3)$$

where:

Q_d = Household water consumption in thousands of gallons

AP = Average price per thousand gallons

I = Annual household income

RS = Number of persons per household

HWH = Percent of households with hot water heat

D_i = Seasonal shifter variables

[12] Donald R. Andrews and Kenneth C. Gibbs, "An Analysis of the Effect of Price on Residential Water Demand: Metropolitan Miami, Florida," Southern Journal of Agricultural Economics (July 1975): 125-130.

The estimated regression coefficients (in natural logs) of this model are:

$$Q_d = 2.02 - 1.07AP + .000064I + .29RS + 3.92HWH + .08D_1 - .02D_2 - .02D_3$$

where:

$R^2 = .46 \qquad F = 176.37 \qquad d.f. = 1404$

Andrews and Gibbs construct a second model which is based on traditional economic theory i.e., "a consumer makes his decision concerning additional purchases of goods and services based on the price of the last unit, i.e., the marginal price."[13] They interpret the marginal price to be any price other than the minimum charge in a declining block pricing system. Consequently, any water customer who uses a small enough quantity of water to be within the first block of the rate structure is not considered as paying a marginal price, but rather a fixed rate, and therefore, is excluded from this second model. These customers account for about 24 percent of the sample residential units.

The general form of the marginal water consumption model is:

$$Q_d = f\,(MP, S_1, I, RS, HWH, D_1, D_2, D_3)$$

where:

Q_d = Household water consumption in thousands of gallons

MP = Marginal price per thousand gallons

S_1 = Zero price shifter

I = Annual household income

RS = Persons per household

HWH = Percentage of homes with hot water heat

D_i = Seasonal dummy variable

[13]Ibid., p. 127.

The estimated regression coefficients of this second model are:

$$Q_d = 3.12 - 1.85MP - 1.93S_1 + .00004I + .14RS + 7.79HWH + .06D_1 - .03D_2 - .03D_3$$

where:

$R^2 = .60$ $F = 267$ d.f. = 1055

The conclusion reached by Andrews and Gibbs is that the average price model is probably more useful than the marginal price model for predicting water demand because most consumers are not aware of the marginal price system and will respond primarily to total water bill. However, the authors do concede that "to be consistent with the economic theory the marginal price model has appeal."[14] Moreover, from an empirical standpoint their marginal price model explains more of the variance than does the average price model.

Another recent study was completed and published by Henry S. Foster Jr. and Bruce R. Beattie who attempted to estimate a generalized model which would allow for the effects of regional and size-of-city differences on urban residential water demand.[15] They suggest the following single equation model as representing the demand function for urban residential water:

$$Q = f(P, Y, R, N)$$

where:

Q = quantity of water demanded per household, i.e., per meter (1000 cubic feet per year);

P = average water price (dollars per 1000 cubic feet);

---

[14] Ibid., p. 130.

[15] Henry S. Foster Jr. and Bruce R. Beattie, "Urban Residential Demand for Water in the United States," Land Economics (February 1979): 43-58.

Y = median household income (dollars per year);

R = precipitation (inches) during the defined growing season; and

N = average number of residents per meter.

While acknowledging the simultenaity problem inherent in the use of average price they note that the use of data from the American Water Works Association precludes avoidance of the problem.

Unlike previous researchers Foster and Beattie specify a model with price in exponential form. This formulation permits price elasticity to vary with price whereas the power formulation of previous studies implicitly assumes that elasticity would remain constant for all price changes. A model is specified for the entire United States but is then disaggregated using dummy variables to account for regional and size-of-city differences. Using ordinary least squares they estimate the following model for the United States:

$$Q = .249e^{-.1278p} Y^{.4619} R^{-.1679} N^{.4345}$$

$$R^2 = .545 \qquad F = 63.69 \qquad df = 213$$

Using a stepdown F-test procedure on the disaggregated models they reject the null hypothesis that demand is invariant among regions of the U.S. but the results are not sufficient to reject the hypothesis that urban residential water demand is invariant among city-size strata.

A tabular summary of the empirical water demand studies is presented in Exhibit 2-1, which is included to demonstrate several points about their development. One of the more noteworthy developments in the past twenty years is the change from aggregated to cross sectional data. The use of aggregated data obscured the interacting effects of various independent or explanatory variables and resulted in poor estimates. The use of cross sectional data in more recent regression models has permitted more disaggregated and, consequently, better statistical estimates. The use of less

EXHIBIT 2-1

SUMMARY RESULTS OF THE EMPIRICAL STUDIES

Author	Publication	Study Area	Type of Data	Type of Model	Price Elasticity	Income Elasticity	R^2 of Best Model
Seidel & Bauman	Journal of American Water Works Association (1955)	441 Cities Nationally	Aggregated Cross-Sectional	Simple Linear	- .45	-	.62
Gottlieb	Land Economics (1963)	Kansas	Cross-Sectional	Exponential	-1.23 - .65	.45 .58	.67 .69
Headley	Land Economics (1963)	14 Cities around San Francisco	Semi-Aggregated Cross-Sectional	Linear Exponential	-	1.29 1.49	.69 .80
Howe & Linaweaver	Water Resources Research (1967)	39 Cities Nationally	Cross-Sectional	Exponential	-1.3	3.47	.72
Ware & North	Atlanta Economic Review (1967)	14 Cities in North Georgia	Cross-Sectional	Linear Exponential	- .61	.36	.86
Hanke	Water Resources Research (1967)	Boulder, CO	Time Series	Plotting	-	-	-
Turnovsky[16]	Water Resources Research (1969)	19 Cities in Massachusetts	Cross-Sectional	Linear	- .3	1.2	.82
Wong[17]	Land Economics (1972)	103 Communities around Chicago	Cross-Sectional Time Series	Exponential	- .28	.195	.81 .57
Andrews Gibbs	Southern Journal of Agricultural Economics	11 Communities around Miami	Semi-Aggregated Cross-Sectional	Exponential	-1.07 -1.85	.00006 .00004	.46(AP) .60(MP)

[16]Turnovsky, Stephen J., "The Demand for Water: Some Empirical Evidence on Consumer's Response to a Commodity Uncertain in Supply," Water Resources Research, Vol. 5 (April 1969), pp. 350-361. (This article is included in the summary of empirical investigation.)

[17]Wong, S.T., "A Model on Municipal Water Demand: A Case Study of Northern Illinois," Land Economics, Vol. 48, No. 1, (February 1972), pp. 34-44. (This article is included in the summary even though it was not reviewed in the body of this chapter because it is an empirical study but added to the methodology of investigating water demand.)

comprehensive data sets has also resulted in models which are able to explain less of the variation in water demand, i.e., the R^2s have decreased. This is the result of using disaggregated data which simultaneously improves the estimates of the coefficients of the model but reduces its explanatory power by increasing the dispersion of observed behavior.

Perhaps more difficult to explain is the apparent diversity of the estimates of price elasticity which range from very inelastic to relatively elastic. It is difficult to reconcile the results of these studies because of the diverse nature of the selected methodologies and regions. One reason for the diversity of estimates is that many of the early studies use distribution rather than consumption data which tends to result in elasticity estimates which are biased downward. Distribution data usually includes commercial, industrial, municipal as well as domestic water demand; additionally, its use implicitly treats leakages in the transmission lines as purchased water.

Another noteworthy development is the narrowing of the definition of price to include only the direct payments the individual customer must pay to receive a given level of water. When defining price many researchers avoid or ignored the problem of an interdependent function of price and quantity based on the prevailing block price system by using an average price for a segment or subset of their sample. This problem has continued to disturb researchers until only a few years ago when Andrews and Gibbs were bold enough to define the price of water in both average and marginal terms. No research has directly addressed the question of interdependence of price and quantity as a result of the declining block pricing systems generally employed; however, this question will surely be investigated using some form of simultaneous equations in the future.

CHAPTER 3
METHODOLOGY AND DATA COLLECTION

Introduction

The primary objective of this study is to determine and compare the residential demand for water in each of the Arkansas Water Resource Planning Areas. In order to accomplish this objective, water demand models will be developed for each AWRPA based on the application of theoretical, empirical and analytical methods which are internally consistent. The first section of this chapter develops the theoretical considerations of the study, while the empirical and analytical considerations are discussed in later sections. The sections pertaining to the pilot study and data collection are more technical and may be skipped without a loss of understanding of the thrust of the overall study.

The development of a water demand model begins with understanding of the term "demand" in an economic sense. The terms "water use", "water requirements", and "water demand" are often used interchangably, although they sometimes mean very different things. For purposes of this study, water demand will be defined as the quantity of water that would be purchased or used at a given price, while the water demand schedule shows the various amounts of water that would be purchased at different prices. Economic theory suggests that more of a good will be consumed by individuals as the price of the good is lowered, and vice versa, if all other factors are held constant. This theory provides only a general conceptual framework in understanding demand. In order to obtain more information about the shape and location of the demand schedule, an understanding of the other determinants of demand is required.

Factor Groups of Demand

The determinants of demand attempt to explain the willingness of consumers to purchase a quantity of a good at a given price. The shape and location of the demand schedule is derived by the interaction of the determinants of demand or these other factors. The determinants of demand, then, are a comprehensive list of those factors which interact to determine the demand schedule. In order to insure that a listing of the determinants of demand is indeed a comprehensive list, the wording is expressed intentionally in nonspecific terms, such as:

Determinants of Demand

1. Price of the good
2. Price of other goods
3. Financial resources of the consumers
4. Uses of the good
5. Consumers' tastes and preferences
6. Quality of the good
7. All other factors that affect the good's demand

The above list is useful in organizing the conceptual framework from which to direct a search for the relevant factors affecting demand of a given good. From this list several factor groups can be specified from which measurable variables can be selected on which to base the water demand models. The relevant factor groups include price, financial resources, demographic characteristics of the household, climatological conditions, internal and external uses of water, and other factors. Each of these will be discussed in turn.

Price. The price a consumer pays for water can affect the amount of water an individual decision-making unit will choose to consume. This idea is basic to economic thoery and the very heart of this study. However, price can be viewed in several different ways. Is the relevant price the average price or the marginal price that the consumer pays? Often there is no distinction between these two prices over the relevant range of quantities purchased by the consumer. However, water as priced by city water systems is an exception. The declining block pricing system used by most water utilities allows more than one value to emerge as the price varies. Although several different values can be determined as the price variables, only one of these variables will be included in the model at a time. A determination must be made as to whether the household unit is aware and responsive to the declining block pricing system. If the assumption of total bill is accepted, the average price is the appropriate price variable, while a marginal price would be the correct variable if the assumption is made that the household is aware of the pricing system.

Financial resources. The financial resources of the decision-making unit can contribute to the amount of water a household will choose to consume since they affect the ability to purchase different kinds and amounts of various goods. The purchasing power of the household is limited by the amount of money and other assets accumulated at a point in time (wealth) and that received during a particular time frame (income). Either wealth or income can be used to pay the water bill, and therefore both are included in the group of financial resources.

Demographics. Since the decision-making unit for this study is the household, a decomposition of the household will be necessary in order to understand the interacting factors that determine residential water demand. Since some water usage is related entirely to the individual and not the household, it is necessary to investigate the relevant demographic characteristics of the members of each household.

Climatological conditions. In the case of water, weather may interact with other determinants of demand to influence the level of usage. Since water can be supplied to the households directly from nature, some of the uses for water supplied by the urban water system may be duplicated. That is, the household may require water from the water system as an alternative or backup system to nature. This would certainly be the case with watering the lawn or garden. Climatological conditions may influence other water uses, such as the number of baths an individual desires. Therefore, weather should be included as an independent variable factor group which may influence the demand for domestic water.

Internal and external uses. The number of different uses that a resource can be put to can influence its desirability. In the case of water, the greater the number of ways it can be employed within the household the more water demand is affected. Water has many basically different types of uses, ranging from drinking, cleaning, and recreation, to enhancing a person's aesthetic surroundings. Each of these uses can influence the level of desirability, and thereby the consumer's willingness to acquire water. Since the uses of water are so diverse, it will be useful to divide this factor group into those uses that exist both inside and outside the home.

Other factors. The six preceding groups of independent variable factors combine to make a rather comprehensive list. However, they do not include all possible factors that could interact within the decision-making unit and contribute to the desirability of water. Some other factors that may also affect domestic urban water demand are the quality of the water, personal hygiene habits, the age of water-using appliances, the maintenance of the household's plumbing system, and so on. This final independent variable factor group will therefore consist of possible independent variables that do not fall within one of the factor groups, but nevertheless will be included as a possible predictive variable.

It should be noted that the preceding list of seven independent variable factor groups excludes the price of other goods. This determinant was not included as a factor group because few other commodities can serve as close substitutes for water in most of its uses. Also, when considering water as a component commodity in a bundle of goods which provides utility, it is assumed that the comparative price of water to these other goods will either be very small, or the ratio of the other good's price to water's price will remain relatively constant. In economic terms it is assumed that the price cross elasticity is not significantly different from zero.

The generalized model to be utilized in this study will take the form:

$$Q_w = f\ (P,\ F,\ D,\ C,\ I,\ E,\ O)$$

where:

Q_w = quantity of water demanded by the individual household from the urban water system.

P = those independent variables that combine to comprise the factor group Price.

F = those independent variables that combine to comprise the factor group Financial Resources.

D = those independent variables that combine to comprise the factor group Demographics.

C = those independent variables that combine to comprise the factor group Climatological Conditions.

I = those independent variables that combine to comprise the factor group Internal Uses.

E = those independent variables that combine to comprise the factor group External Uses.

O = those independent variables that combine to comprise the factor group Other Factors.

Independent Variables

The generalized model and the identified factor groups will be utilized as a framework for selecting prospective dependent variables. A list of these variables follows. In addition, the units of measurement and the variable's expected effect on the dependent variable will be discussed.

Average Price (AVP). The average price for water is the dollar amount paid per month by a household divided by the amount of water consumed during that month in thousands of gallons of water. By design of the pricing schedules used by the water departments, an inverse relationship between this variable and the quantity of water is not only assumed but expected. Since the urban water industry uses a declining block rate pricing structure from tradition, the price of water (regardless of how it is defined) will decline as the volume increases. This relationship will neither negate this variable's usefulness nor automatically establish a useful policy implication. Although the regression analysis will establish a quantitative relationship between average price and quantity of water used by the households, this correlation in and of itself will not establish causation.

Marginal and Incremental Price (MRP and INP). The marginal price can be measured in two different ways. If the relevant unit of measurement of water delivered to the household is a thousand gallons, then the cost of the last thousand gallons of water delivered becomes the marginal cost. However, the declining block rate pricing structure used by the water departments in this study includes the fixed costs of water within the first block rate. Assuming that the first block rate is, in fact, a fixed price, a household would have to consume a larger quantity than is included in the first pricing block before a marginal price is paid. However, if the first block rate is assumed to be only a higher rate, rather than a fixed rate, a marginal cost can be derived if the household does not use a large enough quantity of water to enter the second pricing block. Therefore, two different means of deriving marginal cost are possible, either by including only those block rates that are greater than the first block rate, or by counting the block rate regardless of the amount of water consumed.

Either of the above methods of determining marginal price will produce a variable with units of dollars per thousand gallons of water, and will have an expected inverse relationship between it and the dependent variable. In order to minimize confusion that could arise from having two different marginal prices as variables, the term incremental price will be used. Incremental price will be defined to be the price paid for the last thousand gallons of water when the last thousand gallons is outside of the first block in the pricing structure. The term marginal price will be the price paid for the last thousand gallons of water delivered to the household regardless of whether or not it is in the first block of the pricing structure.

Income (INC). Yearly total income is an excellent variable to measure the ability of a household to purchase water. Since the water bill is a recurring expense, logic dictates that a flow variable should be used to measure a household's financial resources. Perhaps a better measure would be disposable income. However, gross income can be assumed as highly related to disposable income with no major problems developing from the use of this variable. The fact that the time units are different should also present no problems. If one assumes that the household will behave according to its perceived permanent income, then the fact that yearly income is used should enhance rather than detract from the time units of this variable. Income will be measured in hundreds of dollars per year and should be positively related to the amount of water consumed by a household.

House Value (HOV). It is conceivable that a household does not purchase water out of a flow of money but rather from a stock of wealth. If this is the case, some measure of wealth is needed. Since the major investment item for most American families is the house they live in, this can serve as a proxy variable for wealth. The pitfalls of this logic should be obvious to most readers, e.g., overestimating those that have outstanding loans on their house and underestimating the wealth of those who have additional stocks of wealth. However, this measure will be used for its relative data collecting ease, and because house value might also influence the quantity of water consumed, in that house value can be assumed to include the aesthetic beauty of the accompanying lot and yard. This variable will be measured in hundreds of dollars. The expected relationship between house value and water consumption is positive, that is, the greater the house value the more water one should expect the household to consume.

Financial Resources (FIN). Since either a flow of funds or a stock of funds can be used to purchase water, it is also conceivable that some decision-making units combine those two sources of financial resources when deciding on their water consumption levels. Therefore, a third variable will be included in this variable factor group which combines income and wealth into a single variable. The technique of combination will be a simple linear combination or merely the addition of the two previous variables to produce the variable Financial Resources. This variable will have the units of dollars, and is expected to have a positive relationship to the dependent variable.

Number of People per Household (PPH). Since the decision-making unit is assumed to be the household, the composition of its members could affect water demand. The units for this variable will be people per household.

Household Average Age (AGE). The average age of persons in the household is another variable which is included to compensate for the assumption that the household is the decision-making unit. Since older people usually appear to be more interested in the upkeep of their lawn and garden, it is initially assumed that the average age of the people in the household is positively related to the amount of water used. The units of this variable will be years.

Number of Children per Household (CHI). Children have a special affinity for water during the summer months (the season in which this study was conducted) and can significantly affect household water usage by playing with water. To account for this possibility, the variable for children will be added where each person under the age of eighteen is counted as a child. The units for this variable will be children. The initial assumption is that the number of children in any household will be positively related to the level of water usage.

Temperature (TEM). The temperature is one measure of climatological conditions that could effect water usage, since increased temperature may cause an individual to bathe more often, or water his lawn and garden. This independent variable, Temperature, will initially be included with the units of tenths of degrees on the Fahrenheit scale. A-priori logic dictates that an increase in temperature should be associated with an increase in the amount of water consumed.

Precipitation (PRE). A second measure of climatological conditions is the amount of atmospheric moisture an area receives. Although this independent variable should correlate closely with the temperature, it has particular interest for this study and will be included as a second independent variable in the climatological factor group. Precipitation will have the units of tenths of inches of water and should be inversely related to water consumption.

Number of Water Using Appliances (APP). The presence of water using appliances within the household can affect the amount of water an individual household consumes. Instead of identifying and counting each appliance as a separate variable, a single inclusive variable will be used. Water using appliances are defined for this study as dishwashers, clotheswashers, and garbage disposals. Each type of appliance will be given equal weight so that the variable will be simply the sum of each of the appliances listed above. The units for this variable will be appliance, and it is initially assumed that the number of appliances present in the households will be positively related to the amount of water consumed.

Number of Water Using Fixtures (FIX). Similar to the above variable, the number of water fixtures in a household can influence the amount of water a household chooses to consume. Water fixtures will be defined for this study as sinks, bathtubs, shower heads, and toilets. Although these fixtures generally use different amounts of water, each will be assumed to have the potential for equal water usage, and will be added together with equal weights to produce the variable named "Number of Water Using Fixtures". The expected sign of this initial predictive variable's coefficient is assumed to be positive, while the units will simply be fixtures.

Total Number of Water Faucets (FAU). The more access individuals have to water, the more they might use it. Recognizing the truth of this statement, an attempt to include this as an initial predictive variable will be made by counting the total number of water outlets in the household. The counting of the water faucets will not include the hot water outlets because this would amount to double counting, if hot water were assumed to be only a condition of the water. A positive relationship is expected between water demanded and the variable total number of water faucets. Units for this variable will simply be water faucets.

This variable appears to be simply a linear combination of the two previous variables; however, its inclusion allows for the inclusion of water outlets in the household that are either fixtures or appliance plus outlets that are classified as neither of these. Therefore, the value of the total number of water faucets will be equal to or greater than the sum of water fixtures and water using appliances. Other water faucets included in this variable will be those faucets located outside the household, such as the garden hose faucet. Although this expands this variable slightly out of the variable factor group, "inside water uses", it will not substantially affect the analysis. This fact must be considered when conducting the search for the best set of independent variables.

Size of Swimming Pool (SWI). The presence and size of a swimming pool can affect water usage, and will therefore be included in the list of initial independent variables. The units for this variable will be gallons of water. If no swimming pool is present, the variable will have a value of zero. A positive relationship is expected between this variable and the quantity of water demanded.

Lawn Area (LAW). The presence of a lawn and its size can be a major cause for water consumption, and will be included as an intial variable with units of hundreds of square feet. Positive relationship between the dependent variable and lawn area is initially expected.

Garden Area (GAR). Like the lawn area, a garden can affect the level of water consumption and will be included as a variable with characteristics similar to the variable lawn area.

Perceived Water Purity (PUR). The quality of a good can influence its consumption, and in the case of water, purity can be a substitute for quality. Although technical definitions of water purity are available, the relevant factor is purity as the consumer perceives it. Therefore, the variable "Perceived Water Purity" will be included which attempts to measure the consumer's perception of the quality of water. The definition of purity is left to the consumer as he views his water requirements and uses; therefore, conventional units of measurement are not used. Instead, a continuum from high "Perceived Water Purity" to low "Perceived Water Purity" will be employed. It is assumed that the consumer will use a larger amount of water if he perceives that the water purity is good for his uses. This leads to an expected positive sign for the coefficient of this variable.

Town Size (TOW). The size of the town in which an individual consumer lives could contribute to the amount of water the household uses. This relationship may be caused by the willingness of small town consumers to be more family oriented and spend more time at their residence. If this is the case, the small town consumers have more access to the water delivered to their houses than do consumers who live in larger towns. Therefore, the expected relationship between town size and water use for a household will be negative. The town size will be measured by population and will have the units of hundreds of people.

Age of Water Using Appliance (AWA). In an attempt to recognize technology as a factor which might affect water demand, the variable Average Age of Water Using Appliances will be included as a possible predictive variable. Each appliance (dishwasher, clotheswasher, garbage disposal) will be assigned equal weight and their age in years will be averaged to produce units for this variable of years. If technology has produced water efficient appliances over the years, the expected sign of the coefficient of this variable will be positive, the interpretation being that the older appliances use more water than the newer ones.

A summary of the possible variables is presented in Exhibit 3-1 which includes the variables' symbols, names, and units of measurement. The

EXHIBIT 3-1

LIST OF VARIABLES

Symbol	Variable Name	Units of Measurement
Q	Water Quantity	gallons/month
AVP	Average Price	$/gallon
MRP	Marginal Price	$/gallon
INP	Incremental Price	$/gallon
INC	Income	$/year
HOV	House Value	$
FIN	Financial Resources	$
PPH	Number of People Per Household	people
AGE	Household Average Age	years/person
CHI	Number of Children Per Household	children
TEM	Temperature	degrees/Fahr.
PRE	Precipitation	inches of rain
APP	Number of Water Using Appliances	appliances
FIX	Number of Water Using Fixtures	fixtures
FAU	Total Number of Water Faucets	faucets
SWI	Size of Swimming Pool	gallons
LAW	Lawn Area	sq. ft.
GAR	Garden Area	sq. ft.
PUR	Perceived Water Purity	none
TOW	Town Size	people
AWA	Age of Water Using Appliances	years

NOTE: The underlined letters in the Variable Names are the letters used to create the symbol. Since several tables are used in this study, the reader is urged to aquaint himself with the symbols since often space will not permit the use of the full name.

variables are grouped by the factor groups so that the reader can follow the logic behind the development of the factor groups. These groups will aid in the analysis of the data.

Pilot Study

The majority of the variables used in this study were collected using questionnaires mailed to individual households throughout the state of Arkansas. Other sources of data include city water records, and state and federal reports. To determine the size of sample and to validate the method of data collection, a pilot study was conducted. The results of the pilot study provided the guidelines for the majority of the data collection.

The process of data collection started with a pilot study in order to help determine the optimal method to collect the data used in the regression analysis. The objectives of the pilot study were to determine:

1. the best format for the questionnaire,
2. if the use of bulk mail rates would significantly affect the response rate relative to the use of first-class mail rates, and
3. the usable return rate that could be expected from the questionnaire mail-out.

From the list of variables presented in Exhibit 3-1, those variables that were deemed appropriate to be measured directly from the households were selected. For each variable several questions or sets of questions were constructed to elicit a correct response from individuals. The set of possible questions was refined and reduced until two different questionnaires were constructed. Although each questionnaire was designed to achieve the same results, the manner in which the questions were asked differed. Because there was no appropriate way to select one of the questionnaires as optimal without empirical testing, both questionnaires were used in the pilot study.

Another factor which was tested was the effect on the response rate of the method of mail delivery of the questionnaire. The range of available postage handling schemes and rates was extensive. First-class mail with an affixed special issue stamp was chosen as the most elaborate that would be tested. The bulk non-profit organization rate using a rubber stamp to print the postage was the least expensive method tested. Respective mail prices were $0.13 and $0.021 per letter delivered. If the usable return rates were significantly different, the cost differences between these postage rate could justify the first-class postage, even though it would mean an extra $109.00 per 1000 questionnaires mailed. The return postage would be the same regardless of the delivery rate selected. The returned questionnaires were received in pre-addressed envelopes with printed first-class postage. The return postage rate schedule required payment for only those envelopes that were returned at a cost of $0.18 each.

After the letter of introduction and the questionnaires were completed, two hundred residents within the pilot study city were selected randomly. The random procedure for selecting individual households consisted of using every n^{th} name in the local telephone directory, where n was determined by dividing 200 into the total number of names in the telephone directory. Although this method had the possibility of generating a bias into the sample design, it was overlooked in this study, based on the assumption that nearly all households that are connected to the urban water system are also listed in the telephone directory.

The compiled results demonstrated that 21% of the outgoing letters were returned when they included Questionnaire 1 (Exhibit 3-2), while a usable return rate of 27% was observed with Questionnaire 2 (Exhibit 3-3) when used with the introductory letter (Exhibit 3-4).

EXHIBIT 3-2

ARKANSAS RESIDENTIAL WATER USER QUESTIONNAIRE 1

1. What are the ages of the people living in your household?

____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____

2. What are the dimensions of your home lot? (Please estimate in feet and indicate on the drawing or use your own drawing.)

3. What percent of the total area of your home lot would you estimate your lawn to be?

Less than 20% _____	40% to 60% _____	80% to 100% _____
20% to 40% _____	60% to 80% _____	

4. What percent of your home lot is devoted to a garden?

Less than 10% _____	20% to 30% _____	40% to 50% _____
10% to 20% _____	30% to 40% _____	More than 50% _____

5. Do you have any type of swimming pool? Yes No

6. If you have a swimming pook, what are its dimensions?

Length= _____ feet	Depth at shallow end = _____ feet
Width = _____ feet	Depth at deep end = _____ feet

OR

My pool's total capacity is = ________ gallons

7. Did you water your lawn last summer?	Yes	No
8. Did you water your garden last summer?	Yes	No
9. Do you have an automatic dishwasher?	Yes	No
10. Do you have a clotheswasher?	Yes	No
11. Do you have an automatic garbage disposal?	Yes	No

12. Please list the length of time you have owned each of the following.

Automatic dishwasher	_____ years
Clotheswasher	_____ years
Automatic garbage disposal	_____ years

13. How many sinks are in your house? ________

14. How many bathtubs are in your house? ________

EXHIBIT 3-2 Continued

15. How many shower heads are in your house? ______

16. How many toilets are in your house? ______

17. Counting the faucets both inside and outside of your house, what is the total number of faucets you have connected to your water lines?

______

18. Do you have a water purifier attached to your water system? (Do you have a device in your house which is connected to your plumbing that is designed to improve the quality of the water such as filtering, removing minerals, etc.?)

Yes No

19. Does any member of your household use a water conditioner of some type to change the property of the water? (Not including soap and bleach or other cleaning agents, does anyone within your household use some sort of chemical agent in the water to change its quality?)

Yes No

20. How do you rate the purity of the water you receive from your water department?

_____ Much too pure
_____ Probably purer than I need
_____ About as pure as I want
_____ Pure with regard to health, but it contains a few non-harmful particles
_____ I sometimes wonder about the purity of the water
_____ The purity is totally unsatisfactory

21. If you were to sell your home today, what do you think it would bring in today's real estage market?

_____ Less than $15,0000
_____ $15,000 to $30,000
_____ $30,000 to $45,000
_____ $45,000 to $60,000
_____ $60,000 to $75,000
_____ More than $75,000

22. What is your total household yearly income?

_____ Less than $5,000
_____ $5,000 to $8,000
_____ $8,000 to $11,000
_____ $11,000 to $14,000
_____ $14,000 to $17,000
_____ Over $17,000

23. According to your water bill receipts for last summer, how much water did you use for the following months? (This question will require that you locate your old water bills and read the monthly usage for each month. If you cannot locate these receipts, please go on to the next question. If you can locate this information I will very much appreciate it.)

EXHIBIT 3-2 Continued

June _______ gal.   July _____ gal.   August _____ gal.

24. What is the source of your water? (What company or what city delivers the water to your home?)

____

Thank you for your cooperation! If you will please send this form back using the envelope provided, I will be very much in your debt and shall try to repay you by providing the best possible study results to your community leaders.

EXHIBIT 3-3

ARKANSAS RESIDENTIAL WATER USER QUESTIONNAIRE 2

1. What are the ages of the people living in your household?

______ ______ ______ ______ ______ ______ ______ ______ ______ ______

2. Do you have any of the following items connected to the plumbing of your house? If so, how long have you had them?

Dishwasher	Yes	No	_______	Years
Clotheswasher	Yes	No	_______	Years
Automatic garbage disposal	Yes	No	_______	Years

3. What is the size of the lot your house is located on? (Please estimate the dimensions.)

Width = ________ feet Length = ____________ feet

4. What percent of your lot would you estimate your lawn to be? Please place a check along the line at the point that best indicates your estimate.

0% —— 25% —— 50% —— 75% —— 100%

5. If you have a garden on your lot, what would you estimate its portion of your total lot to be?

0% —— 10% —— 20% —— 30% —— 40%

6. Do you have a swimming pool of any kind? Yes No

7. If you have a swimming pool, what would you estimate its size to be in gallons of water when it is full? _______ gal.

8. Did you water your lawn last summer? Yes No

9. Did you water your garden last summer? Yes No

10. How many of the following items do you have connected to the plumbing in your house?

Sinks _____
Toilets _____
Bathtubs _____
Shower heads _____

EXHIBIT 3-3 Continued

11. Counting all the fixtures in the above question and any other fixtures you have whether inside or outside your house, what is the total number of water faucets connected to your water line? _______

12. Do you have a water purifier attached to your water system? (Do you have a device in your house which is connected to your plumbing that is designed to improve the quality of the water such as filtering, making the water "soft" or "hard", etc.?)

    Yes   No

13. Does any member of your household use a water conditioner of some type to change the property of the water? (Not counting soap and bleach or other cleaning agents, does anyone within your household use some sort of chemical agent in the water to change its quality?)

    Yes   No

14. How do you rate the purity of the water you get from your water department? (Pace a check at the point along the line that most accurately describes your feelings.)

Totally Unfit for My Use		Safe to Drink, but Has Impurities		Much too Pure

15. If you were to sell your home today, what would you guess it would be worth in today's real estate market?

Less than $10,000	$30,000	$50,000	$70,000	More than $90,000

16. What is the amount of your total household income for last year?

Less than $5,000	$8,000	$11,000	$14,000	More than $17,000

17. According to your water bill receipts for last summer, how much water did you use for the following months? (This question will require that you locate your old water bills and read the monthly usage for each month. If you cannot locate these receipts, please go to the next question. If you can locate this information, I will very much appreciate it.)

    June ____ gal.   July _____ gal.   August _______ gal.

EXHIBIT 3-3 Continued

18. What is the source of your water? (What is the name of the water company or city water department that delivers the water to your home?)

____

Thank you for your cooperation! If you will please send this form back by using the envelope provided, I will be very much in your debt and shall try to repay you by providing the best possible study results to your community leaders.

EXHIBIT 3-4

UNIVERSITY OF ARKANSAS
COLLEGE OF BUSINESS ADMINISTRATION
FAYETTEVILLE, ARKANSAS 72701

January 9, 1978

Dear Arkansan:

I am writing to ask for your help. Currently, members of the University of Arkansas faculty are trying to gather facts and information about water usage within this state. This effort is being undertaken in order to develop a better economic understanding of the water situation within various communities of Arkansas. When this study is completed, its results will be made available to those authorities and individuals that are concerned with this type of information. If this study is successful, many communities within the state will benefit from its results through more efficient water systems. However, in order to complete the study, it is necessary to acquire some information from individual citizens like yourself.

Your name has been selected at random from an extensive list of people that obtain water from a community water system. I an interested in finding some answers to some rather specific questions, so that an overall profile can be constructed of Arkansas water users. The information obtained from the questionnaires will be programmed into a computer for analysis in such a manner that it will not be possible to trace individual answers. All the individual questionnaires will be destroyed after the information is placed anonymously into the computer.

If you would take a few minutes and complete the accompanying questionnaire, you will be helping not only the University faculty but possibly yourself, since your community could be directly helped by the study's results. You will probably notice a typed line on the front of the return envelope that you may wonder about. I want to assure you that this is for mail handling purposes only, since many research projects use these same envelopes, some means of sorting returned letters at the mail room is necessary.

If for some reason you would prefer not to answer any question, I would appreciate it if you would skip that question and continue with the other questions and return the partially completed questionnaire. For your convenience, an addressed envelope with prepaid postage is included which I hope you will use to return the completed questionnaire.

Sincerely,

Joseph A. Ziegler
Professor of Economics
University of Arkansas

Although a 2% difference was observed by using the first-class postage as compared to bulk rate postage (25% versus 23% respective overall usable return rate), the cost to obtain the higher return rate ($0.109 per letter) outweighed the benefits. By increasing the size of the initial mailing by about 2%, the bulk rate postage would produce the same number of usable responses at a substantial cost savings over the use of first-class postage.

The results of the pilot study demonstrated that the optimal combination for generating usable responses from a mail questionnaire was to send Questionnaire 2 using bulk rate postage. However, two critical questions still remained to be answered concerning the pilot study:

1. Was the sample selected for the pilot study a randomly-selected sample?

2. Did the non-respondents of the pilot study posess similar characteristics to the respondents?

To test for randomness in the selection procedure used for the pilot study, the sample statistics were compared to the population parameters. The usable respones received from Questionnaire 2 were chosen as the relevant sample since this questionnaire was preferred. Two different characteristics of both the population and the sample were compared in order to test for randomness. If the differences between the corresponding parameters and the statistics are statistically insignificant, the assumption of random sample household selection could be made. The two characteristics to be compared are:

1. The number of people per housing units in occupied housing units.

2. The income of families and unrelated individuals living in a housing unit.

Both of the above characteristics are confined to the city which was used for the pilot study and will be referred to as: (1) Pop. per house; and (2) Family income. The sample statistics were derived from the pilot study in which 100 questionnaires were mailed and 33 usable responses were returned. The parameters were derived from the 1970 United States Census Report for the city in which the pilot study was conducted.[1] Exhibit 3-5 compares the unadjusted values of the parameters and statistics.

In using the 1967 consumer price index, an adjustment factor of 1.54 was applied to the dollar value of income to adjust the parameter to current terms, i.e., [179(1977 index)-116(1970 index)]/116=1.54. Family income was adjusted as presented in Exhibit 3-6.

EXHIBIT 3-5

UNADJUSTED PARAMETERS AND STATISTICS OF THE PILOT SAMPLE STUDY

	n	s	x	u
Pop. per house	27	1.01	2.77	2.7
Family income ($)	27	4,317	14,285	9,263

EXHIBIT 3-6

ADJUSTED PARAMETERS AND STATISTICS OF THE PILOT SAMPLE STUDY

	n	s	x	u
Pop. per house	27	1.25	2.77	2.7
Family income ($)	27	4,317	14,285	14,265

An hypothesis test was performed on the above data to test whether the sample was selected from the population. Using a student t_{test} with 26 degrees of freedom and 10% error level, both statistics were statis-

[1] 1970 Census of Housing, General Housing Characteristics (Washington, D.C.: U.S. Government Printing Office, 1971), No. HC-1-A5, p. 543.

tically demonstrated to come from the population with a .90 probability. The acceptance limits for the t_{test} value using d.f.=32 and error level of .1 for a two-tail test are:

$$-1.697 \leq t_{test} \leq 1.697$$

The t_{test} for pop. per house is:

$$\frac{X - u}{\frac{s}{\sqrt{n}}} = \frac{2.77 - 2.7}{\frac{1.01}{\sqrt{27}}} = .360$$

The t_{test} for family income ($) is:

$$\frac{X - u}{\frac{s}{\sqrt{n}}} = \frac{14,285 - 14,265}{\frac{4,317}{\sqrt{27}}} = .024$$

Because the statistical test above did not demonstrate a statistically significant difference between the characteristics of the sample and the population from which it was intended to be selected, it is assumed that the techniques used in the pilot study generated a random sample and further, that the same sampling techniques, when applied to other cities within the state of Arkansas, will produce similar results.

One final question remains to be answered with respect to the results of the pilot test. Namely, do the households that responded to the questionnaire have similar characteristics to those that chose not to respond? Although this question was partly answered by the above hypothesis test, a more direct approach was taken. The 73 households which did not return Questionnaire 2 were contacted by telephone in an attempt to generate responses from them. A total of 19 additional responses were received by this second wave technique. Assuming that those households that responded to the second wave were a representative sample of all households that did not respond to the initial mailing, an hypothesis test was again performed. This time the question to be answered was whether or not the second wave respondents had similar characteristics to the first wave respondents. Assuming the population to be the first wave respondents, and using the same set of characteristics, Exhibit 3-7 presents the parameters and statistics internal to the pilot sample study.

EXHIBIT 3-7

INTERNAL PARAMETERS AND STATISTICS OF THE PILOT SAMPLE STUDY

	n	s	x	u
Pop. per house	19	1.6	2.5	2.77
Family income ($)	19	5,421	12,382	14,285

No adjustment in the data is needed this time since the dollar terms were collected during the same time period. Therefore, the hypothesis test using the above data was performed. The range of acceptance is -1.734 to 1.734 for a t_{test} (with d.f.=18 and an error level of .1 for a two-tail test.

The t_{test} for pop. per house is:

$$\frac{X - u}{\frac{s}{\sqrt{n}}} = \frac{2.59 - 2.77}{\frac{1.6}{\sqrt{19}}} = -.49$$

The t_{test} for family income is:

$$\frac{X - u}{\frac{s}{\sqrt{n}}} = \frac{12,382 - 14,285}{\frac{5,421}{\sqrt{19}}} = -1.53$$

Since the t_{test} values for both the pop. per house and the family income ($) characteristics are within the acceptance limits of the decision rule, the assumption was made that the characteristics of people who responded and those that did not were not statistically different.

The pilot study statistically demonstrated that the sampling techniques to be used in collecting most of the data for this project were valid. Assuming that the two characteristics chosen for the hypothesis tests were representative of all characteristics with respect to variability in a random sample, the use of the mail questionnaire sent to individuals randomly selected from the local telephone directories will produce statistically random samples of each city selected.

Data Collection

The pilot study provided most of the information needed to design the sample for this study. However, two relevant questions remained before the questionnaires could be mailed out, i.e., the size of the sample and the selection of sample cities. The answers to these questions involved consideration of the trade-off between the precision and cost to obtain a sample. The cost of increasing the sample size is directly related to the costs of mailing and receiving each questionnaire, while the precision of the sample may increase as the sample size increases. The cost of varying the number of cities is not as obvious. Each response from a household must be matched with data that can only be obtained from the corresponding city water records. The marginal costs of obtaining additional observations from the same city is relatively small compared to the start-up costs to obtain the first observations from the additional city's water records. However, the precision considerations of limiting the number of cities rest in the lack of variability with each variable whose variation is a function of city, i.e., water price, temperature, and precipitation are three variables which will remain fixed for all households selected from a city in a given time period.

In considering sample size, both the accuracy and the level of acceptable error must be considered. In performing the earlier hypothesis tests, an error level of 10% was used. Although this level was arbitrarily chosen, it is consistent with similar studies. The error level is a quantitative statement of the "willingness to be wrong" on the part of the researcher, or the level of confidence that is desired in the study. An error level of 10% corresponds to a confidence level of 90%. The accuracy is the amount of variation between the estimate of the parameter and the parameter itself. Accuracy can be stated in either absolute or relative terms and will specify the numerical distance between the statistic and the parameter. An accuracy level of 12% has been chosen after several interactions of solving the formula for sample size. A larger accuracy level would reduce the size of the sample and therefore cost less, while a small accuracy level would increase the costs. A 12% accuracy level was the smallest possible given the budget constraint of this study. The sample size required from a single city can be determined by the following formula:

$$n = \frac{Z^2 \ s^2}{(X-u)^2}$$

where:

n = sample size
z = error level expressed in standard deviates
s = standard deviation obtained from pilot study
(X-u) = accuracy level desired, expressed in the maximum difference between the estimate and the actual central tendency

Using the characteristic of population per household from both the pilot study and the U.S. Census reports as the respective statistic and parameter in the above equation, a sample size of 26 is obtained in the following manner:

$$n = \frac{Z^2\ s^2}{(X-u)^2} = \frac{(1.64)^2\ (1.01)^2}{(.324)^2} = 26$$

Therefore, when considering a single city, a sample size of 26 was needed to insure that the results would be accurate within 12% of the true state of the world, with a confidence level of 90%. This still left unanswered the question of the number of cities to be sampled. Since the selection of a city would dictate the watershed region as a variable in the study, some consideration should be given to the relative weights of the regions. Population was selected as the relevant characteristic to assign relative weights on the assumption that the other characteristics are generally comparable with respect to their proportions by regions and because the proportions themselves are not critical to this study other than as a means of structuring the sample so that each region is generally proportionally represented. Given considerations of cost constraints and estimation requirements the relative weights of the number of cities included in the sample design coincides with the population distribution by region as precisely as possible. The relative population distribution and distributions of cities is presented in Exhibit 3-8.

EXHIBIT 3-8

NUMBER OF CITIES TO BE INCLUDED IN THE SAMPLE
DESIGN BY WATERSHED REGION

Region	Relative Pop. Distribution	Number of Cities	Relative Distributions of Cities
1	.25	5	.23
2	.22	4	.18
3	.11	4	.18
4	.36	6	.28
5	.06	3	.13
Total	1.00	22	1.00

The information from Exhibit 3-8 and the pilot study enabled a determination of the total mailing size which became the number of cities multiplied by the size of the sample needed per city, multiplied by the inverse of the usable return rate, or:

$$n_L = (n_C \cdot n_H) \frac{1}{RR} = (22 \cdot 26) (\frac{1}{.34}) = 1682$$

where:

n_L = total number of questionnaires to be mailed

n_C = number of cities

n_H = number of households per city

RR = usable return rate

A total mailing of 1,682 was deemed necessary to generate the required sample size in order to collect the necessary data for the regression analysis used in this study. This number was derived from the results of the pilot study which suggested that a 34% usable response rate could be expected, using bulk postage rates to send Questionnaire 2 (herein after referred to as simply the questionnaire). However, other researchers at the University of Arkansas reported that a slightly higher response rate could be expected from questionnaires mailed to households in Northwest Arkansas (the area in which the pilot study was conducted) than in other regions of the state.[2] Allowances for lower response rates were suggested by these researchers.

In addition, a sample size of twenty-six per town assumed not only twenty-six usable returned questionnaires per city, but also the corresponding data matched up by households from the city water records. That is,

---

[2]In 1978, Dr. William R. Darden established a consumer research panel for the State of Arkansas. While conducting this project, Dr. Darden gained an appreciation of the comparative household questionnaire response rates for the state.

less than twenty-six usable responses per city could be expected when the corresponding information from the city water records was unavailable. To account for these possible sources of a reduction in the number of responses needed per city, two additional factors were added to the calculations of the number of questionnaires to be mailed out. First, a postcard reminder was determined to be the most cost-efficient way of increasing usable response rate. The postcard was mailed to each household that received a questionnaire to encourage the individuals to return the completed questionnaire (see Exhibit 3-9). This action was designed to increase the usable response rate by a large enough margin to adjust for the loss of paired information from the city water records. The expected usable return rate was adjusted by town to comply with the return rate difference experienced by the other researchers. Also, the actual number of questionnaires mailed was adjusted so that the budgeted funds would be exhausted on the first mailing, rather than to risk an insufficient usable return rate and thus require a second mailing.

The selection of cities to be surveyed in each watershed region required that each city meet several preconditions. The individual city's water department must employ some form of rate pricing structure, have a metered water distribution system, and maintain records of water distribution to individual households accessable for purposes of this study. In addition, geographical dispersion within the watershed regions was considered desirable. To insure geographical dispersion, only one city per county was initially considered as a possible candidate for a sample city. This restriction led to the selection of using only the largest city in each county. Many of the counties' largest cities are relatively small (under 5,000 people), which was considered the lower limit of usable city size.

By identifying the largest city in each Arkansas county, seventy-five using other sources, such as city water records and government publications. Values for the three variables in the variable factor group "climatological conditions" were collected from the monthly climatological data reports of the U.S. Department of Commerce.[3] The values for the variable town size were extracted from an unpublished working paper of the Arkansas Department of Health.[4] The values for the remaining four variables - water quantity, average price, marginal price and incremental price, were collected from a search of the records maintained by the city water departments.

After the questionnaires were returned and the usable ones were identified, statistics on the amount of water delivered and the price paid for the water by the individual households were collected from the city water records. Research assistants were employed in this process since the dispersion of sample cities is rather extensive in the state of Arkansas. The water pricing schedules were obtained from each sample city along with the other information from the cities which allowed for the derivation of the marginal prices and the incremental prices.

Values for the variables not included in the questionnaire, except for the variable town size, were collected for three different months. That is, values for the variables in both the price and the climatological variable factor groups as well as the values for the dependent variable were collected for the months of June, July and August of 1977. Within the entire data set, these are the only variables that vary significantly by month. Since all the variables whose values change by month have been recorded separately by their

---

[3]U.S. Department of Commerce, Climatological Data, National Oceanic and Atmospheric Administration Environmental Data Service. June 1977, Vol. 82, No. 6; July 1977, Vol. 82, No. 7; August 1977, Vol. 82, No. 8.

[4]Arkansas Department of Health, Division of Engineering, Public Water Supplies (internal working paper, April 1977).

EXHIBIT 3-10

SAMPLE RETURN RATES

Area	No. in Mailing	No. of Returns	No. of Usable Returns	Usable Return Rates	No. of Complete Usable Observations
Blytheville	165	42	37	22	30
Forest City	165	45	27	16	27
Jonesboro	165	54	30	18	28
Paragould	165	55	33	20	31
Stuttgart	165	40	36	22	33
Region 1	825	236	163	20	149
Arkadelphia	165	68	36	22	28
El Dorado	165	45	35	21	32
Hot Springs	165	46	29	18	22
Pine Bluff	165	62	32	16	22
Region 2	693	221	132	19	104
Batesville	165	60	37	22	34
Harrison	165	54	38	23	35
Pochahontas	165	55	34	20	30
Walnut Ridge	165	50	38	23	30
Region 3	660	219	147	22	129
Benton	132	45	35	26	35
Conway	132	44	31	23	30
Fayetteville	200	62	46	23	35
Ft. Smith	132	40	27	20	25
Little Rock	132	32	28	21	26
Van Buren	132	34	28	21	24
Region 4	860	257	197	23	175
Magnolia	165	35	27	16	27
Mena	165	37	34	21	32
Texarkana	165	28	26	16	24
Region 5	495	100	87	18	83
State	3533	1033	726	21	640

respective month, the size of the input data set can be tripled (and therefore the effective sample size) by rearranging the values. The rearranging process employed with the data set simply inserted three observations where one was originally. Each of the variables that did not vary by month were recorded in the new data set three times while those values that did change by month were entered in the appropriate new observation, depending if it were the first, second, or third month. Because the dependent variable in the regression model was one of the variables whose values changed by month, i.e., June water quantity for household 1, July water quantity for household 1, August water quantity for household 1, the rearranging of the data set was not merely a tripling of the numbers. The variation in the dependent variable as been increased since the dependent variable is one of those variables that was recorded by month. Since the underpinnings of regression analysis provide a method of explaining variation in the dependent variable based on the variable is recorded by month is critical when tripling the data set in this manner. This procedure increased the effective sample size from 640 total observations to 1,920 observations.

Additional data editing reduced the effective sample size because of obvious clerical errors, i.e., when the price and quantity of water values were not aligned with the cities' proportioned rate structure, and by missing values which were not observed until the data set was computerized. The final result was that a data set with 1,773 observations for 22 different variables identified by watershed region in Arkansas was created. The majority of the observations were collected by primary research method while only 18% of the values were collected using secondary sources. Exhibit 3-11 presents some of the descriptive statistics of the data set. The mean, standard deviation, and the number of observations of each variable is presented by region and for the state so that the reader can gain an initial appreciation for the values before the regression analysis is presented.

EXHIBIT 3-11

DESCRIPTIVE STATISTICS

	State			Region 1		
Variable	N	Mean	Standard Deviation	N	Mean	Standard Deviation
Q	1770	81.009	67.766	426	102.877	90.961
AVP	1767	117.581	103.524	424	89.619	55.702
MRP	1770	85.323	93.817	426	64.395	52.580
INP	1692	70.364	18.118	411	55.756	9.283
INC	1773	142.714	53.238	426	143.978	51.627
HOV	1773	365.939	180.719	426	369.225	190.156
FIN	1773	508.653	210.672	426	513.204	217.450
PPH	1773	2.751	1.212	426	2.816	1.137
AGE	1773	43.977	19.913	426	41.667	19.911
CHI	1773	0.639	0.934	426	0.718	0.892
TEM	1773	794.166	30.621	426	796.917	39.608
PRE	1773	426.612	199.434	426	443.469	170.146
APP	1773	1.905	1.000	426	1.887	1.008
FIX	1773	6.670	2.388	426	6.584	2.397
FAU	1773	9.353	3.213	426	9.218	3.299
SWI	1773	6.419	48.397	426	1.190	11.893
LAW	1773	94.002	100.048	426	76.651	69.646
GAR	1773	8.993	20.108	426	5.224	11.395
PUR	1773	5.248	1.271	426	5.443	1.191
TOW	1773	250.263	349.631	426	177.816	75.317
AWA	1773	6.176	5.311	426	5.877	5.283

EXHIBIT 3-11 Continued

	----------Region 4---------			----------Region 5---------		
Vari- able	N	Mean	Standard Deviation	N	Mean	Standard Deviation
Q	472	81.387	60.393	239	79.050	60.442
AVP	471	99.683	39.063	239	142.747	133.166
MRP	472	78.696	27.881	239	112.256	138.004
INP	464	75.400	14.571	222	79.757	26.613
INC	474	149.436	49.454	240	127.975	61.462
HOV	474	384.639	174.464	240	346.562	185.375
FIN	474	534.075	201.001	240	474.537	226.462
PPH	474	2.822	1.325	240	2.437	1.225
AGE	474	41.608	19.269	240	50.777	17.828
CHI	474	0.708	1.027	240	0.400	0.817
TEM	474	789.649	24.364	240	792.158	27.847
PRE	474	553.843	255.771	240	380.495	91.757
APP	474	2.082	1.007	240	1.612	0.957
FIX	474	6.784	2.392	240	6.337	2.257
FAU	474	9.791	3.195	240	8.712	3.052
SWI	474	15.601	77.508	240	6.875	61.234
LAW	474	116.497	135.542	240	97.544	75.293
GAR	474	11.144	21.096	240	15.382	32.806
PUR	474	5.170	1.138	240	5.337	1.434
TOW	474	490.873	579.503	240	114.550	67.117
AWA	474	6.329	5.308	240	6.204	5.534

EXHIBIT 3-11 Continued

	----------Region 2---------			---------Region 3----------		
Vari-able	N	Mean	Standard Deviation	N	Mean	Standard Deviation
Q	279	59.451	45.615	354	72.502	55.349
AVP	279	188.034	189.396	354	102.369	41.360
MRP	279	109.736	174.186	354	81.918	38.025
INP	255	69.745	11.108	340	75.760	17.527
INC	279	145.193	52.436	354	140.228	52.886
HOV	279	354.032	170.761	354	359.466	180.466
FIN	279	499.225	195.460	354	499.694	212.330
PPH	279	2.763	1.223	354	2.779	1.091
AGE	279	42.557	19.980	354	46.435	20.809
CHI	279	0.634	0.853	354	0.618	0.966
TEM	279	808.767	25.192	354	786.759	27.112
PRE	279	401.939	111.710	354	313.457	176.264
APP	279	1.731	1.019	354	2.025	0.926
FIX	279	6.645	2.162	354	6.864	2.603
FAU	279	9.193	3.018	354	9.491	3.312
SWI	279	2.139	14.313	354	3.483	25.856
LAW	279	80.149	94.683	354	93.281	87.679
GAR	279	5.354	16.252	354	9.186	16.694
PUR	279	5.354	1.208	354	4.974	1.406
TOW	279	303.763	160.270	354	65.118	17.608
AWA	279	5.681	5.529	354	6.703	4.983

CHAPTER 4

DATA ANALYSIS

Introduction

The previous chapter presented the information collection methodology used to obtain the data for this study. The individual variables were extracted from independent variable factor groups, defined with their accompanying units, and collected from either primary or secondary sources. The chapter concluded with the presentation of the data's basic descriptive statistics by state and watershed regions.

This chapter centers on the analysis of the data set. As stated previously the objectives of this study are to find various residential water demand elasticities and to determine if the water demand schedules between the watershed regions are similar. The remainder of this chapter will discuss statistical considerations in utilizing appropriate tests to achieve the stated objectives. In addition it will discuss the statistical procedures as well as analyze the results.

Statistical Considerations

The first objective of this study is to determine whether the individual regional residential water demand predictive models are significantly different from each other. Or, stated differently, when the water demand models are developed for each region, are the differences between them so small as to be statistically insignificant? To address this question, predictive models for each region must be developed. Once the models are developed, statistical tests can be applied to determine the probability that the differences between the models are due to random sampling error. If the probability that the differences exist between the regional water demand models is small, the conclusion can be drawn that the models are similar within the confidence levels defined by the probability.

Developing comparable regional residential water demand predictive models requires the use of regression analysis. Ordinary least squares regression analysis will generate a mathematical relationship between the dependent variable (quantity of water demanded) and the independent variables (the 20 variables contained within the seven independent variable factor groups). However, the goal will be to generate the models that have the greatest amount of predictive capability. Predictive capability for a multivariant model can be measured by the coefficient of multiple determination (R-squared, or R^2).

In addition to the level of predictive capability for the models, each model should contain only those independent variables that are significant in contributing to the model. The method of determining significance for individual independent variables is to perform separate hypothesis tests for each variable with a predetermined level of error that will be used in the study. From the sample design, a .12 level of significance will be used throughout this study.

The use of both linear and exponential models of integrating the independent variables within a regression equation will be explored. The method that generates the models with the highest amount of predictability, within acceptable levels of significance, for the independent variables will be used to test for differences between the regional models. In order to avoid an incorrect conclusion from comparing only one set of regional models if they fail to demonstrate that they are statistically comparable, two additional sets of models will be compared.

The logic for this precautionary measure is that the specific set of regional demand models developed with the restriction of including only variables with a specific level of significance may, by design, exclude the model which would serve for all regions. Therefore, the test for significant differences between regional residential water demand models will be performed for a set of models as all independent variables included, as well as a set of models with only a limited number of variables included.

In estimating regional residential water demand models, care must be taken to insure that the models are statistically comparable. If each region's model is developed completely independently, the resulting set of regional models may not be comparable because the possiblity that they each contain a different mix of independent variables. Therefore, a generalized model solution must be found that contains the same mix of independent variables for all regions, while at the same time the models must achieve a high degree of predictability for each individual region. After constructing such a generalized model solution, the individual models can be statistically tested to determine if there are differences in the regional water demand schedules. If no significant differences exist, a single model can be used to describe the residential water demand for the entire state of Arkansas. However, if significant differences do exist between the generalized models for each region, then the logical approach would be to abandon the generalized model solution and develop regional models independently so that each region's residential water demand model approaches the maximum amount of predictive capability possible from the data set available.

The second objective of this study is to determine empirically various residential water demand elasticities. Elasticities are useful tools in policy making and should be derived from a model which is capable of explaining the interactions of the independent variables on the dependent variable. A predictive model, on the other hand, may or may not be useful in explaining the interaction of the independent variables on the dependent variable.

A multi-variant regression model derived by the use of the sum of the least squares will serve as a predictive model. If the coefficient of multiple determination is high, the model will have a high degree of prediction; if no multicollinearity exists between the independent variables in the regression model, the model will also serve as an explanatory model.

Multicollinearity, however, is a matter of degree rather than a question of absolutes. As Exhibits 3-11, 3-12, and 3-13 demonstrate, the vast majority of the variables are correlated to each other in some degree. The question becomes one of determining if multicollinearity exists to a sufficient degree to preclude the use of predictive model as an explanatory model. Or more specifically, at what level of multicollinearity does the use of a predictive model as an explanatory model become invalid? Although this question has not been completely answered in the statistical literature, a pragmatic approach is available.

Multicollinearity exists when one independent variable correlates with any other independent variable or with any linear combination of the other independent variables. The correlation coefficient tables presented in Chapter 3 demonstrate that multicollinearity exists since none of the individual cells have a zero value. However, the vast majority of the values are quite small, and therefore insignificant. A general rule for a pragmatic solution eliminates all values under the absolute value of .8 for a correlation coefficient between independent variables as insignificant with respect to high levels of multicollinearity within the regression model.

A careful examination of the correlation coefficient tables reveals that an absolute value of .8 or higher is found only when two variables from the same independent variable factor group are involved. For example, the coefficient of correlation between HOV and FIN (for all regions) is .97. This, of course, was entirely expected since the variable FIN is a linear combination which includes HOV (i.e., FIN = INC + HOV). Instead of removing one or the other of these variables, which could lower the predictive capability of the model, care will be exercised while including independent variables in the models to prevent both variables appearing in the same explanatory model.

After determining the model which provides the best predictive capability for each region, elasticities will be determined for each of the independent variables if the model can serve as an explanatory model as well as a predictive model. Since the only time this would not be the case is when an unacceptable level of multicollinearity is involved (because of the inclusion of independent variables from the same factor group), an additional step in the analysis is required conditionally. An examination of the best predictive model so that variables in the same factor group with high levels of correlation are not included, will suffice to convert the predictive models into explanatory models. Finally, it is recognized that the approach of using the correlation coefficients between the independent variables to exclude the possibility of unacceptable levels of multicollinearity implicitly assumes that a three or more linear correlation of independent variables will not exist without some indication at the two-variable level.

Statistical Analysis Procedures

This section is devoted to an explanation of the statistical analysis employed in this study and the accompanying results. The previous section presented the objectives of this study and a few of the statistical concepts to be employed. The remaining methods of analysis will be presented, followed immediately by the results of the procedures described. References will be made to the exhibits at the end of this chapter. The reader may find it useful to initially acquaint himself with the location and general contents of the exhibits, since frequent references will be made to these throughout this section.

Before beginning the statistical explanation and its accompanying analysis, an overview of this section is presented. The first objective of this study is to determine if the five watershed regions within Arkansas have similar residential water demand models. To accomplish this, the best available predictive models must be found. Therefore, a search is conducted to identify the best model form, for predictive purposes, using the given set of independent variables. After finding the model form which has the greatest predictive capability, a statistical test is described and performed to determine if the regional models are statistically different from each other. After that, the best predictive models are converted into explanatory models and elasticities of each independent variable in the regional models are determined. Finally, an explanation and comparison of the elasticities are presented.

Selection of Predictive Model Form

A procedure to determine the model form for combining the independent variables into a single mathematical expression that best provides predictive capabilities is direct correlated comparisons. Fortunately, the current statistical literature provides a measurement of a regression model's ability to predict. This measurement is the coefficient of determination (r^2), or the coefficient of multivariate determination (R-squared or R^2) if more than two independent variables are included in the regression model. This measurement is a ratio of the total amount of variation explained by the model to the total amount of variation of the dependent variable. When working with stochastic data, as in this study, the result of regression analysis is to provide a model which increases the probability to predict the value of an unobserved occurrence of the dependent variable. By using these statistics as the base level of predictability without the aid of a regression model, a ratio of the improved predictability with the aid of a regression model is available.[1]

Using the same set of independent variables in varied forms can yield different levels of predictability. By comparing these different models' predictive capabilities, the best available predictive model can be determined. The methods of combining the independent variables into regression models are restricted to three basic variations: linear, exponential and polynomial regression model forms. All three of these methods were explored for this study. The first two model forms, linear and exponential, were constructed, since the number of terms within these models with all independent variables is limited to simply the total number of variables in the model. However, a second order polynominal model with 19 independent variables would have 190 terms, while higher order polynomial models would have a substantially larger number of terms. Therefore, a method to predict the efficiency of polynomials were employed prior to the constructing involved models for comparable predictive purposes.

---

[1] John Neter and William Wasserman, <u>Applied Linear Statistical Models</u> (Homewood, ILL: Richard D. Irwin, Inc., 1974), page 228.

An accepted search method for locating possible polynomial relationships within the data set is to examine the graphical relationship between the residuals resulting from a linear or exponential model and each individual independent variable. If a polynomial model form is inherent within the data set, a curvilinear pattern will usually appear on the plot. Although this method can be employed relatively quickly with the aid of an electronic computer, the results are extremely lengthy to present since each independent variable requires a separate graph for each model (linear and exponential) and for each region, or a total of 684 separate graphs. Although the individual graphs are not presented in this paper, they were plotted and examined with no significant evidence to suggest that the use of polynomial model forms would improve the predictive capability of the linear and exponential forms of models.

Since the objective at this stage is to simply identify the model form which has the highest predictive capability, all independent variables that influenced the coefficient of multiple determination were included in the regression models for comparison purposes. Eighteen models were constructed for both the linear and the exponential forms, using a stepwise regression procedure that maximized the R-squared value.[2] The results of

---

[2]Each time a particular method of examining the data is explored, 18 different models must be generated. The state of Arkansas is divided into five watershed regions. When counting one model for the state, the total is six. These six models are then multiplied by three, one for each variable in the Price independent variable factor group, to yield a total of 18. The reason for finding a separate model for each region with the three different price variables lies in the fact that they are not proxies for each other. One of the three price variables will be included in a model depending upon the assumption made concerning the consumer's decision process as he evaluates his opportunities for acquiring goods and services. It is assumed that the consumer will either consider the average price he pays for water or a marginal price. In addition, the high amount of correlation between these variables requires the exclusion of all but one of the price variables to avoid the problem of multicollinearity as explained in the previous section. Therefore,

these models indicate that the exponential form is a superior method of combining the independent variables into a predictive model. In each individual case when the same price variable assumption is combined with a particular area, the exponential model produces a larger R-square value than the linear model. The mean R-square value for the 18 linear models is .3753, compared to a mean R-square value of .5962 for the exponential models. The .2209 difference between these mean values, represents a 58 percent increase in predictive capability available by using the exponential model form over the linear model form. Overall, eighteen exponential models were estimated (one for each of the six regions and three price variables) but the models utilizing the average price variable provided the highest predictive capability. The results of these models are reported in Exhibits 4-1 through 4-6.

Comparison of Regional Models

In the previous section it was demonstrated that the exponential regression model provided the best means of predicting water demand from the set of independent variables used in this study. The question that must be addressed now is whether there are any significant differences between the regional regression models in exponential form. A method of approaching this question is to compare the results of the individual regional regression exponential models to a combined state regression exponential model. By comparing separate regression models for each region and the entire state it is possible to determine if the state regression

---

for both statistical reasons and because of a-priori assumptions, one and only one price variable will be included in the models at a time. However, the selection of the appropriate price variable to use can be approached more efficiently after the empirical data has been analyzed to assist in the decision process. For these reasons, the inclusion of 18 models for each method to be employed should be expected by the reader.

model would serve as well as the five separate regional regression models for predictive purposes. Or stated differently--do the regression coefficients of the regional models represent significant differences at a given level of confidence?

To address this question the null hypothesis can be stated that a single regression model is adequate for the individual regions. That is, a single coefficient can be found for each independent variable so that the regional variable coefficients represent only insignificant differences.

Can β_{S1} be determined such that:

$$\beta_{S1} \simeq \beta_{11}, \beta_{S1} \simeq \beta_{21} \ldots \beta_{S1} \simeq \beta_{51}$$

and

can β_{S2} be determined such that:

$$\beta_{S2} \simeq \beta_{12}, \beta_{S2} \simeq \beta_{22} \ldots \beta_{S2} \simeq \beta_{52}$$

and

can β_{S3} be determined such that:

$$\beta_{S2} \simeq \beta_{13}, \quad \beta_{S3} \simeq \beta_{23} \ldots \beta_{S5} \simeq \beta_{53}$$

etc.

Where: β is a coefficient identified by subscripts where the first subscript denotes the area (S = state and the numbers correspond to the regions) and the second subscript represents the variable that this coefficient modifies.

Note the approximation signs (not equal signs) between the coefficients to denote that the coefficients are equivalent at some level of significance.

The procedure that accomplishes this test relies on the amount of error that the individual regression models have by composing a F statistic test which can be compared to a F value from the predetermined F distribution at n specific level of confidence for the associated levels of degrees of freedom[3]. The algebraic test is:

---

[3]James E. Dunn, "Methods of Multivariate Analysis," (unpublished manuscript), p. 58.

$$F_{test} = \frac{(SSE_R - SSE_F) \div (V_r - V_f)}{(SSE_F) \div V_f}$$

where:

SSE_R = Square sum of the errors for the state model

SSE_F = Square sum of the errors for all the regional models

V_r = Degrees of freedom for the state model $V_f = n_s - p_s$

V_f = Degrees of freedom for the regional models or $V_r = V_1 + V_2 + V_3 + V_4 + V_5$ where $V_1 = n_1 - p_1$, where n is the number of observations and p is the number of estimated regression coefficients.

After the F_{test} has been calculated, it can be compared to the $F_{statistic}$ with the appropriate parameters to conclude the procedure with the following decision rule:

Reject the null hypothesis if:

$$F_{test} > F_{statistic}$$

where:

$F_{statistic} = F(V_r - V_f;\ V_f;\ 1-\alpha)$, determined from the F distribution.

If the null hypothesis is rejected, then individual regional regression models are statistically different from each other, and the state regression model would not provide an efficient approximation for the individual watershed regions in predicting residential water demand.

To calculate this test, it is imperative that all the regression models used have the same set of independent variables. This requirement is satisfied if all the independent variables available (except the duplicated price variables) are used in each area's equation. The F_{test} is calculated for each of the three price variables. The resulting values are 8.76 for the model which includes average price as the independent variable and 6.82 and 1.42 for the marginal and incremental price models respectively.

When the values of the F_{test} are compared to the appropriate $F_{statistic}$ ($F_{(80; 1667; .88)} = 1.32$), the null hypothesis is rejected at the 88 percent confidence level. That is, using an alpha level of .12 (determined from Chapter 3), the single exponential regression model for the state with 19 independent variables provides statistically significant different estimates for the coefficients than the individually calculated regional regression models.

The possibility still exists that a less specific exponential regression model constructed for the state using a fewer number of independent variables could be statistically acceptable in explaining the individual regional residential water demand models. Therefore, a set of regression models will be constructed using the three most constantly significant variables in predicting water demand. Based on the average order of entry into the stepwise regression, these variables include the different measures of price, total number of water faucets, and town size. Again, three sets of six models each must be derived so that each price variable is entered for the various price related assumptions. Following the same testing procedure as above the values of the F_{tests} are 20.44, 14.63, and 11.13 for the average, marginal, and incremental price variable models respectively. All three F_{tests} are larger than their corresponding $F_{statistics}$ determined from the F distribution ($F_{(12, 1752; .88) = .88)} = 1.33$) and therefore, the null hypothesis in each case is rejected. This means that even by reducing the exactness of the models, and thereby forfeiting some predictability, the regional models are statistically significant as regional predictors, compared to a state regression model.

Before concluding that independently developed regional models are significantly different from a state regression model, independently developed, one final approach will be employed. Although models with all independent variables included and with only three independent variables included have been demonstrated to be significantly different from the state regression model similarly developed with respect to their estimation of the parameter coefficients, the possibility remains that some inbetween number of variables could prove to be statistically insignificant in the state model as an approximation of the regional models. To construct such a model, the number and selection of included variables must be addressed so that a generalized state model would be an approximation of the mean of the various contributory variables in the regional model. To accomplish this task, the order in which the variables are entered into the model by a stepwise regression procedure will be considered, along with the adjusted coefficient of multiple determination and the $F_{statistic}$ for each independent variable.

By employing a stepwise regression procedure which maximizes the R-squared value, each independent variable is entered into the model according to its contribution to the model's R-squared value. Using the variable's entry order number as the basic means of constructing an average optimal state model, and then considering the other two factors when needed, a composite model will be developed and then tested.

Since the coefficient of multiple determination can only increase or remain constant by the adding of independent variables to the regression model, the adjusted coefficient of multiple determination is useful when examining the relationship between the number of variables in the model and the amount of total variation in the dependent variable, which is explained by the interacting of the independent variables.

The adjusted coefficient of multiple determination accounts for the number of independent variables used in the model which reduces the degrees of freedom it has. If an additional variable does not add to the predictive capability of the model enough to offset the loss of a degree of freedom, the value of the adjusted coefficient of multiple determination will decline. The value of the adjusted coefficient of multiple determination can be determined by:

$$R_a^2 = 1 - \frac{(n - 1)}{(n - p)} \frac{SSE}{SSTO}$$

where:

R_a^2 = adjusted coefficient of multiple determination

SSE = error sum of squares, or $\Sigma(Yi - \hat{Y})^2$

SSTO = total sum of squares, or $\Sigma(Y_1 - \overline{Y})^2$

n = number of observations

p = number of independent variables in the regression model

A two-dimensional graph of the adjusted coefficient of multiple determination plotted against the number of independent variables used in the model can be a useful tool in selecting the optimal number of independent variables to include in a model. Assuming the independent variables are added in order of the increasing value of the R^2, the abovementioned plot will indicate the optimal number of independent variables. When the marginal value of the adjusted coefficient of multiple determination reaches zero, the associated mix of independent variables is optimal in statistical terms, given the set of possible independent variables from which to select.

The third factor to consider while trying to construct a generalized state model is the level of significance of each independent variable. If an individual independent variable has a low level of confidence associated with its estimated coefficient, its usefulness as a predictor must be questioned. By identifying the variables that have an alpha level of more than .12 associated with their estimated coefficient, a third criterion for compiling a composite model is identified.

Based on the three criteria outlined above a generalized state model is estimated by first averaging each independent variable's entry order number in the stepwise regression in order to provide an array of the most common entry order. The number of variables to include in the generalized model is determined by averaging the total number of variables in each regional model that entered solution with a positive R_a^2 value and were significant at the .12 alpha level.

An exponential regression model was computed for the state and each AWRPA for each separate price variable. The results of these models are presented in Exhibits 4-8, 4-9, 4-10. The F_{tests} for the average, marginal, and incremental price models are 13.45, 10.04, and 8.24 respectively. As before, the above F_{tests} are larger than their corresponding $F_{statistics}$ determined from the F distribution, and therefore the null hypothesis is rejected in each case. At this time a conclusion can be drawn that the individually developed regional residential water demand models are significantly different with respect to the estimates of their coefficients than the state residential water demand model.

Development of Explanatory Models

The methods explained in the preceding section provide the necessary tools for developing individual regional models for residential water demand for predictive purposes. Explanatory models can be estimated by simply adjusting these models as necessary to account for problems of multicollinearity. By specifying the variables in exponential form the estimated regression coefficients are also the elasticities of the variables they modify with respect to water. The estimated explanatory models, specified in terms of natural logs, are presented in Exhibit 4-11. Note that the resulting regional explanatory residential water demand models do not contain the same set of independent variables. This result is fully expected, based on discussions in the preceding sections. The results and usefulness of these models will be discussed in the next chapter.

EXHIBIT 4-1

AVERAGE PRICE EXPONENTIAL STEPWISE REGRESSION RESULTS
FOR
STATE

The following exhibit is a summary of the equations generated by the stepwise regression using the maximum R-squared improvement technique. The variable AVP was the only variable from the Price Factor Group included in the data set. The variable values were transformed into their natural logrithmic form.

Varible	Entry Number	Variable F_{value}	Variable F_{prob}	R^2	R_a^2	ΔR_a^2	Coef-ficient
AVP	1	894.6	.0001	.4180	.4180		-1.080
PPH	2	24.6	.0001	.4995	.4989	.0809	.221
FIN	3	19.6	.0001	.5002	.4996	.0007	.225
PRE	4	16.8	.0001	.5158	.5147	.0151	.086
PUR	5	11.5	.0007	.5189	.5176	.0029	- .030
TOW	6	8.9	.0031	.5214	.5197	.0021	- .045
APP	7	4.6	.0326	.5231	.5212	.0015	- .041
AWA	8	2.8	.0964	.5247	.5226	.0014	.030
FAU	9	1.7	.1937	.5248	.5227	.0001	.154
FIX	10	3.5	.0603	.5256	.5232	.0005	.197
TEM	11	2.9	.1401	.5263	.5236	.0004	.486
HOV	12	1.5	.2235	.5271	.5239	.0003	- .016
AGE	13	4.0	.0468	.5275	.5240	.0001	.093
CHI	14	2.9	.0886	.5278	.5243	.0003	.004
SWI	15	1.5	.2227	.5282	.5244	.0001	- .003
INC	16	.7	.4088	.5283	.5243	.0001	.036
GAR	17	.4	.5248	.5285	.5241	-.0002	.001
LAW	18	.0	.9047	.5285	.5239	-.0002	- .001

The following statistics were computed for the last equation calculated by the stepwise procedure:

SSE = 504.614 F_{value} = 108.85

n = 1767 F_{prob} = .0001

EXHIBIT 4-2

AVERAGE PRICE EXPONENTIAL STEPWISE REGRESSION RESULTS
FOR
REGION 1

The following exhibit is a summary of the equations generated by the stepwise regression using the maximum R-squared improvement technique. The variable AVP was the only variable from the Price Factor Group included in the data set. The variable values were transformed into their natural logrithmic form.

Variable	Entry Number	Variable F_{value}	Variable F_{prob}	R^2	R^2_a	ΔR^2_a	Coef-ficient
AVP	1	111.4	.0001	.3302	.3302		- .854
TOW	2	74.9	.0001	.4166	.4138	.0836	.635
TEM	3	24.8	.0001	.4192	.4379	.0241	2.645
PRE	4	17.4	.0001	.4677	.4626	.0247	.281
PUR	5	18.1	.0001	.4839	.4777	.0151	.519
FIN	6	14.8	.0001	.4883	.4822	.0045	2.88
CHI	7	6.9	.0091	.4961	.4888	.0066	.014
AGE	8	6.1	.0138	.5087	.505	.0117	.232
HOV	9	13.6	.0001	.5182	.5077	.0072	-1.954
INC	10	10.0	.0017	.5203	.5099	.0022	- .668
LAW	11	5.1	.0248	.5262	.5148	.0049	.035
APP	12	7.0	.0084	.5344	.5196	.0048	- .108
AWA	13	6.0	.0150	.5371	.5224	.0028	.094
FAU	14	1.1	.3053	.5381	.5234	.0010	.291
PPH	15	2.7	.1023	.5413	.5256	.0022	.172
SWI	16	1.1	.2888	.5423	.5255	-.0001	.010
GAR	17	.3	.5590	.5428	.5248	-.0007	- .001
FIX	18	.0	.9572	.5428	.5236	-.0012	- .013

The following statistics were computed for the last equation calculated by the stepwise procedure:

SSE = 114.619 F_{value} = 26.72

n = 424 F_{prob} = .001

EXHIBIT 4-3

AVERAGE PRICE EXPONENTIAL STEPWISE REGRESSION RESULTS
FOR
REGION 2

The following exhibit is a summary of the equations generated by the stepwise regression using the maximum R-squared improvement technique. The variable AVP was the only variable from the Price Factor Group included in the data set. The variable values were transformed into their natural logrithmic form.

Variable	Entry Number	Variable F_{value}	Variable F_{prob}	R^2	R^2_a	ΔR^2_a	Coef-ficient
AVP	1	423.1	.0001	.4965	.4965		-1.615
TOW	2	112.7	.0001	.6205	.6191	.1226	.501
FIN	3	14.8	.0002	.6691	.6667	.0476	.486
FAU	4	4.8	.0295	.7121	.7079	.0412	.596
PUR	5	2.9	.0872	.7195	.7133	.0054	- .188
PPH	6	2.2	.1433	.7249	.7157	.0024	.136
CHI	7	2.8	.0944	.7251	.7158	.0001	- .009
AGE	8	3.0	.0831	.7252	.7160	.0002	- .176
APP	9	.5	.4640	.7255	.7163	.0003	- .029
SWI	10	2.0	.1564	.7272	.7170	.0007	- .009
LAW	11	2.7	.1020	.7303	.7181	.0011	- .015
INC	12	1.5	.2225	.7314	.7182	.0001	- .125
HOV	13	1.9	.1610	.7316	.7184	.0002	- .019
FIX	14	1.5	.2182	.7333	.7191	.0007	- .3127
GAR	15	.2	.6371	.7335	.7183	-.0008	.001
AWA	16	.1	.7043	.7337	.7173	-.0010	.014
PRE	17	.0	1.0000	.7337	.7163	-.0010	.275
TEM	18	.0	1.0000	.7337	.7152	-.0011	- .089

The following statistics were computed for the last equation calculated by the stepwise procedure:

SSE = 52.996 $\qquad F_{value}$ = 37.57

n = 279 $\qquad F_{prob}$ = .0001

EXHIBIT 4-4

AVERAGE PRICE EXPONENTIAL STEPWISE REGRESSION RESULTS
FOR
REGION 3

The following exhibit is a summary of the equations generated by the stepwise regression using the maximum R-squared improvement technique. The variable AVP was the only variable from the Price Factor Group included in the data set. The variable values were transformed into their natural logrithmic form.

Variable	Entry Number	Variable F_{value}	Variable F_{prob}	R^2	R_a^2	ΔR_a^2	Coef-ficient
AVP	1	137.4	.0001	.3116	.3116		-1.035
PPH	2	17.6	.0001	.4634	.4619	.1503	.430
LAW	3	9.2	.0026	.5465	.5426	.0807	.079
FAU	4	5.6	.0181	.5502	.5446	.0020	.574
SWI	5	5.3	.0223	.5683	.5633	.0187	.017
INC	6	5.7	.0173	.5905	.5834	.0201	.275
APP	7	8.7	.0034	.5972	.5890	.0056	- .116
AWA	8	6.3	.0128	.6050	.5958	.0068	.092
GAR	9	2.5	.1160	.6080	.5977	.0019	.004
PUR	10	1.3	.2523	.6095	.5981	.0004	- .011
CHI	11	.7	.4202	.6108	.5983	.0002	.004
FIN	12	.3	.6040	.6108	.5983	.0000	- .161
TEM	13	.7	.4128	.6114	.5977	-.0006	- .579
TOW	14	.4	.5251	.6118	.5969	-.0008	- .058
HOV	15	.2	.6570	.6120	.5959	-.0010	- .090
PRE	16	.1	.7311	.6122	.5949	-.0010	.012
FIX	17	.1	.7616	.6123	.5938	-.0011	- .063
AGE	18	.0	.9954	.6123	.5926	-.0012	- .001

The following statistics were computed for the last equation calculated by the stepwise procedure:

SSE = 62.754 $\quad F_{value}$ = 29.39

n = 354 $\quad F_{prob}$ = .0001

EXHIBIT 4-5

AVERAGE PRICE EXPONENTIAL STEPWISE REGRESSION RESULTS FOR REGION 4

The following exhibit is a summary of the equations generated by the stepwise regression using the maximum R-squared improvement technique. The variable AVP was the only variable from the Price Factor Group included in the data set. The variable values were transformed into their natural logrithmic form.

Variable	Entry Number	Variable F_{value}	Variable F_{prob}	R^2	R_a^2	ΔR_a^2	Coef-ficient
AVP	1	378.0	.0001	.3700	.3700		-1.565
TOW	2	103.1	.0001	.4788	.4777	.1077	- .255
FIN	3	5.8	.0169	.5699	.5681	.0904	1.395
LAW	4	47.2	.0001	.5940	.5914	.0233	- .079
PPH	5	20.0	.0001	.6250	.6217	.0303	.338
FAU	6	8.0	.0048	.6405	.6367	.0150	.569
TEM	7	7.6	.0062	.6472	.6426	.0059	-1.964
AGE	8	1.4	.2354	.6489	.6436	.0010	.092
APP	9	.6	.4526	.6507	.6446	.0010	- .023
INC	10	4.3	.0389	.6527	.6452	.0006	- .382
HOV	11	3.4	.0655	.6544	.6468	.0006	- .740
PUR	12	.8	.3759	.6552	.6469	.0001	- .073
GAR	13	1.3	.2633	.6559	.6469	.0000	- .002
FIX	14	.3	.5716	.6565	.6460	-.009	- .099
AWA	15	.2	.6613	.6567	.6454	-.0006	.013
CHI	16	.1	.7574	.6568	.6447	-.0007	.001
SWI	17	.0	.9231	.6568	.6439	-.0008	.001

The following statistics were computed for the last equation calculated by the stepwise procedure:

SSE = 89.873 $\quad F_{value}$ = 48.07

n = 471 $\quad F_{prob}$ = .0001

EXHIBIT 4-6

AVERAGE PRICE EXPONENTIAL STEPWISE REGRESSION RESULTS FOR REGION 5

The following exhibit is a summary of the equations generated by the stepwise regression using the maximum R-squared improvement technique. The variable AVP was the only variable from the Price Factor Group included in the data set. The variable values were transformed into their natural logrithmic form.

Variable	Entry Number	Variable F_{value}	Variable F_{prob}	R^2	R_a^2	ΔR_a^2	Coef-ficient
AVP	1	255.7	.0001	.7122	.7122		-1.285
SWI	2	14.4	.0002	.7566	.7546	.0424	- .036
FAU	3	5.8	.0168	.7590	.7569	.0023	.692
AWA	4	2.0	.1569	.7741	.7702	.0133	- .051
TEM	5	2.4	.1213	.7781	.7733	.0031	1.311
PUR	6	1.8	.1817	.7842	.7776	.0043	- .017
CHI	7	1.7	.1925	.7864	.7779	.0003	.007
GAR	8	1.1	.2986	.7871	.7787	.0008	.003
LAW	9	1.0	.3314	.7879	.7786	-.0001	.039
APP	10	.9	.3470	.7883	.7780	-.0006	.037
AGE	11	.8	.3757	.7885	.7782	.0002	.103
PPH	12	.8	.3697	.7891	.7778	-.0004	.081
FIN	13	.2	.6996	.7896	.7774	-.0004	- .399
HOV	14	.1	.7420	.7897	.7775	.0001	- .249
FIX	15	.3	.5898	.7900	.7768	-.0007	- .122
TOW	16	.1	.7047	.7901	.7760	-.0008	.019
HOV	17	.0	.8380	.7902	.7750	-.0010	.062
PRE	18	.0	.8995	.7902	.7790	-.0010	.013

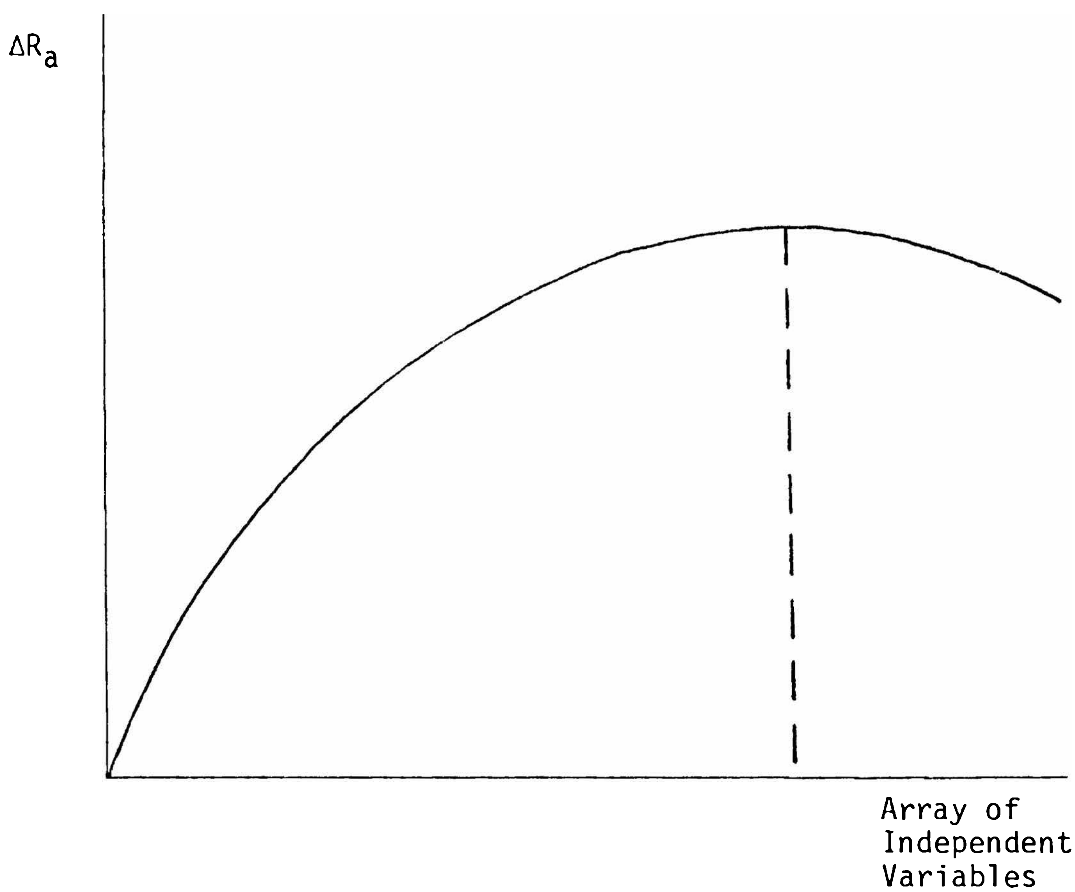
The following statistics were computed for the last equation calculated by the stepwise procedure:

SSE = 35.009 $\quad F_{value}$ = 46.05

n = 239 $\quad F_{prob}$ = .0001

EXHIBIT 4-7

RELATIONSHIP BETWEEN THE ADJUSTED COEFFICIENT OF MULTIPLE DETERMINATION AND THE NUMBER OF INDEPENDENT VARIABLES ENTERED BY A STEPWISE REGRESSION PROGRAM



The apex of this curve locates the optimal mix of independent variables (measured on the abscissa) when the variables are entered by a maximum R^2 improvement stepwise regression technique.

EXHIBIT 4-8

AVERAGE PRICE GENERALIZED EXPONENTIAL REGRESSION MODELS

This table presents the results of least-squared multiple regressoin methods with the nine independent variables which have been determined to most likely produce a generalized solution for the separate regions. Only the variable AVP was included from the Price Factor Group, while all variable values are in their natural logrithmic form.

Model Number	55	56	57	58	59	60
Area	State	Region 1	Region 2	Region 3	Region 4	Region 5
R^2	.5192	.4809	.6883	.5737	.6499	.7667
Regression F_{value}	210.85	42.63	66.01	51.45	95.10	83.67
n	1767	424	279	354	471	239
SSE	514.523	130.126	62.036	68.991	91.679	38.922

Variables	Coefficient Values					
a**	3.246	-11.833	6.882*	10.642	23.449	- .643*
AVP	-1.071	- .893	-1.520	- 1.002	- 1.543	-1.168
FIN	.269	.287	.359	- .096*	.321	.118*
PPH	.219	.181	.061*	.559	.268	.198
TEM	.510*	2.412	- .202*	- .430*	- 2.029	1.276*
APP	- .012	- .004*	- .017	- .002	- .005*	- .010
FAU	.354	.230	.358	.538	.411	.511
LAW	.001*	.002	.002*	.066	- .079	- .062*
PUR	- .028	- .489	.196	- .023	- .096*	- .013*
TOW	- .045	.447	.477	- .113*	- .249	.028*

* Variable is not significant at an alpha level of .12.

** a = Intercept term

EXHIBIT 4-9

MARGINAL PRICE GENERALIZED EXPONENTIAL REGRESSION MODELS

This table presents the results of least-squared multiple regression methods with the nine independent variables which have been determined to most likely produce a generalized solution for the separate regions. Only the variable MRP was included from the Price Factor Group, while all variable values are in their natural logrighmic form.

Model Number	61	62	63	64	65	66
Area	State	Region 1	Region 2	Region 3	Region 4	Region 5
R^2	.5046	.4628	.7212	.6088	.5160	.7026
Regression F_{value}	199.20	39.83	77.35	59.50	54.73	60.14
n	1770	426	279	354	472	239
SSE	530.324	134.714	55.479	63.311	126.827	49.621

Variables	Coefficient Values					
a**	5.357	2.668	4.051	8.060	9.206	5.041
MRP	- .990	- .902	-1.342	-1.131	-1.513	- .865
FIN	.271	.236	.029	.054*	.357	.193
PPH	.263	.228	.014*	.584*	.396	.249
PRE	.096	.202	.287	.021	.071	- .030*
APP	- .010	- .005*	- .012	- .022	.002*	- .017
FAU	.400	.436	.404	.474	.359	.589
LAW	- .007*	.012*	- .010*	.048	- .093	- .045*
PUR	- .026	- .391	- .253	- .016	- .137*	- .012*
TOW	- .049	.415	.320	- .151	- .299	.148

* Variable is not significant at an alpha level of .12.

** a = Intercept term

EXHIBIT 4-10

INCREMENTAL PRICE GENERALIZED EXPONENTIAL REGRESSION MODELS

This table presents the results of least-squared multiple regression methods with the nine independent variables which have been determined to most likely produce a generalized solution for the separate regions. Only the variable INP was included from the Price Factor Group, while all variable values are in their natural logrithmic form.

Model Number	67	68	69	70	71	72
Area	State	Region 1	Region 2	Region 3	Region 4	Region 5
R^2	.4017	.3612	.5734	.5512	.4621	.6214
Regression F_{value}	125.49	25.20	26.60	45.05	43.35	38.67
n	1692	411	255	340	464	222
SSE	501.668	134.089	53.221	58.090	124.423	46.418

Variables	Coefficient Values					
a**	4.727	2.069	3.678	7.267	9.339	4.773
INP	- .875	- .828	-1.347	-1.095	-1.584	- .779
INC	.059*	.003*	- .144	.321	- .033*	.089*
FIN	.253	.219	.397	- .298	.403	.178*
PPH	.239	.194	- .007	.513	.400	.217
PRE	.089	.210	.314*	.018*	.071	- .092*
FAU	.410	.464	.399	.532	.324	.527
LAW	- .009*	.015*	- .009*	.065	- .094	- .044*
TOW	.050	.348	.312	- .040*	- .315	.168
AWA	- .011	- .003*	- .011	- .014	.001*	- .019

* Variable is not significant at an alpha level of .12.

** a = Intercept term

EXHIBIT 4-11

EXPLANATORY EXPONENTIAL MODELS

Using Average Variable Price

State

Natural logrithmic model for the State using AVP as the price variable.

$$Q = 6.659 \ln - 1.074 \ln AVP + .269 \ln FIN + .219 \ln PPH - .012 \ln APP$$
(.0001) (.0001) (.0001) (.0001) (.0001)

$$+ .353 \ln FAU - .028 \ln PUR - .043 \ln TOW$$
(.0001) (.0015) (.0041)

R^2 = .5185 $\quad$ F_{value} = 270.70

n = 1767 $\quad$ F_{prob} = .0001

Region 1

Natural logrithmic model for Region 1 using AVP as the price variable.

$$Q = -18.069 \ln - .929 \ln AVP - 1.640 \ln HOV + 2.644 \ln FIN + .277 \ln AGE$$
(.0001) (.0001) (.0018) (.0003) (.0012)

$$+ .019 \ln CHI + 2.68 \ln TEM = .291 \ln PRE - .558 \ln PUR + .575 \ln TOW$$
(.0001) (.0001) (.0001) (.0001) (.0001)

R^2 = .5203 $\quad$ F_{value} = 44.81

n = 424 $\quad$ F_{prob} = .0001

Region 2

Natural logrithmic model for Region 2 using AVP as the price variable.

$$Q = 5.749 \ln - 1.562 \ln AVP + .443 \ln FIN + .472 \ln TOW$$
(.0001) (.0001) (.0001) (.0001)

R^2 = .6691 $\quad$ F_{value} = 185.38

n = 279 $\quad$ F_{prob} = .0001

Region 3

Natural logrithmic model for Region 3 using AVP as the price variable.

$$Q = 6.410 \ln - 972 \ln AVP + .218 \ln INC + .496 \ln PPH - .092 \ln APP$$
(.0001) (.0001) (.0013) (.0001) (.0083)

EXHIBIT 4-11 Continued

$$+ .335 \ln FAU + .021 \ln SWI + .051 \ln LAW + .005 \ln GAR + .066 \ln AWA$$
(.0016) (.0001) (.0001) (.0263) (.0449)

$R^2 = .5986$ $F_{value} = 57.02$

$n = 354$ $F_{prob} = .0001$

Region 4

Natural logrithmic model for Region 4 using AVP as the price variable.

$$Q = 29.421 \ln - 1.354 \ln AVP + .160 \ln INC + .323 \ln PPH - 3.161 \ln TEM$$
(.0001) (.0001) (.0168) (.0001) (.0001)

$$+ .513 \ln FAU - .060 \ln LAW$$
(.0001) (.0001)

$R^2 = .5519$ $F_{value} = 95.28$

$n = 471$ $F_{prob} = .0001$

Region 5

Natural logrithmic model for Region 5 using AVP as the price variable.

$$Q = 8.207 \ln - 1.25 \ln AVP + .541 \ln FAU - .040 \ln SWI$$
(0001) (.0001) (.0001) (.0001)

$R^2 = .7590$ $F_{value} = 246.77$

$n = 239$ $F_{prob} = .0001$

Using Marginal Price

State

Natural logrithmic model for the State using MRP as the price variable.

$$Q = 5.187 \ln - .977 \ln MRP + .235 \ln FIN + .293 \ln PPH + .089 \ln PRE$$
(.0001) (.0001) (.0001) (.0001) (.0001)

$$- .064 \ln APP + .411 \ln FAU - .028 \ln PUR + .051 \ln AWA$$
(.0003) (.0001) (.00019) (.0022)

$R^2 = .5042$ $F_{value} = 223.89$

$n = 1770$ $F_{prob} = .0001$

EXHIBIT 4-11 Continued

Region 1

Natural logrithmic model for Region 1 using MRP as the price variable.

$$Q = \underset{(.0001)}{15.937 \ln} - \underset{(.0001)}{.871 \ln MRP} - \underset{(.0022)}{1.696 \ln HOV} + \underset{(.0004)}{2.743 \ln FIN} + \underset{(.0732)}{.188 \ln PPH}$$

$$+ \underset{(.0016)}{.303 \ln AGE} + \underset{(.0001)}{2.182 \ln TEM} + \underset{(.0007)}{.236 \ln PRE} - \underset{(.0985)}{.068 \ln APP}$$

$$- \underset{(.0001)}{.478 \ln PUR} + \underset{(.0001)}{.591 \ln TOW} + \underset{(.1212)}{.061 \ln AWA}$$

$R^2 = .5026$ $\quad F_{value} = 32.03$

$n = 426$ $\quad F_{prob} = .0001$

Region 2

Natural logrithmic model for Region 2 using MRP as the price variable.

$$Q = \underset{(.0001)}{6.307 \ln} - \underset{(.0001)}{1.346 \ln MRP} + \underset{(.0026)}{.246 \ln FIN} + \underset{(.0012)}{.337 \ln FAU} - \underset{(.0100)}{.269 \ln PUR}$$

$$+ \underset{(.0001)}{.309 \ln TOW}$$

$R^2 = .7017$ $\quad F_{value} = 128.46$

$n = 279$ $\quad F_{prob} = .0001$

Region 3

Natural logrithmic model for Region 3 using MRP as the price variable.

$$Q = \underset{(.0001)}{6.630 \ln} - \underset{(.0001)}{1.105 \ln MRP} + \underset{(.0001)}{.250 \ln INC} + \underset{(.0001)}{.515 \ln PPH} - \underset{(.0085)}{.088 \ln APP}$$

$$+ \underset{(.0004)}{.303 \ln FAU} + \underset{(.0184)}{.015 \ln SWI} + \underset{(.0060)}{.033 \ln LAW} + \underset{(.0218)}{.005 \ln GAR} + \underset{(.0497)}{.062 \ln AWA}$$

$R^2 = .6321$ $\quad F_{value} = 65.69$

$n = 354$ $\quad F_{prob} = .0001$

Region 4

Natural logrithmic model for Region 4 using MRP as the price variable.

EXHIBIT 4-11 Continued

$$Q = 10.615 \ln - 1.523 \ln MRP + .418 \ln PPH + .084 \ln PRE + .615 \ln FAU$$
(.0001) (.0001) (.0001) (.0035) (.0001)

$$- .098 \ln LAW - .303 \ln TOW$$
(.0001) (.0001)

$R^2 = .4951$ $F_{value} = 76.03$

$n = 472$ $F_{prob} = .0001$

Region 5

Natural logrithmic model for Region 5 using MRP as the price variable.

$$Q = 5.244 \ln - .957 \ln MRP + .654 \ln FAU - .050 \ln GAR + .183 \ln TOW$$
(.0001) (.0001) (.0001) (.0001) (.0002)

$R^2 = .6935$ $F_{value} = 132.37$

$n = 239$ $F_{prob} = .0001$

Using Incremental Price

State

Natural logrithmic model for the State using INP as the price variable.

$$Q = 6.059 \ln - .908 \ln INP + .278 \ln PPH + .095 \ln PRE + .634 \ln FAU$$
(.0001) (.0001) (.0001) (.0001) (.0001)

$$- .039 \ln TOW - .011 \ln AWA$$
(.0128) (.0001)

$R^2 = .3850$ $F_{value} = 175.83$

$n = 1692$ $F_{prob} = .0001$

Region 1

Natural logrithmic model for Region 1 using INP as the price variable.

$$Q = -16.539 \ln - .739 \ln INP - 1.661 \ln HOV + 2.453 \ln FIN + .265 \ln AGE$$
(.0001) (.0001) (.0022) (.0013) (.0052)

$$+ .022 \ln CHI + 2.289 \ln TEM + .237 \ln PRE - .107 \ln APP + .490 \ln FAU$$
(.0001) (.0001) (.0006) (.0108) (.0001)

$$- .376 \ln TEM + .544 \ln TOW + .091 \ln AWA$$
(.0027) (.0001) (.0217)

EXHIBIT 4-11 Continued

R^2 = .4486 F_{value} = 24.85

n = 411 F_{prob} = .0001

Region 2

Natural logrithmic model for Region 2 using INP as the price variable.

$$Q = \underset{(.0001)}{6.350\ ln} - \underset{(.0001)}{1.363\ ln\ INP} + \underset{(.0050)}{.259\ ln\ FIN} + \underset{(.0016)}{.349\ ln\ FAU} - \underset{(.0123)}{.274\ ln\ PUR} + \underset{(.0001)}{.300\ ln\ TOW}$$

R^2 = .5467 F_{value} = 60.07

n = 255 F_{prob} = .0001

Region 3

Natural logrithmic model for Region 3 using INP as the price variable.

$$Q = \underset{(.0001)}{5.970\ ln} - \underset{(.0001)}{.993\ ln\ INP} + \underset{(.0001)}{.303\ ln\ INC} + \underset{(.0001)}{.517\ ln\ PPH} - \underset{(.0051)}{.094\ ln\ APP} + \underset{(.0051)}{.304\ ln\ FAU} + \underset{(.0845)}{.010\ ln\ SWI} + \underset{(.0021)}{.007\ ln\ GAR} + \underset{(.0359)}{.066\ ln\ AWA}$$

R^2 = .5606 F_{value} = 52.01

n = 340 F_{prob} = .0001

Region 4

Natural logrithmic model for Region 4 using INP as the price variable.

$$Q = \underset{(.0001)}{10.966\ ln} - \underset{(.0001)}{1.578\ ln\ INP} + \underset{(.0001)}{.407\ ln\ PPH} + \underset{(.0062)}{.080\ ln\ PRE} + \underset{(.0001)}{.606\ ln\ FAU} - \underset{(.0001)}{.098\ ln\ LAW} - \underset{(.0001)}{.313\ ln\ TOW}$$

R^2 = .4375 F_{value} = 59.26

n = 464 F_{prob} = .0001

EXHIBIT 4-11 Continued

Region 5

Natural logrithmic model for Region 5 using INP as the price variable.

$$Q = \underset{(.1074)}{-10.630 \ln} - \underset{(.0001)}{.798 \ln INP} + \underset{(.0011)}{.293 \ln HOV} + \underset{(.0065)}{.010 \ln CHI} + \underset{(.0313)}{2.13 \ln TEM}$$

$$+ \underset{(.0313)}{.427 \ln FAU} - \underset{(.0001)}{.051 \ln SWI} + \underset{(.0187)}{.130 \ln TOW}$$

$R^2 = .6446$ $\quad$ $F_{value} = 55.47$

$n = 222$ $\quad$ $F_{prob} = .0001$

CHAPTER 5

SUMMARY AND CONCLUSIONS

Introduction

Water is required not only to sustain life, but also satisfy wants which generally enhance the quality of life. Because it is used by households, industries, and commercial activities to satisfy a variety of demands its quality, availability, and distribution are important considerations in water resource management. This study is concerned with one aspect of this management, i.e., demand, for one type of user, i.e., households. Previously, it was assumed that the demand for water was very inelastic for individual households and, consequently, that varying the price of water was not a very effective way to ration it. Over the past twenty years, however, studies which have investigated the assumption of inelastic water demand have found repeatedly found that the evidence simply does not support such a position.

This study contributes to previous efforts and adds several new insights to the existing body of knowledge. It reports the estimates of regional residential water demand models for each of the five Arkansas Water Resource Planning Areas within the state. These models can be used to forecast future residential water demand for the state as well as to estimate the effects of different rate policies on water demand and water system revenues. Additionally, the techniques employed in this study of comparing similarly developed regional residential water demand models provide a means of reconciling other research efforts. Each of these contributions will be discussed in the following two sections of this chapter. The results of this study cannot be stated in deterministic terms since the methods employed to derive them relied heavily upon statistical techniques which, at best, can provide only stochastic relationships. However, the type of problem and the nature of the data within this study are best approached through the use of statistical methods. In addition to the qualifications that must necessarily accompany the results of statistical studies, some other limiting considerations of this study must be mentioned. Both of these topics are addressed in the final section of this chapter.

Results of the Study

One result of this study is based on the inconsistency of the results of the estimated water demand models from region to region. This observation is particularly interesting from a methodological point of view. Initially, the authors anticipated that estimated water demand models for each of the various watershed regions would be nearly alike. However, when the regional models failed to include a similar set of significant variables, much less similar coefficients of the corresponding variables, further study was conducted into the methodology of the earlier water demand research studies. The inconsistency of the results of the earlier studies was at first attributed to the slightly different research methods employed by the investigators. The formalation of the present study enabled a unique opportunity to test this hypothesis by comparing the results of estimated water demand models from different geographical areas using similar research techniques.

Exhibit 5-1 is a summary of the results of the individual regression equations for each Water Resource Planning Area similarly derived as

EXHIBIT 5-1

COEFFICIENTS OF EXPLANATORY REGRESSION MODELS
(AVERAGE PRICE)

Variable	Model Area					
	State	R-1	R-2	R-3	R-4	R-5
AVP	-1.07	-.929	-1.56	-.976	-1.35	-1.25
INC				+.218	+.160	
HOV		-1.64				
FIN	+.269	+2.64	+.443			
PPH	+.219			+.496	+.323	
AGE		+.277				
CHI		+.019				
TEM		+2.68			-3.161	
PRE		+.291				
APP	-.012			-.092		
FIX						
FAU	+.353			+.335	+.513	+.541
SWI				+.021		-.040
LAW				+.051	-.060	
GAR				+.005		
PUR	-.028	-.558				
TOW	-.043	+.575	+.472			
AWA				+.066		

EXHIBIT 5-1

COEFFICIENTS OF EXPLANATORY REGRESSION MODELS
(MARGINAL PRICE)

Variable	Model Area					
	State	R-1	R-2	R-3	R-4	R-5
MRP	-.977	-.871	-1.34	-1.10	-1.52	-.957
INC				+.250		
HOV		-1.696				
FIN	+.235	+2.74	+.246			
PPH	+.293	+.188		+.515	+.418	
AGE		+.303				
CHI						
TEM		+2.18				
PRE	+.089	+.236			+.084	
APP	-.064	-.068		-.088		
FIX						
FAU	+.411		+.337	+.303	+.615	+.645
SWI				+.015		
LAW				+.033	-.098	
GAR				+.005		-.050
PUR	-.028	-.478	-.269			
TOW		+.541	+.309		-.303	+.183
AWA	+.051	+.061		+.062		

EXHIBIT 5-1

COEFFICIENTS OF EXPLANATORY REGRESSION MODELS
(INCREMENTAL PRICE)

Variable	Model Area					
	State	R-1	R-2	R-3	R-4	R-5
INP	-.908	-.739	-1.36	-.993	-1.57	-.798
INC				+.303		
HOV		-1.66				+.293
FIN		+2.45	+.259			
PPH	+.278			+.517	+.407	
AGE		+.265				
CHI		+.022				+.010
TEM		+2.28				+2.13
PRE	+.095	+.237			+.080	
APP		-.107		-.094		
FIX						
FAU	+.634	+.490	+.349	+.304	+.606	+.427
SWI				+.010		
LAW					-.098	
GAR				+.007		
PUR		-.376	-.274			
TOW	-.039	+.544	+.300		-.313	+.130
AWA	-.011	+.091		+.066		

explained in Chapter Four. Although Chapter Four dealt in depth with the statistical testing of these models, Exhibit 5-1 shows clearly that the water demand model of each region is different and not comparable. The rejection of the hypothesis that the residential demand for water is invariant amoung the sub-regions of the state is significant in water resource planning. A particular policy may have different results in different regions and any inferences applicable to one region should not be used indiscriminately in another. Moreover, the results indicate factors other than those considered in this study influence the demand for water and that the use of a single model to explain the water demand in different regions could result in biased estimates. Perhaps the most useful result of this study is the development of economic models which identify the major factors which affect the residential demand for water. The significance of these models lies in their ability to explain how various factors affect the quantity of water that individual households would be willing to purchase. This information can be used to predict future water demand if it is assumed that the historical circumstances affecting the variables when the models were constructed will remain unchanged during the projection period. As explained in Chapter 1, unlike water requirement models these models are capable of generating estimates of future water demand under changing conditions. This feature allows for greatly improved predictive capabilities since the assumption of "all things will remain constant into the future", which is implicit in many predictive models, is often unrealistic.

In order to use the explanatory models for predictive purposes it is necessary to estimate future expected values of all the independent variables in the relevant model. For example, if an estimate of future residential water demand is desired for the state ten years from now it would be necessary to estimate the values of the independent variables listed in the first equation in Exhibit 4-11, i.e., average price (AVP), income and house value (FIN), persons per household (PPH), etc. Since the model is specified in exponential form it would be necessary also to convert the expected values of these independent variables into natural logrithims using a log conversion table. The final step in estimation is to multiply the expected values of the independent variables converted to logs by the appropriate coefficient and then add together. The result will be an estimate of future residential water demand stated in logs; this figure must then be reconverted using the log table once again.

The procedure outlined above can be used for each of the Arkansas Water Resource Planning Areas by using the relevant regional model specified in Exhibit 4-11. Obviously, the choice of expected values for the independent variables will affect the results. Consequently, least likely and most likely outcomes can be projected based on the selection of least likely and most likely values for the independent variables. Moreover, the use of the explanatory models enables policy makers to estimate and compare expected changes on water demand of various changes in water pricing policy. Used in this way it permits policy makers to exercise some degree of control over future water demand and introduces another element in the development of water resource management plans which are based very often on accommodating water "requirements". The efficient allocation of water resources requires not only considerations of supply but also of demand.

Reliable estimates of water demand are important for good resource management and development projects which involve large capital expenditures. The use of the requirements approach to estimate future water needs is inadequate during periods of rapidly rising prices, i.e., it is likely to result in excess capacity. The effect of a price change upon demand and revenue should be considered as part of an evaluation of future capital development projects. In the past utilities have often followed policies aimed at accommodating future "needs" through the augmentation of water supply. These types of policies are likely to become less feasible and more expensive in the future. An alternative policy to ration water through price increases may not only have the effect of decreasing the quantity consumed but also of increasing revenue if the demand for water is inelastic, and consequently, militate against the need for capital expansion.

Uses of Results for Rate Setting

Policy makers have two inherent problems to consider when determining a price schedule for water. The first is to insure that the revenues raised from the sale of water will cover the costs of providing that water. The second consideration is concerned with conservation. Because the time required to increase the capacity of plant and equipment used in the production of water is usually long, sometimes water must be priced to insure that individual households will exert the appropriate degree of conservation until additional facilities can be constructed.

Although it is difficult to recommend the appropriate price and pricing schedule given the regulatory constraints imposed on most water utilities, the results of this study can be used to estimate the effects on quantity of water demanded and revenues collected from the sale of water of different pricing policies. The effect on quantity was explained in the preceding section of this chapter; the effect on revenues necessitates an understanding of the interrelationship between price elasticity of demand and total revenue.

The price elasticity of demand is computed as the percentage change in price divided by the percentage change in quantity demanded. Because quantity demanded is inversely related to price, the price elasticity of demand is generally negative. If its value is less than - 1.00 demand is elastic, i.e., the percentage change in price exceeds the percentage change in quantity demanded. If its value is greater than - 1.00 demand is inelastic. For example, from Exhibit 4-11 it can be seen that the estimated average price elasticity of demand for the state is equal to -1.074 which means that the residential demand for water is elastic and that if the price of water changes by 1.000 percent quantity demanded will change in the opposite direction by 1.074 percent. The average price elasticity of demand for AWRPA 1 is -.929 which means that the demand is inelastic and that quantity demanded will change by .929 percent for every 1.000 percent change in price.

Since price and quantity sold determine total revenue collections, the degree of responsiveness of quantity demanded to price changes (price elasticity of demand) will determine whether total revenue increases or decreases in response to a price change. See Exhibit 5-2 for a summary of the relationship between price changes, total revenue, and elasticity.

EXHIBIT 5-2

RELATIONSHIP BETWEEN THE PRICE ELASTICITY
OF DEMAND AND TOTAL REVENUE

If η_x is Elastic:		If η_x is Inelastic:	
As P↑	TR↓ Q ↓	As P↑	TR↑ Q ↓
As P↓	TR↑ Q ↑	As P↓	TR↓ Q ↑

Where η_x = Price elasticity
P = Price
Q = Quantity

TR = Total revenue
↑ = Increases
↓ = Decreases

A study of this exhibit shows that a price increase will result in a lower quantity demanded and greater total revenues only when demand is inelastic. A review of Exhibit 4-11 shows this to be the case for the average price of water only in AWRPA 1 and 3, i.e., the Mississippi - St. Francis and White regions respectively. Whether or not utilities in regions facing elastic demands should raise the price of water to effect conservation will depend, in part, on whether or not the reduced revenues from the sale of water are offset by the reduced costs of providing less water. Moreover because these utilities face elastic demands a given percentage increase in the average price of water will reduce the demand for water more than proportionate to the reduction in regions where the demand is inelastic.

Limits of the Study

As with most studies of this nature, some remarks are in order concerning the limits of the study. The remarks that follow are not intended to quality the results to the point of oblivion, but merely to direct the reader's attention to four particular areas that should be considered when using the results of this study.

The reader should be aware that the water demand models presented in Chapter Four have not been time tested. Although these models have been derived using generally accepted statistical techniques, their validity should still be questioned to some degree since the variables have not been manipulated in the "real world" to confirm the ability of the models to predict and explain the underlying relationships that they are purported to explain. This would be a costly procedure, and usually not required, since the level of statistical significance is high enough for most practical applications.

Although the size of the data base is rather large for this type of study, there is one area where the sample size could have been increased. Namely, during the stratifying of the population into appropriate sample units, the number of towns could have been larger. The small number of different towns presented some problems in data manipulation, since the number of different pricing structures in each regional model is limited to the number of towns in that region. This is particularly troublesome problem where the number of towns in a region is only three, as in the fifth watershed region. This problem is addressed partially by defining average price in terms of the individual household's water usage, but is still evident in the marginal and incremental price variables. If additional work is considered based on the methodology of this study, the authors suggest strongly that the number of towns be enlarged - even if it is at the expense of the number of different households that would be included in the total sample.

The issue of interdependency between price and quantity resulting from the declining block pricing schedules used by the individual water systems was addressed during the course of the study. The possibility of constructing a set of simultaneous equations that could be solved by the use of two stage least squared techniques was explored to get around this problem. However, these efforts did not produce satisfactory results since the models that were developed proved to be over identified to an extent that the results of the two stage models could not be interpreted. Therefore, ordinary least squares was used to estimate the models and some interdependency of price and quantity is inherent in the estimates. However, this should not jeopardize seriously the results of the study, particularly with respect to the primary objective which was to determine and compare residential water demand models among the AWRPAs. The authors feel that further research into the area of interdependency resulting from the declining block pricing system is a fruitful area for further research.

The type of analytical procedure used in this study requires a few words of warning. Although the nature of the problem and the type of data contained within this study demand the use of statistical techniques, the reader should be reminded that there is a difference between stochastic and deterministic models. Deterministic models, once derived, will produce certain solutions, whereas stochastic models leave some room for error. Regardless of how little statistical error is included into a statistical analysis, there is always some latitude for doubt in the results. This is particularly true if the models derived in this study are used singularly. That is, if the quantity of water demanded by a particular household is determined from the use of the models set forth in this study, the probability for error is quite large. However, if the quantity of water is determined for the aggregate, the probability for error is smaller. The reader is reminded that the models will not generate the precise solution to a particular forecasting problem, but merely to provide a reasonable estimate.

BIBLIOGRAPHY

BIBLIOGRAPHY

Books

Allen, R.G.D. Mathematical Economics. London: MacMillan Press, 1973.

Awh, Robert Y. Microeconomics Theory and Applications. New York: John Wiley Co., 1976.

Babcock, Robert E. Annual Report of the Arkansas Water Resources Research Center. Fayetteville, AR: University of Arkansas, 1974.

Barr, Anthony J.; Goodnight, James H.; Sall, John P.; and Heiwig, Jane T. A User's Guide to SAS 76. Raleigh, NC: Sparks Press, 1976.

Camp, Robert Carter. "An Analysis of the Demand for Water by Domestic and Selected Commercial Users in Northern Mississippi Cities." Ph.D. dissertation, University of Mississippi, 1975.

Chiang, Alpha C. Fundamental Methods of Mathematical Economics. New York: McGraw-Hill Co., 1967.

Christ, Carl F. Econometric Models and Methods. New York: John Wiley & Sons, Inc., 1966.

Clark, Charles T. and Schkade, Lawrence L. Statistical Methods for Business Decisions. Cincinnati: South-Western Pub. Co., 1969.

Denn, Robert D.; Leahy, William H.; and McKee, David L. Spatial Economic Theory. New York: The Free Press, 1970.

Doyle, Rachel Jane. "An Empirical Investigation into the Factors Determining the Demand for Hospital Services in Arkansas." Ph.D. dissertation, University of Arkansas, 1976.

Fair, Gordon M.; Geyer, John C.; and Okun, Daniel A. Water and Wastewater Engineering. New York: John Wiley & Wons, Inc., 1966.

Ferguson, C.E. and Gould, J.P. Microeconomic Theory. Homewood, IL: Richard D. Irwin, Inc., 1975.

Gandolfo, Giancarlo. Mathematical Methods and Models in Economic Dynamics. Amsterdam: North-Holland, Inc., 1971.

Gisser, Micha. Introduction to Price Theory. Scranton, PA: International Textbook Co., 1966.

Hirshleifer, J.: DeHaven, J.C.; and Milliman, J.W. Water Supply. Chicago: The University of Chicago Press, 1960.

Hirshleifer, J.; DeHaven, J.C.; and Milliman, J.W. Water Supply Economics, Technology, and Policy. Chicago: University of Chicago Press, 1969.

Hirshleifer, Jack. Price Theory and Applications. Englewood Cliffs: Prentice-Hall, Inc., 1976.

Kemonta, Jan. Elements of Econometrics. New York: Macmillan Press, 1971.

Laidler, David. Introduction to Microeconomics. New York: Basic Books, Inc., 1974.

Mach, Leslie E. Water Law in Arkansas. Little Rock: Arkansas Industrial Development Commission, 1963.

Marshall, Alfred. Principles of Economics. London: Macmillan Press, 1936.

McKee, David L.: Dean, Robert D.; and Leahy, William H. Regional Economics. New York: The Free Press, 1970.

Mills, Richard L. Statistics for Applied Economics and Business. New York: McGraw-Hill Co., 1977.

Neter, John and Wasserman, William. Applied Linear Statistical Models. Homewood, IL: Richard D. Irwin, Inc., 1974.

Newman, Peter. The Theory of Exchange. Englewood Cliffs: Prentice-Hall, Inc., 1965.

Oh, Ho-sung. "Economics of Urban Water Demand: A Case Study of the Honolulu Board of Water Supply." Ph.D. dissertation, University of Hawaii, 1973.

Paik, C.M. Quantitative Methods for Managerial Decisions. New York: McGraw-Hill Co., 1973.

Pfaffenberger, Roger C. and Walker, David A. Mathematical Programming for Economics and Business. Ames, Iowa: The Iowa State University Press, 1976.

Pindyck, Robert S. and Rubinfeld, Daniel L. Econometric Models and Economic Forecasts. New York: McGraw-Hill Co., 1976.

Pope, Robert Martin, Jr. "Effects of Price Level and Price Change upon the Domestic Use of Water Over Time." Ph.D. dissertation, Clemson University, 1973.

Saunders, John L. and Billingsley, G.A. Surface-Water Resources of Arkansas. Little Rock: Resources and Development Commission, 1950.

Seckler, David. California Water. Los Angeles: University of California Press, 1971.

Siegel, Robert Andrew. "An Economic Analysis of Residential Water Use: A Case Study of Rhode Island." Ph.D. dissertation, University of Rhode Island, 1973.

Singer, Neil M. Public Microeconomics. Boston: Little, Brown and Co., 1972.

Sniegocki, R.T. Water Resources Summary Number 10. Little Rock: Arkansas Geological Commission, 1976.

Todd, David Keith. Ground Water Hydrology. New York: John Wiley & Sons, Inc., 1967.

Turban, Efrain and Meredith, Jack R. Fundamentals of Management Science. Dallas: Business Publications, Inc., 1977.

U.S. Department of Commerce Bureau of the Census. Statistical Abstract of the United States 1970. Washington, DC: U.S. Government Printing Office, 1970.

Village Water Supply. Washington, DC: World Book, 1976.

Wallis, Kenneth F. Introductory Econometrics. Chicago: Aldine Co., 1972.

Articles

Andrews, Donald R. and Gibbs, Kenneth C. "An Analysis of the Effect of Price on Residential Water Demand: Metropolitan Miami, Florida." Southern Journal of Agricultural Economics 12 (July 1975): 121-142.

Batchelor, R.A. "Household Technology and the Domestic Demand for Water," Land Economics 51 (1975): 208-223.

Berry, Dale W. and Bonem, Gilbert W. "Predicting the Municipal Demand for Water." Water Resources Research 10 (December 1974): 56-79.

Bromley, Daniel W.; Butcher, Walter R.; and Smith, Stephen C. "Policy and Research Implications of the National Water Commission's Recommendations." Land Economics 50 (1974): 25-43.

Burke, Thomas R. "A Municipal Water Demand Model for the Conterminous United States." Water Resources Bulletin 6 (July-Aug. 1970): 661-681.

"1970 Census of Housing, General Housing Characteristics, Arkansas." U.S. Department of Commerce Bureau of the Census. 8 (1971): 326-341.

"Climatological Data, Arkansas." U.S. Department of Commerce National Oceanic and Atmospheric Administration Environmental Data Service. 5 (1977): 172-193.

Conley, Bryan C. "Price Elasticity of the Demand for Water in Southern California." Annals of Regional Science 52 (Dec. 1967): 180-189

Darr, Peretz; Feldman, Stephen; and Kamen, C.S. "A reconsideration of the 'Requirement Approach' to Residential Water Use Water Resources Research 11 (Dec. 1974): 13-45.

Dunn, Dorothy F. and Larson, Thurston. "Relationships of Domestic Water Use to Assessed Valuation. Journal of American Works Association 55 (April 1963): 72-85.

Gottlieb, Manuel. "Urban Domestic Demand for Water: A Kansas Case Study." Land Economics 39 (May 1963): 204-210.

Hanke, Steve H. "Dynamic Aspects of Urban Water Demand." Proceedings of the Fourth Annual American Water Resources Conference, Proceedings Series No. 4 (1968): 21-28.

Hanke, Steve H. "Some Behavioral Characteristics Associated With Residential Water Price Changes." Water Resources Research-Brief Reports 6 (Oct. 1970): 1383-1386.

Hanke, Steven H. "Demand for Water Under Dynamic Conditions." Water Resources Research 6 (Oct. 1970): 1253-1261

Hanke, Steven H. and Flack, J. Ernest. "Effects of Metering Urban Water." Journal of American Water Works Association 60 (Dec. 1968): 1359-1366.

Headley, J.C. "The Relation of Family Income and Use of Water For Residential and Commercial Purposes in the San Francisco-Oakland Metropolitan Area." Land Economics 39 (Nov. 1963): 441-449.

Himes, Lawrence G. "The Long-run Cost Function of Water Production for Selected Wisconsin Communities." Land Economics 3 (Feb. 1969): 133-140.

Hirshleifer, Jack and Milliman, Jerome W. "Urban Water Supply: A Second Look." American Economic Review 57 (1967): 33-58.

Hogarty, Thomas F. and Mackay, Robert J. "The Impact of Large Temporary Rate Changes on Residential Water Use." Water Resources Research 11 (Dec. 1975): 116-123.

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

Kiker, Clyde and Lynne, Gary D. "Water Allocation Under Administrative Regulation: Some Economic Considerations." Southern Journal of Agricultural Economics 10 (Dec. 1976): 26-33.

Lane, M.N. and Littlechild, S.C. "Weather-Dependent Pricing for Water Resources in the Texas High Plains." Water Resources Research 12 (Aug. 1976): 133-138.

Larson, B.O. and Hudson, H.E. "Residential Water Use and Family Income." Journal of American Water Works Association 3 (Aug. 1951): 603-611.

Leser, Conrad E.V. "The Measurement of Elasticities of Demand." Applied Statistics 7 (June 1954): 74-84.

Linaweaver, F.P.: Geyer; and Wolff. "Summary Report on the Residential Water Use Research Project." Journal of American Water Works Association 59 (March 1967): 267-282.

Metcalf, I. "Effects of Water Rates and Growth in Population upon per Capita Consumption." Journal of American Waterworks Association 15 (Jan. 1955): 1-22.

Milliman, J.W. "Policy Horizons for Future Urban Water Supply." Land Economics 39 (May 1963): 109-132.

North, R.M. "Consumer Responses to Price of Residential Water." Proceedings of the Third Annual American Water Resources Conference 3(1967): 10-16.

"Public Water Supplies." Arkansas Department of Health, Division of Engineering, (April 1977): 13-17.

Reynolds, John E. and Conner, Richard J. "Spatial and Temporal Water Allocation in the Kissimmee River Basin." Southern Journal of Agricultural Economics 13 (July 1977): 63-76

Saunders, Robert J. "Urban Area Water Consumption: Analysis and Projections." Quarterly Review of Economics and Business 18 (1969): 5-20.

Seidel, H.F.; Johnson, S.S.; and Dencker, D.O. "A Statistical Analysis of Waterwork Data for 1950." Journal of American Waterworks Association (Dec. 1953): 1316.

Seidel, H.F. and Bauman, E.R. "A Statistical Analysis of Water Works Data for 1955." Journal of American Water Works Association 49 (Dec. 1957): 1531-1566.

Seidel, H.F. and Bauman, E.R. "A Statistical Analysis of Water Works Data 1960." Journal of American Water Works Association 58 (Dec. 1966): 1113-1123.

Shulstad, Robert N.; Ziegler, Joseph A.; and Cross, Eddie D. "Projected Water Requirements and Surface Water Availability for Arkansas." Arkansas Soil and Water Conservation Commission Special Report 61 (April 1978): 1-49.

Turnovsky, Stephen J. "The Demand for Water: Some Empirical Evidence on Consumers Response to a Commodity Uncertain in Supply." Water Resources Research 5 (April 1969): 350-361.

Ware, James E. and North, Ronald M. "Price and Consumption of Water For Residential Use in Georgia." The Atlanta Economic Review 15 (Oct. 1968): 9-13.

Wong, S.T. "A Model on Municipal Water Demand: A Case Study of Northeastern Illinois." Land Economics 48 (Feb. 1972): 98-113.

Whitford, Peter W. "Residential Water Demand Forecasting." Water Resources Research 8 (1972): 12-13.

Wolff, Jerome B. "Peak Demands in Residential Areas." Journal of American Water Works Association 5 (Oct. 1961): 22-32.

Wollman, Nathaniel. "A Preliminary Report on the Supply of and Demand for Water in the United States as Estimates for 1980 and 2000." Water Resources Activities in the United States 7 (1960): 1-10.

Dunn, James E. (Unpublished Manuscript) "Methods of Multivariate Analysis."

ACKNOWLEDGEMENTS

The authors would like to acknowledge the assistance of the following organizations and individuals in completing this report: Bureau of Business and Economic Research, College of Business Administration, and Water Resources Research Center of the University of Arkansas; individuals who took the time to complete the questionnaires in the towns of Blytheville, El Dorado, Hot Springs, Pine Bluff, Batesville, Harrison, Pochahontas, Walnut Ridge, Van Buren, Magnolia, Mena, and Texarkana, and Ms. Deborah Keeling who typed the final manuscript.