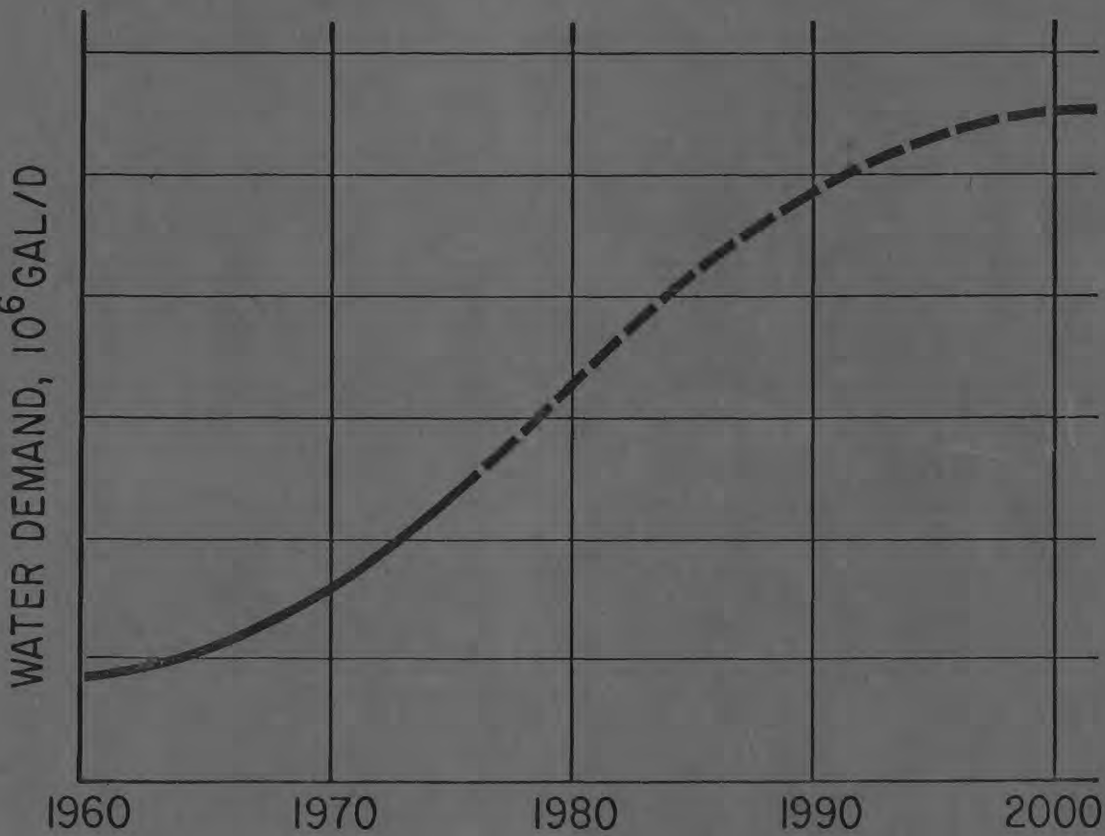
UNITED STATES
DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY

TECHNIQUES FOR WATER DEMAND ANALYSIS AND FORECASTING: PUERTO RICO, A CASE STUDY



Open-File Report 75-94

Prepared in cooperation with the
Commonwealth of Puerto Rico

UNITED STATES
DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY

TECHNIQUES FOR WATER DEMAND ANALYSIS AND FORECASTING: PUERTO RICO, A CASE STUDY

By E. D. Attanasi, E. R. Close, and M. A. López

Open-File Report 75-94

Prepared in cooperation with the
Commonwealth of Puerto Rico

San Juan, Puerto Rico

1975

CONTENTS

ILLUSTRATION

TABLES

TABLES--Continued

Acknowledgments

This report would not have been possible without the cooperation of the Puerto Rico Planning Board, the Puerto Rico Aqueduct and Sewer Authority, and the Puerto Rico Environmental Quality Board. In Particular, Ms. Sarah López of the Puerto Rico Planning Board aided in helping us obtain data sources. Puerto Rico Aqueduct and Sewer Authority also supplied all of the billing records over a 10-year period, from which the water use data base was developed. The work of María Suárez of the Puerto Rico Environmental Quality Board was also indispensable.

TECHNIQUES FOR WATER DEMAND ANALYSIS AND FORECASTING: PUERTO RICO, A CASE STUDY

by
E.D. Attanasi, E.R. Close and M.A. López
U.S. Geological Survey, WRD

ABSTRACT

The rapid economic growth of the Commonwealth of Puerto Rico since 1947 has brought public pressure on Government agencies for rapid development of public water supply and waste treatment facilities. Since 1945 the Puerto Rico Aqueduct and Sewer Authority has had the responsibility for planning, developing and operating water supply and waste treatment facilities on a municipio basis. The purpose of this study was to develop operational techniques whereby a planning agency, such as the Puerto Rico Aqueduct and Sewer Authority, could project the temporal and spatial distribution of future water demands.

This report is part of a 2-year cooperative study between the U.S. Geological Survey and the Environmental Quality Board of the Commonwealth of Puerto Rico, for the development of systems analysis techniques for use in water resources planning. While the Commonwealth was assisted in the development of techniques to facilitate ongoing planning, the U.S. Geological Survey attempted to gain insights in order to better interface its data collection efforts with the planning process.

The report reviews the institutional structure associated with water resources planning for the Commonwealth. A brief description of alternative water demand forecasting procedures is presented and specific techniques and analyses of Puerto Rico demand data are discussed. Water demand models for a specific area of Puerto Rico are then developed. These models provide a framework for making several sets of water demand forecasts based on alternative economic and demographic assumptions. In the second part of this report, the historical impact of water resources investment on regional economic development is analyzed and related to water demand forecasting. Conclusions and future data needs are in the last section.

INTRODUCTION

Over the past 2 years the U. S. Geological Survey and the Environmental Quality Board of the Commonwealth of Puerto Rico carried out a cooperative study to develop systems analysis techniques for use in water resource planning. One part of the study concentrated on water-supply features of a site-selection model formulated as a mixed-integer program (Moody and others, 1973) and another part concentrated on water demand analysis. The latter part, which was extended to developing and presenting operational techniques for forecasting water demand, is reported on below. Water demand analysis is a vital part of water resource planning because it serves to identify where future development of supplies will provide the greatest benefit. In addition, a topic which was also investigated and related to forecasting water demands concerns how water resource development influences economic growth of an area. This latter subject was investigated by examining the historical experience of Puerto Rico.

Water resource planners require hydrologic data for the efficient, sound design and siting of water resource facilities. Since 1957 the U. S. Geological Survey has maintained an island-wide network of surface water stations and observation wells in Puerto Rico through the Federal-Commonwealth cooperative program. From the perspective of the data collectors, the two purposes of the cooperative study were: 1) to assist the Commonwealth in the development of systems analysis techniques to facilitate their on-going planning efforts, and, 2) to gain insight into the water resources planning process in order to provide a framework for data collection prior to project design and implementation. The overall Water Resource Planning Model Study also provides an opportunity to evaluate the adequacy of data collection programs in light of water resource planning practices.

This report begins with a description of the institutional structure associated with the supply of water for the Commonwealth. After a brief review of alternative water demand forecasting procedures, the specific techniques and their results are analyzed and discussed. Following this, the historical impact of public works and in particular water resource investments on regional economic development is analyzed. This analysis is considered from the perspective of how such investments might affect future water demands. The final section of the report summarizes the water use and economic data developed for the study and outlines steps to implement the procedures described.

For use of those readers who may prefer to use metric units rather than English units, the conversion factors for the terms used in this report are listed below:

Multiply English unit	By	To obtain metric unit
million gallons (Mgal)	3785	cubic metre (m^3)
million gallons per day (Mgal/d)	0.04381	cubic metre per second (m^3/s)
acre-foot	0.001233	cubic hectometre (hm^3)
mile (mi)	1.609	kilometre (km)

INSTITUTIONAL SETTING

Since 1945 the Puerto Rico Aqueduct and Sewer Authority (PRASA) has been responsible for planning, developing and operating water-supply and water and waste treatment facilities on a <u>municipio</u>[1] basis. Public irrigation and electric power generation facilities are developed and operated by the Puerto Rico Water Resource Authority (PRWRA). Both authorities are Government-owned corporations with several appointed Government officials acting on their Boards of Directors.

The municipal water-supply system consists of 62 urban and 176 rural water-supply systems. With the exception of the San Juan metropolitan area, only six systems serve more than one municipio. Currently, about 56 percent of these systems obtain water from reservoirs or stream diversions, 39 percent from ground-water wells and 5 percent from both ground water and surface water. At present Puerto Rico Aqueduct and Sewer Authority supplies approximately 195 Mgal/d (8.5 m^3/s) to 430,000 urban customers and 178,000 rural customers (Moody and others, 1973). Extrapolation of historical trends suggest that by 1990 Puerto Rico Aqueduct and Sewer Authority may be supplying as much as 494 Mgal/d (21.6 m^3/s) (Puerto Rico Aqueduct and Sewer Authority, 1969) implying that water demand management may be required if the water resources of the Commonwealth are not developed rapidly enough. It has been estimated that heavy industry self-supplies as much as 20 Mgal/d (0.9 m^3/s). Water rights, granted when the Island was under Spanish rule, are still honored. In 1960 the Public Service Commission was given authority to grant and control transfer of future water rights.

Puerto Rico Water Resource Authority currently supplies irrigators

1 A municipio is a local government entity which is roughly equivalent in size to a county in the continental United States. There are 76 municipios which comprise Puerto Rico.

approximately 218,710 acre-feet (270 hm^3) of water per year with 272,000 acre-feet (335 hm^3) projected by 1984 (Puerto Rico Department of Natural Resources, 1973). However, any forecasting of irrigation water demand remains tenuous because the demand for irrigation water is perhaps more sensitive to political decisions, such as crop subsidization, and cost sharing formulas for public construction, than to natural trends in the economic development process (Cordero, 1969). Moreover, resistance by older farmers to the introduction of new farming techniques, has resulted in instances where farmers may not even bother to irrigate, or irrigate in a wasteful fashion (Cordero, 1969). For these reasons and because of the difficulty of establishing a data base, irrigation water demand projections were not made in this study. This report does, however, present a demand analysis along with operational procedures for forecasting residential, commercial, and industrial water demand on a municipio basis. Discussion in the second part of the report relates to the regional economic impact of all kinds of water resource investments, including irrigation investments, on regional economic development.

DEMAND ANALYSIS AND FORECASTS

Nature of Demand Modeling and Applications to Water Resource Planning

The following discussion is presented to provide a general view of the nature of demand modeling and the procedures used to generate the forecasts. Along with the nature of the demand models, model parameterization techniques and forecasting performance criteria are also considered. Finally, the discussion relating to the application of the models indicates the relevance of such models to actual planning problems.

Models for forecasting values of economic variables can generally be classified as predictive (unconditional) or descriptive (conditional) in nature (Armstrong and Graham, 1972). Predictive models are constructed under the assumption that conditions determining values of the variable of interest will remain unchanged in the future. Historical values of the variable, water use for example, are used to fit or parameterize a model which characterizes a stochastic process. This parameterized model is then used to generate future values of the variable. Another example of a predictive forecasting tool is a simple linear extrapolation of historical water use trends.

Descriptive or conditional models are specified to characterize a behavioral response between the dependent variable, for example, water demand, and a set of independent or explanatory variables such as price and income. Because demand functions are behavioral relationships, the explanatory variables, and in some cases the functional forms of equations, are selected to be consistent with economic theory. Within the context of forecasting, descriptive economic models are often preferred to predictive models for the following reasons. While the basic demand relationship may be unchanged, the dynamic influence of the growth of individual conditioning variables may be examined in forecasting levels of water demand. Demand analysis conducted with descriptive economic models provide the opportunity for either reinforcing or questioning assumed theoretical relationships. Because the descriptive models are frequently estimated using standard statistical techniques, measures of uncertainty in future projections can be explicitly accounted for. In particular, uncertainty resulting from model misspecification and estimation can theoretically be separated from the uncertainty associated with future value of the explanatory (independent) variables. Confidence intervals will then reflect the range of values where future projections may fall. Descriptive economic models can be used by planners for policy analysis because the effects of specific policy variables may be separated from uncontrolled variables. For instance, the influences on water demand of changes in prices, metering policies, and industrial zoning policies can be projected with such models.

An additional distinction between forecasting techniques can also be made by specifying whether the model is static or dynamic in nature. Static models frequently do not recognize that changes in explanatory variables do not produce immediate responses in quantity demanded. Realistically, such responses are generally spread over time, and this behavior may be theoretically related to the consumer's inventory of the commodity or habit formation. The standard (static) approach to demand analysis involves estimating parameters of the following nature

$$q_t = f(x_t, p_t, z_{1t}, \ldots, z_{nt}, u_t), \qquad (1)$$

where q_t is a measure of consumption or withdrawal, $f(\cdot)$ is a general mathematical function or form, x_t is a measure of income or output, p_t is deflated price of the commodity, z_i is any other explanatory variable,[1] t is a specified time period and u_i is a disturbance or stochastic term. In general, the shortness of time series, lack of data, and independent variation in the explanatory variables limit the number of predictors which can be introduced. Rarely does economic theory specify a priori a functional form. Therefore, choosing a functional form is based on a combination of factors including fit and consistency with a dynamic formulation. A general functional form for the dynamic relationship is expressed in terms of rates per unit time where

$$q(t) = f(s(t), x(t), p(t), z_1(t), \ldots z_n(t)), \qquad (2)$$

where q(t) is the rate of consumption, x(t) the income rate or output rate, s(t) is the inventory stock of the commodity or measure of habit persistence (psychological stock) at time t, and so forth. In general, s(t) is not observable and the coefficient or parameter corresponding to this variable is estimated indirectly. Equation (2) is generally referred to as a structural equation, theoretically characterizing the actual demand relationship, while the equation which is empirically estimated is the reduced form equation. Estimates of parameters of the reduced form are used to calculate the structural equation parameters which may be solved for in terms of the reduced form parameters. In the discussion of residential or household water demand a simple dynamic model is examined in detail. Specific problems which are encountered when dynamic models are used in the forecasting context will be discussed in relation to residential water-demand forecasting.

Economic theory and classical-statistical techniques are applied here to develop parameterized water-demand functions. These empirical functions are analyzed and shown to provide information regarding the responsiveness of consumers to changes in water prices, income or output. Furthermore, demand functions provide a basis from which to project future water demands. The process of generating forecasts based on alternative

---

[1] Because prices are set by the Government-owned utility, prices are taken as exogenous (determined apart) to the demand equation.

policies and economic and demographic conditions also provides information which can aid the planner in several ways:

First, economic water-demand functions provide a basis for assessing the sensitivity of forecasts to various economic, social, and demographic variables. This may be done by systematically varying these influences and observing changes in values of the forecasted variable.

Secondly, because benefits associated with water-supply features of projects can be related to water demand functions (Turnovsky, 1973), empirical-demand functions provide a means for evaluating whether specific investments are economically justified. Moreover, for multipurpose projects, benefits need to be developed in order to compare economic gains and losses for alternative project uses.

Finally, the demand functions provide a means of predicting the success of attempts at demand management by indicating the responsiveness of demand to alternative pricing policies. For a developing area, the models can be applied to forecast water demand for areas which are to be supplied in the near future by the municipal water-supply system.

Before presenting the specific water-demand models, several preliminary comments to explain the methodology are now made. In the following sections water-demand models and projections for residential, commercial and manufacturing water use are presented. Each subsection begins with a brief review of existing demand models along with their shortcomings. The particular models used in this study are then developed along with an analysis of the estimation results corresponding to the metropolitan San Juan Region. Following this, alternative assumptions concerning growth of population, income, and other explanatory variables are used to generate alternative forecasts which are subsequently interpreted. Because much of the explanation of the models and practical problems encountered with them are best viewed in the context of the actual applications, discussions of these points are presented in relation to residential water demand and are not repeated in subsequent sections.

Residential Water Demand

Nature of Residential Water Demand

Economists have probably given more attention to residential water demand than other water uses. Table 1 provides a summary of previous residential water demand studies. Although early studies were confined to the analysis of cross-sectional data from individual water districts (see Howe and Linaweaver, 1967), later writers have utilized time series information (Wong, 1972 and Young, 1973). However, all of these models are static in nature. As previously alluded to, one problem which cannot be addressed with a static model is the question of the degree of dependence of commodity demand in one period on consumption in previous periods. In particular, the commodity may be influenced by consumer "habit" buying or by "stock or inventory" effects whereby a component of current demand is largely independent of current economic conditions. These influences are particularly significant when commodities are narrowly defined, as in the case of water (Houthakker and Taylor, 1971). These considerations are important because it would be useful to planners to know how responsive consumers will be to immediate income and price changes or if consumers will take a long period of time to adjust consumption to new price levels.

"Inventory" or "habit" effects (two effects precisely opposite in nature) imply that current consumption of the commodity is dependent not only on current income or prices but on the stock of the goods held by the consumer. In the case of durable goods an "inventory" effect is interpreted as the adjustment by the consumer of durable goods to some desired level of consumption, given his current stocks of the goods, income, and prices. Alternatively, for habit persistence the interpretation is that the consumer has built up a psychological stock of habits whereby, the more he has consumed of the commodity in the past, the more he will currently consume with tastes, income and prices given. In the case of household water demand, it is possible to argue a priori that either effect may prevail. Although individual personal water use may be subject to habit persistence, that part of water use which is complementary to consumer durable goods may exhibit fluctuations in demand which reflect fluctuations in the demand for consumer durable goods such as washing machines and dishwashers. If "inventory" effects are reflected in commodities in such complementary goods, then "inventory" effects might be induced in household water use as families acquire new luxury goods.

While economic theory is useful in identifying the relevant determinants of demand, it does not suggest a specific functional form for the dynamic demand relationship. Suppose $q(t)$ is the rate of consumption at time t, $x(t)$ is a measure of income rate, $p(t)$ is the unit price rate, and s(t) is a stock variable of complementary goods or a psychological stock of services determined by habit. For simplicity the following functional form of the demand relationship is considered here

$$q(t) = \alpha + \beta s(t) + \gamma x(t) + \eta p(t) + u(t), \qquad (3)$$

where u(t) is the stochastic component of the relation. Houthakker and Taylor (1971) argue on an <u>a priori</u> basis that $\beta < 0$ if demand exhibits an inventory effect and $\beta > 0$ if habit persistance is present. It is reasonable to expect that the component of household demand, which is highly complementary to new durable goods, might also exhibit an inventory effect as these appliances serve to increase in-house water use. Although several water-saving technologies are available, the incentive, in terms of reduced capital costs, is to install heavier (more) water-using devices (Howe and others, 1971). Thus β, when calculated for household water use, reflects "inventory" influences attributable to changes in water use resulting from acquisition of durables and the coefficient of water use for the new durables. Therefore, until water saving devices are widely installed, one can interpret a result of $\beta < 0$ to suggest that water demand is dominated by an inventory effect induced by purchases of consumer appliances.

Because s(t) is not observable, its coefficient must be indirectly estimated. The stock variable will be eliminated and parameters of equation (3) estimated indirectly by utilizing a procedure developed by Houthakker and Taylor (1971). The rate of change of the services of the stock variables depends on the rate of purchase and wearout (or depreciation), δ, of the stock of services

$$\dot{s}(t) = q(t) - \delta \dot{s}(t). \qquad (4)$$

Then from equation (3)

$$s(t) = 1/\beta \, [q(t) - \alpha - \gamma x(t) - \eta p(t)] \qquad (5)$$

$$\dot{s}(t) = q(t) - \frac{\delta}{\beta} [q(t) - \alpha - \gamma x(t) - \eta p(t)] \qquad (6)$$

Because $\dot{q}(t) = \beta s(t) + \gamma x(t) + \eta p(t)$, and substituting for s(t) from equation (5)

$$\dot{q}(t) = \alpha\delta + (\beta - \delta) q(t) + \delta\gamma x(t) + \gamma x(t) + \delta\eta p(t) + \eta p(t) \qquad (7)$$

Reference	Location	Type of study	Results
Bain, Caves and Margolis (1966)	Northern California	Cross-sectional study of 41 communities; related demand to price.	Relatively high price elasticity (-1.099) for water demand.
Conely (1967)	Southern California	Regressed logarithm of per capita demand against logarithm of price for cross-sectional data.	Price elasticity estimates were -1.02 to -1.09.
Guilbe (1969)	Mayagüez, Puerto Rico	Cross-sectional analysis for sample households during 12-month period.	Found relation between monthly water use and housing value.
Hanke (1970)	Selected communities in Colorado	Observed communities during transition from unmetered to metered water policy.	Reduction in water demand with installation of meters was found to be permanent.
Headley (1963)	San Francisco Oakland	Related water demand from 9-year period to family income.	Higher income associated with higher levels of water demand.
Howe and Linaweaver (1967)	Northeast United States	Cross-sectional study of 35 communities; related water demand to price and income.	Price elasticity estimates were -0.21 to -0.23 and income elasticity estimates were 0.31 to 0.37.
Turnovsky (1969)	Massachusetts townships	Cross-sectional study; related demand to housing and water price.	Price elasticity estimates were -0.05 to -0.40.
Wong (1972)	Northern Illinois	Related price and income to water demand using 10 years of data.	Price elasticity estimates were -0.02 to -0.28 and income elasticity constraints were 0.20 to 0.26.
Young (1973)	Tucson, Arizona	Time series analysis for 1947-71 for municipal water supply of Tucson, Arizona	Relate water demand to price and climatic variables with estimated price elasticity of -0.4 to -0.6.

With s(t) eliminated in equation (7) and making the following discrete approximation of the continuous variables

$$\dot{q}(t) \approx q_t - q_{t-1} \ , \quad q(t) \approx 1/2\ (q_t + q_{t-1})$$

then equation (7) becomes

$$q_t = \frac{\alpha\ \delta}{1-1/2\ (\beta-\delta)} + \frac{1+1/2\ (\beta-\delta)}{1-1/2\ (\beta-\delta)}$$

$$+ \frac{(1 + 1/2)}{1-1/2\ (\beta-\delta)}\ x_t\ \frac{\gamma\ \delta}{1-1/2\ (\beta-\delta)}\ x_{t-1}$$

$$+ \frac{(1 + 1/2)}{1-1/2\ (\beta-\delta)}\ p_t\ \frac{\eta\ \delta}{1-1/2\ (\beta-\delta)}\ p_{t-1} \qquad (8)$$

Regression coefficients of equation (8) may be solved to provide estimates of α, β, γ, and η.

Finally, parameters of the following reduced form equation were estimated by several regression techniques:

$$q_t = A_O + A_1\, q_{t-1} + A_2 \Delta x_t + A_3 x_{t-1} + A_4 \Delta p_t + A_5 p_{t-1} \qquad (9)$$

Parameters of equation (9) yield the structural equation parameters of equation (5) when solved by using equation (8). Several statistical problems arise with the parameter estimation of equation (7). First, the presence of the lagged dependent variable implies a degree of autocorrelation is present. This suggests that a straightforward application of ordinary least squares would produce biased and inconsistent coefficient estimators (Goldberger, 1964). Secondly, pooling cross-sectional and time-series information without appropriate adjustments in the estimation procedure could also produce biased and inefficient estimates (Kmenta, 1971). In order to overcome these difficulties, an iterative regression technique developed by Balestra and Nerlove (1966) was employed which provides asymtotically consistent estimates. A description of this estimation procedure is provided in Appendix A.

Initially, it was felt that meaningful demand relationships could be derived for unmeasured or flat-rate customers, but this was not done for several reasons. Probably, the most important reason is the method by which sales to flat-rate customers are calculated by Puerto Rico Aqueduct and Sewer Authority. Unmetered use was calculated as a residual representing the part of water production not accounted for by sales to metered customers. Because of the high water system leakage rates, as much as 50 percent in some areas of Puerto Rico (Buck, Seifert and Jost, 1971), the data for unmetered water demand might be more representative of the condition of the local system than of the actual water used. Experimentation with the data for unmetered customers did not produce any meaningful demand relationships. Therefore, it was decided to utilize the demand functions for metered water demand with marginal (average) prices set to zero.

The data for the region under study, the San Juan metropolitan area, included 13 cross-sectional units (municipios), and a time span of 11 years (Figure 1). This region was chosen because of its rapid economic growth from 1960 to 1971 and the relatively developed stage of the municipal water system. Moreover, this area experienced uniform climatic conditions. Prices for metered water users were computed from monthly data and averaged over each year in order to obtain the effective annual price. Although this procedure resulted in an average price rather than a marginal price, the modified block rate[1] in effect for Puerto Rico may not have produced a substantial difference between average and marginal prices because the block rate was not necessarily monotonically decreasing. However, appropriate qualifications should still be made when interpreting empirical estimates of price elasticities[2] and income elasticities[3] (Verleger, 1972), when average price data is used to estimate parameters of demand functions.

---

[1] Block rate pricing simply means that the commodity is priced at different levels for broad ranges. That is, there is not a unique price corresponding to each quantity.

[2] Price elasticity represents the percentage change in demand at a fixed level of income induced by a specific percentage change in price.

[3] Income elasticity represents the percentage change in demand at a fixed price which is induced by a specific percentage change in income.

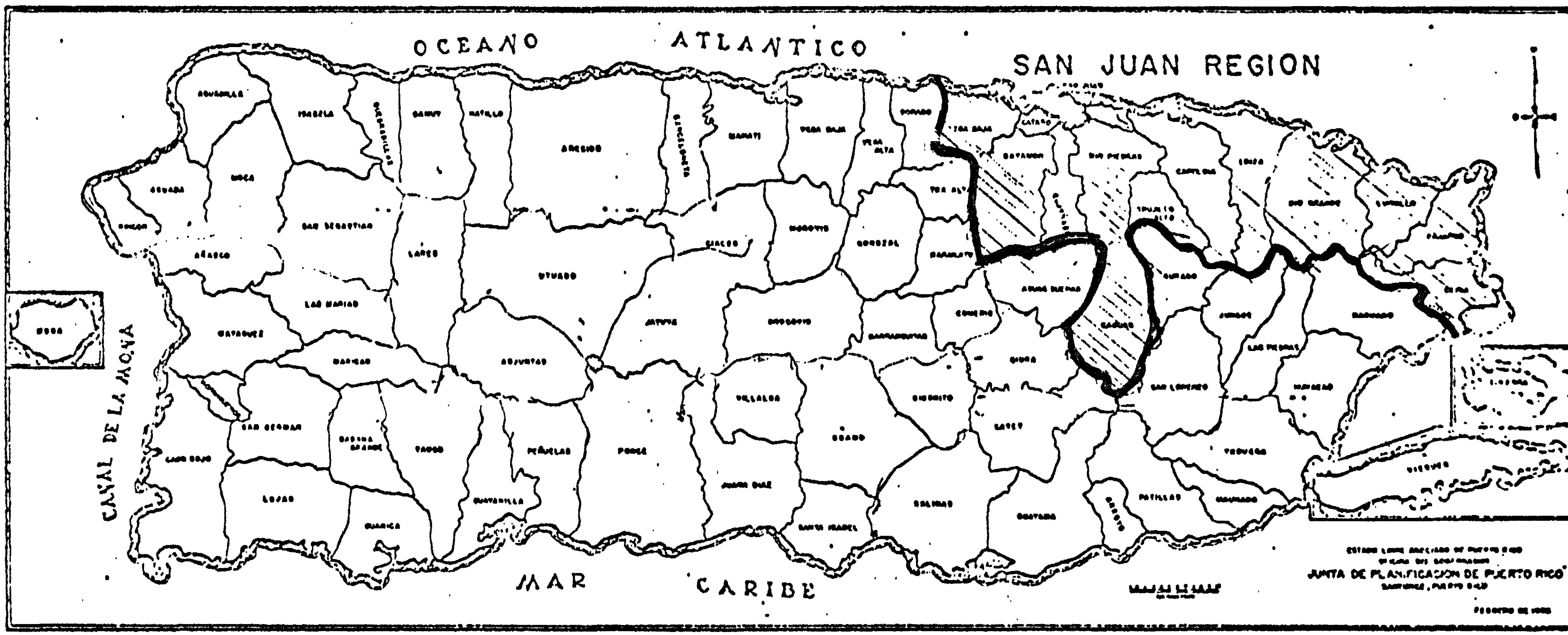


Figure 1.--Delimitation of San Juan test area.

Analysis of Empirical Results

Residential water demand functions were constructed which relate demand to per capita water use and household water use. The ordinary least squares results are presented along with the results of the Nerlov's error component method of parameter estimation from a time series of cross sectional data. Monte Carlo experiments comparing small sample properties of alternative methods of estimation indicate that the error components procedure described in Appendix A compares favorably with all other methods of estimation (including maximum likelihood) (Nerlove, 1971b). Small sample bias also appeared for error component estimates of coefficients for all explanatory variables except the lagged dependent variable (Nerlove, 1971a). The importance of coefficient bias and estimation procedure for forecasting purposes are discussed later. Coefficient estimates for the demand model specified in equation (9) along with their standard errors are shown in table 2 for the two methods of estimation. Ordinary least squares regression provided the better fit of the data as measured by the coefficient of determination. Signs of the coefficients are consistent with what theory would suggest. That is, coefficients of the income variables A_2 and A_3 are positive while the price coefficients A_4 and A_5 are negative. However, the coefficients were not always statistically significant. In contrast to the results of the water demand equation estimated by Houthakker and Taylor (1971), the estimate of β of equation (3) suggests habit persistence rather than an induced inventory effect for water demand. However, the habit formation parameter has a relatively large standard error. Estimated coefficients exhibit that pattern found in Nerlove's Monte Carlo studies where A_1 was consistently estimated larger by the ordinary least squares procedure.

Variables in table 2 denoted by γ and η, respectively, are the structural parameters of equation (3) and the long term derivative of the q with respect to income and price. [1] That is, from equation (3) it is argued that the short term effect of changes in income, x, or price, p, are measured by γ and η respectively while the long-term effects are measured by K_γ and K_η where

$$K_\gamma = \frac{\hat{\gamma}\hat{\delta}}{\hat{\delta}-\hat{\beta}}$$

$$K_\eta = \frac{\gamma\delta}{\delta-\beta}$$

with δ, β and η defined in equations (3) and (4). δ is interpreted as the rate of depreciation of the stock variable in equation (3)[2] while β is

---

[1]Along with being estimates of the structural equation, coefficients and represent the short term derivative of q with respect to income and price, respectively.

[2]In the development of the model by Houthakker and Taylor (1971, 10-24), the parameter, δ, is over-identified, i.e., it may be calculated using A_3 and A_2 and also A_4 and A_5 of equation (7) which result in two different values. In their estimation procedure an identification restriction was imposed by solving the roots of a nonlinear equation (see pp. 47-51). However, in this study, the unrestricted estimates of δ were calculated separately (using income and price coefficients) and employed to estimate K_γ and K_η, respectively.

Table 2.--Estimated regressions and structural coefficients for per capita and household residential water demand (in gallons per month)

	A_0 Constant	A_1 q_{t-1}	A_2 Δx_t	A_3 x_{t-1}	A_4 Δp_t	A_5 p_{t-1}	SEE	R^2	β	γ	η	K_γ	K_η
Per Capita													
OLS	401 (114)	0.848 (.050)	162 (97)	54 (34)	-1790 (287)	-500 (230)	97.0	0.88	.235 (.359)	146 (106.2)	-0.17 (.027)	356	-0.598
EC	189 (47)	.614 (.077)	204 (97)	106 (54)	-1390 (298)	139 (278)	141	.54	.225 (.561)	187 (117)	-.16 (.029)	274	.145
Household													
OLS	1720 (493)	.861 (.051)	158 (102)	43.4 (34.6)	-7770 (1220)	-2660 (982)	412.5	.87	.169 (.320)	147 (110.6)	-.71 (.113)	312.5	-1.42
EC	793 (200)	.610 (.078)	217 (101)	109 (57)	-5910 (1250)	-570 (1170)	591.5	.53	.184 (.530)	201.8 (122.2)	-.699 (.122)	278.1	-.861

OLS = ordinary least squares method of estimation

EC = error components method of estimation

SEE = standard error of estimate

R^2 = coefficient of determination

β, γ, η, K_γ and K_η as defined in the text

Numbers in parenthesis are standard deviations of coefficients the habit persistence parameter. Because γ, η, $K\gamma$ and $K\eta$ represent derivatives, these estimated values can be used to compute income and price elasticities.[1] Calculated elasticities are shown in table 3.

Table 3.--Calculated elasticities

	Short Run		Long Run	
	Income	Price	Income	Price
Per Capita				
OLS	0.0825	-0.61	0.2008	-2.19
EC	.1388	- .79	.2041	.7
Household				
OLS	0.0848	-0.65	0.1808	-1.29
EC	.1487	.81	.2050	-1.00

Calculated income and price elasticities can be employed to predict the consequence of changes in price and income on household and per capita water demand. Signs of the short run elasticities are as expected for both methods of estimation with income exerting a positive influence and price a restraining force on water demand. Long-run income elasticities are approximately twice the short-run elasticities suggesting that

---

1 The particular formulas for price and income elasticities are

$$e_p = \frac{dq}{dp} \quad \frac{p}{q}$$

$$e_I = \frac{dq}{dI} \quad \frac{I}{q}$$

The estimates γ, η, $K\gamma$ and $K\eta$ represent short and long-term derivatives of q with respect to price and income, respectively. The value q in the rates was computed from the regression equation at the mean values of the variables.

residential water demand for Puerto Rico is more sensitive to changes in permanent income than short-run fluctuations in income.

Calculated price elasticities by the two estimation techniques were significantly different. The error components procedure produced short-run elasticities which were larger (in absolute value) than the ordinary least squares technique. Surprisingly, the calculated long-run price elasticity of the error components estimate is positive while the ordinary least squares equation provided a negative elasticity which was larger in absolute value than the respective short-run elasticity. The latter result suggests that pricing policy is more effective for restraining long-run than short-run residential water demand. This interpretation seems reasonable in light of the greater responsiveness to changes in permanent income as opposed to short-run changes in income. There is no obvious explanation for the positive long-run coefficient for the error component model except that the relatively large standard error associated with A_5 might suggest that the point estimate of the elasticity is not very efficient and therefore could be an artifact of the estimation procedure. In comparison with aggregate United States estimates obtained by Houthakker and Taylor (1971)[1], the estimates of income elasticity for Puerto Rico are larger than calculated short-run elasticities found by Houthakker and Taylor (1971).

Additional information relating to residential water demand was obtained by re-estimating and comparing the demand equations employing data from the two poorest and the two wealthiest municipios. The estimated equations are shown in table 4 for both per capita and household water demand. The estimated equations for the poorer municipios produced better fits than the full set of data in table 1 while the opposite was true for the wealthier municipios. Comparison of the estimated coefficients for the rich and poor municipios indicates significant differences in the estimated model coefficients and elasticities. The coefficients exhibiting the most significant differences are A_2, which is associated with price variables. There are also substantial departures from estimates using the full 13 municipios. For example, the estimates of β for the poorest municipios indicate an induced inventory effect which is statistically significant for household water demand. The estimated elasticities of poorer municipios indicate much greater responsiveness to price and income changes than that of the rich municipios. These rather significant differences between estimates of rich and poor areas suggest that models which assume constant elasticities of price and income grossly misspecify the nature of residential water demand.

---

1 Comparison of income elasticities with those found by Houthakker and Taylor is not strictly valid because they used total expenditures which was less than total income. On this basis, the higher income elasticity found by Houthakker and Taylor might be justified.

Table 4 . -- Estimated coefficients for rich and poor municipios

	A_0 Constant	A_1 q_{t-1}	A_2 Δx_t	A_3 x_{t-1}	A_4 Δp_t	A_5 p_{t-1}	SSE	R^2	β	γ	η	K_γ	K_η	E_{sx}	E_{sp}	E_{1x}	E_{1p}
RICH																	
Per capita																	
OLS	1110 (525)	0.490 (.359)	218 (204)	52.4 (106)	-2100 (1220)	-757 (889)	119.4	.27	-0.410 (1.01)	256.8 (278.6)	-0.23 (.107)	102.7	-0.119	0.217	-0.56	0.087	-0.29
EC	956 (428)	.495 (.361)	222 (185)	58.2 (106)	-2120 (1170)	-797 (874)	115.5	.33	-.375 (1.079)	257.9 (264.5)	-.23 (.101)	11501	-.128	.240	-.62	.107	-.34
Household																	
OLS	2790 (1800)	.818 (.298)	171 (231)	-30.2 (124)	-11700 (5260)	-2980 (4700)	520.4	.40	-.363 (1.39)	204.5 (248.8)	-1.12 (.358)	-165.4	-.50	.170	-.67	-.138	-.30
EC	2050 (992)	.604 (.322)	233 (218)	68.2 (139)	-10200 (4970)	4010 (4400)	581.0	.33	-.152 (1.23)	248.0 (262.7)	-1.03 (.386)	172.0	-.78	.174	-.52	.121	-.39
POOR																	
Per capita																	
OLS	-69.1 (204)	.632 (.099)	1140 (413)	-755 (271)	-500 (482)	1550 (548)	66.8	.93	-.948 (.477)	1859 (511.9)	-.156 (.048)	-2052	-.71	.781	-.43	-.863	-1.95
EC	-5.6 (85.9)	.544 (.138)	1190 (405)	-493 (395)	-518 (468)	1460 (542)	99.9	.77	-.934 (.488)	1855.8 (525.2)	-.161 (.05)	-1079.6	-.805	.733	-.44	-.45	-2.19
Household																	
OLS	-7.34 (895)	.528 (.117)	1190 (396)	-489 (204)	-2420 (2070)	6500 (2280)	288.5	.88	-.959 (.493)	1873.9 (523)	-.742 (.224)	-1036.9	-4.6	.524	-.71	-.290	-4.34
EC	23.1 (502)	512 (.129)	1200 (393)	-431 (248)	-2460 (2040)	6380 (2280)	328.9	.79	-.950 (.490)	1872.8 (521.5)	-.747 (.224)	-883.3	-4.7	.526	-.71	-.248	-4.50

OLS = ordinary least squares method of estimation

EC = error components method of estimation

SEE = standard error of estimate

R^2 = *coefficient of determentation*

β, γ, η, K_γ and K_η as defined in the text

Numbers in parenthesis are standard deviations of the coefficients

E_{SX}, E_{LX} = Short and long-term income elasticities

E_{SP}, E_{LP} = Short and long-term price elasticities

The criterion of performance of alternative demand models and estimation procedures for forecasting water demands should be determined by the situation in which the projections will be used. In particular, the decision-maker's utility or loss function should be specified. The structure of the utility function would indicate the relative losses or tradeoffs between increased bias or smaller projection variance. Because the purpose of this report is to present procedures of projecting water demand, it would not be appropriate to select a method of projection. If the decision maker was very risk averse or if the relative loss of system overdesign were small relative to economic losses resulting from potential water shortages, the decision maker would probably choose the demand function which indicated the largest responses to income and population growth, that is, the error components procedures. The opposite might be true if economic losses of system overdesign were large compared to potential water shortages.

Projection with a static model is somewhat routine because the value of the predictors have only to be substituted into the equation. The dynamic model, however, provides a means for incorporating the most recent realization of the dependent variable as part of its initial condition. The effect of including the lagged dependent variable is to lessen the influence of the stochastic elements of the other predicted explanatory variables. As Houthakker and Taylor (1971) have indicated, however, it is not possible to provide an estimatable closed form expression for the standard error of projection with dynamic models (not just this dynamic model)[1]. Because prediction errors accumulate in the dynamic model due to the recursive nature of the method of projection, the actual variance of the dynamic model projection is more

---

[1] This argument follows the one provided by Houthakker and Taylor (1971). To illustrate why the projection variance cannot be estimated, consider the following general dynamic model;

$$y_{t+1} = \psi + \lambda y_t + hx_{t-1} + u_{t-1}$$

If $x_{t+1} + h$ for all t then the general solution for the above equation can be be written

$$\hat{y}_{t+n} = \psi(1 + \lambda + \lambda^2 + \ldots \lambda^n) + \lambda^n y_t + h(\lambda^n + 2\lambda^{n-1} + 3\lambda^{n-2} + \ldots + (n-1)\lambda + n)\, hx_t + \sum_{i=1}^{n} \lambda^{n-1} u_{t+1}.$$

Projections for y_{t+n} will be

$$y_{t+n} = a(1 + b + b^2 + b^n) + b^n y_t + c(b^n + 2b^{n-1} + \ldots + (n-1)b + n)\, hx_t$$

sensitive to the length of the projection period than a corresponding static model. Another measure of projection error which also does not appear for the static model is the degree of model misspecifications. Insofar as the dynamic model more properly characterizes consumer behavior, for a given fit of the sample data, the actual projection variance of the static model is understated relative to the dynamic model. After considerable experimentation with alternative forms of parameter restrictions, the models used in this study were chosen on the basis of conformance to theory, statistical fit, and the standard error of the estimate.

The forecasting performance of the two estimation procedures was measured by splitting the sample, reestimating the parameters and by comparing the departure of the predicted values of the dependent variable. In particular, the data for 1960-69 were employed to reestimate the equations while the 1970 observations of the 13 municipios were compared with the predicted values of 1970. Forecast performance was measured by the computed value of R^2 where

$$R^2 = 1 - \frac{\sum e_i^2}{\sum (q-\hat{q})^2}$$

where e_i is the residual difference of the actual q and predicted values. The Thiel coefficient U is a statistic which measures the goodness of fit of a set of forecasts with specific realized values and is defined as

$$U = \frac{\{\sum_i (P_i - A_i)^2\}^{1/2}}{\sqrt{\sum_i P_i^2} + \sqrt{\sum_i A_i^2}}$$

(Continuation of footnote 1, page 19)

The variance of projection is then given by

$$E(\psi_{t+n} - \hat{\psi}_{t+n})^2 = E\{(a-\psi)^2 + (ab - \psi\lambda)^2 + \ldots + (ab^n - \psi\lambda^{n})^2 = (b^n - \lambda^{n})^2\} y_t^2$$

+ other terms and the cross product terms

In the final equation, the terms involving the powers of the coefficients do not appear to be estimatable for the sample data utilizing the traditional tools.

where P_i is the predicted value and A_i is the actual value. Thiel U must be between zero and one where zero denotes a perfect forecast and one denotes no forecasting ability (Thiel, 1961). Although the forecast performance of the ordinary least squares estimators were significantly better than the error components model; both procedures appeared adequate for short run forecasts.

Table 5.--Forecast performance of alternative residential water demand models

	OLS model		Error components model	
	Thiel U	R^2	Thiel U	R^2
Household income	0.048422	0.86376	0.2001	0.43775
Per capita income consumption	.046077	.90996	.19786	.45695

In choosing a particular forecasting model Houthakker and Taylor (1971) were guided by a number of considerations including fit of the model and signs of the model coefficients. In fact, for the Houthakker and Taylor study there were instances when the problem of autocorrelation was ignored and ordinary least squares estimates were used for projections because of the higher explanatory power of the least squares regression equation.

The estimated demand model for the San Juan region was employed recursively to project per capita residential water demand on the basis of alternative growth rates of income and prices. These projections are used with the alternative assumptions about population growth, extension of services to areas not now served, and substantial changes in the mix of metered versus unmetered customers, to calculate total municipio residential water demand. The observed value of q_{1970} is utilized as one of the initial conditions of the projection. The 1970 observation was forced to lie on the regression line by adjusting the intercept term for each cross-sectional unit.[1] This procedure had essentially the same effect as imposing

[1] This might have also accounted for the poorer predictive performance of the error components model.

an additive constant dummy variable for each cross-sectional unit.[1] If data for individual municipios were based on longer time series (25 or 30 years), demand equations could be developed for each without losing a large proportion of the available degrees of freedom.

The most important feature of utilizing conditional economic models for projection purposes is that such models facilitate sensitivity analysis of the projections by permitting the systematic variations in underlying economic and demographic assumptions. In Appendix B several sets of projections are presented which were generated by varying the growth rates of per capita income, water prices, population, and policies relating to the extension of service areas and water metering.

A subset of the projections given in Appendix B are presented in table 6. Comparison of projection set (1) with (2) and (3) illustrates the relative sensitivities or insensitivities of the projections to the changes in the rate of price increases from 1 to 3 percent. It is evident that a large proportion of the differences between (1) and (4) may be attributed to the difference in population growth rates. Of course, it is highly unlikely that all the municipios in this area would experience annual growth rates of 5 percent in population over a 20-year period. As might be expected from the low income elasticities reported in table 3 the projections are relatively insensitive to changes in income. However, a comparison of these sets of projections serve to illustrate the relative responsiveness of water demand to underlying economic and demographic assumptions.

Further comparison of these projections with those found in the report by the Puerto Rico Aqueduct and Sewer Authority (1969) should be very limited. While the per capita water use information (data) for the year of 1965 is in basic agreement (see table 1, page 1), the projections may diverge substantially because of different assumptions about population growth and the basic determinants of residential per capita water use. Moreover, earlier projections did not have the advantage of the data generated from the 1970 Census of Population.[2]

Caution must be used in specifying alternative sets of assumptions about annual growth rates of economic and demographic variables because such rates have compounding effects. That is, while a 3 percent annual growth rate in population may seem small, over a 20-year period the cumulative impact is substantial. Moreover, for the relatively short time series

---

[1] Houthakker, Verleger and Sheeham (1973) employed the error components estimation procedure on a dynamic model used to forecast demand for gasoline. They implicitly adjusted the forecast models for cross-sectional variations by regionalizing the regression models. Houthakker and Taylor (1971) discuss this procedure when the observed lagged dependent variable is not subject to revision or is known to be without measurement error.

[2] For example, population estimates for Loíza and Río Grande for 1970 were 32,500 and 23,000; respectively, while actual census figures were approximately 38,700 and 21,900 (Puerto Rico Aqueduct and Sewer Authority, 1969).

Table 6a .— Illustrative projections for residential water demand (In million gallons per month). See table 6b for assumptions on forecasts 1 to 4

Municipios	1975				1980				1985				1990			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Bayamón	287	288	280	257	348	354	340	281	444	435	412	306	504	540	503	332
Caguas	132	132	128	118	162	165	157	130	212	207	193	144	241	262	239	158
Carolina	135	136	133	121	176	182	175	142	244	239	227	164	284	316	295	187
Cataño	39.3	39.4	38.5	35.2	49.3	50.3	48.3	39.8	65.0	63.6	60.1	44.4	75.0	80.9	75.1	49.3
Ceiba	10.7	10.7	10.4	9.6	14.0	14.3	13.7	11.3	19.1	18.7	17.5	13.0	22.7	24.4	22.5	14.9
Fajardo	35.2	35.3	34.2	31.4	43.2	44.2	42.0	34.8	56.7	553	51.7	382	64.3	70.0	64.1	42.0
Guaynabo	147	148	145	132	180	186	180	146	239	235	226	161	269	300	285	178
Loíza	46.8	46.4	44.9	41.7	58.9	58.6	55.5	47.4	74.4	723	668	52.1	87.6	88.9	79.9	57.0
Luquillo	12.6	12.6	12.1	11.3	15.2	15.3	14.3	12.2	19.4	18.7	17.0	13.1	22.0	23.0	20.3	14.2
Río Grande	18.7	18.7	17.9	16.6	53.6	24.0	22.3	18.8	32.2	31.0	280	21.2	37.0	40.4	35.4	23.9
San Juan	990	1013	993	889	1241	1323	1282	1006	1770	1743	1673	1137	1948	2326	2214	1289
Toa Baja	79.9	79.5	74.5	71.6	96.7	96.5	92.6	78.1	118	116	109	84.1	167	140	129	90.3
Trujillo Alto	34.9	34.7	33.6	31.2	44.2	44.0	42.0	35.6	56.6	55.2	51.5	39.8	67.6	69.0	62.7	44.0

Table 6b.--Assumptions used to make forecasts 1 to 4

Forecast	Per Capita Income Growth (percent)	Population Growth (percent)	Growth in Prices (percent)
1	3	3	0
2	5	3	1
3	5	3	3
4	3	1	1

Also all projections assumed a 3 percent reduction in areas not served which was made up by additional metered customers.

used to estimate the models, the absolute value of the economic variables will rapidly be outside the sample range experienced from 1960 to 1970. As will be seen later, this problem is particularly troublesome in the manufacturing and commercial water use projections. Because of the very high price elasticities found in these sectors for later time periods, even a 3 percent annual rate of increase in price may result in negative amounts of water use being predicted by the models, which, of course,is nonsense.

Caution must also be exercised when interpreting the projections. Because it was not possible to provide an analytical form for classical confidence bands for the projections and recalling that the variances are cumulative with the dynamic model, a relatively high degree of uncertainty should be attached to projections at the end of the projection period. Monte Carlo studies performed by Houthakker and Taylor (1971) on dynamic demand models of this type suggest that forecasting variances increase sharply from the sixth period[1] thereafter. From this latter observation, the importance of a planned program of model updating can be inferred. Model updating will be discussed later.

[1]The relevant time period is a year if the sample data are annual.

Commercial Water Demand

Nature of Commercial Water Demand

Water purchased by commercial (non-manufacturing and non-governmental establishments including construction and mining, transportation, commercial trade, and financial and service economic sectors), is conventionally defined by Commonwealth agencies as commercial water use. With the exception of mining and thermal electric power generation[1], water is used by these establishments for cleaning and sanitary purposes for workers, customers, and machinery. However, Puerto Rico presently has no active mining industry and thermal electric power generation is vested in a Government owned corporation, not in a public utility. Because water is not an integral part of production of services and there are limited substitution possibilities, commercial water demand at the municipio level is principally determined by the mix and level of economic activity. Generally, water used by commercial establishments must be potable.

By nature the commercial sector serves the surrounding community and is itself primarily determined by income generated from manufacturing, agriculture, and other basic industries. Without a mechanism linking population, manufacturing, and other activities to development of a commercial or secondary economic sector, there is little direct application of these models to planning decisions. That is, planning decisions which affect industrial and population concentration also influence the location and growth of economic activity, and, thus, they will influence commercial water demand.

Because Commonwealth agencies aggregate water use for non-manufacturing and non-governmental establishments, individual commercial sectors do not have separate water-use statistics. There are two approaches which may be used to relate commercial water use to economic activity. The water demand may be related to aggregate commercial economic activity or to an index or constructed variable which also reflects the relative magnitudes of the individual economic sectors. The particular index variable constructed for this study was based upon a principal component analysis of the data from the San Juan test area (see Appendix C). Results utilizing both the aggregated income and the principal components measures of commercial economic activity are presented. Because water is not a basic part of the production process there is no theoretical basis for deriving a functional

[1]Public utilites are generally classified under the transportation sector.

relationship for demand. In this absence, the demand relationship estimated had the following function form;[1]

$$q_t = A_o + A_1 q_{t-1} + A_2 \Delta y_t + A_3 y_{t-1} + A_4 \Delta p_t + A_5 p_{t-1}$$

q_t = commercial water use per customer

y_t = aggregate municipio commercial income or the first principal component of the individual sector incomes

p_t = per unit water price

---

1 A similar functional form was applied by Balestra and Nerlove (1966) in the joint analysis of residential and commercial demand for natural gas. The firm was rationalized on the basis of indicating the effects on demand of a fixed technology set of appliances for household and commercial users.

Analysis of Empirical Results and Forecasts

The use of principal components analysis for this specific set of data was not particularly productive. While the data were highly colinear and the first principal component explained 97 percent of the variation, the characteristic vector assigned nearly equal-weights to income generated from each sector. This had the approximate effect of multiplying the sum of the individual sector income by a constant. Under a different circumstance the use of principal components analysis would provide a means of tracing through the effects on the constructed index of unbalanced growth on individual economic sectors. Results of the regression equation are presented in table 7.

The coefficients of the estimated demand equation for commercial water use exhibits expected coefficient signs. That is, income coefficients are positive and price coefficients are negative. Because information relating to the total number of establishments for each sector was not available, commercial income was left in its aggregate form. Calculated income and price elasticities for the two estimated equations are found in table 8.

Price and income elasticities shown are also of the expected sign. Income elasticities are relatively low and the price elasticities are unrealistically high. This appears to be the result of a limited range of variations in municipio income and the aggregation of water demands for extremely diverse water users such as banks, laundries and car washes. One would expect a high degree of price responsiveness for heavy water users (such as laundries and car washes) because they are likely to institute water reuse and other water saving systems. For example, restaurants may forego serving water to customers unless they ask for it and institute water saving procedures for dishwashing. Although one would expect such establishments to account for a large proportion of commercial water use there is no way to disaggregate their demand from other users with available data.

The forecast performance of the commercial water demand equations were established by repeating the procedure outlined in the section on residential water demand. That is, the sample data was split to cover the period from 1960 to 1969 and the coefficients were reestimated for this period. Forecasts generated from the reestimated model were compared to the actual 1970 observations. Table 9 presents the forecast performance of the commercial water demand model. As shown in table 9, the performance of the model utilizing ordinary least squares produced significantly better forecasts than the error components model.

Table 7.--Coefficients of commercial water demand per customer

	A_0 Constant	A_1 q_{t-1}	A_2 Δy_t	A_3 y_{t-1}	A_4 Δp_t	A_5 p_{t-1}	R^2	SEE
Commercial OLS	16200 (2700)	0.718 (.047)	0.00454 (.0152)	0.00245 (.00206)	-57100 (6630)	-31800 (6030)	0.89	2529.5
EC	7080 (1160)	.573 (.068)	.00751 (.0147)	.00184 (.00302)	-52300 (6880)	-28500 (1160)	.55	3578
Commercial principal components								
OLS	162000 (2480)	.718 (.047)	.0075 (.0356)	.0057 (.0048)	-59300 (7040)	-31700 (6460)	.88	2601.4
EC	7350 (1240)	.576 (.071)	.0144 (.0035)	.0044 (.0069)	-54700 (7310)	-28800 (6790)	.56	3559

Standard deviation in parethesis

OLS = ordinary least squares method of projection.
EC = error components method of projection.
q_t = commercial water use per customer (thousand gallons per month).
y_t = aggregate municipio commercial income or the first principal component based on individual sector commercial income, respectively
p_t = per unit water price.

Table 8.--Calculated elasticities

Total income	Short run		Long run	
	Income	Price	Income	Price
OLS	0.0167	-1.08	0.0377	-2.30
EC	.1006	-3.02	.0515	-2.23
Principal component				
OLS	.0171	-1.08	.0376	-2.21
EC	.1031	-3.02	.0520	-2.22

Table 9.--Forecast performance of commercial water demand models

	OLS Model		EC Model	
	Theil U	R^2	Theil U	R^2
Aggregated commercial income	0.0486	0.9400	0.3977	0.6029
Commercial principal components	0.0486	0.9400	0.3975	0.6045

High price elasticities present a real problem in forecasting when assumed annual growth in prices results in prices which are well outside of those of the sample information. For example, with an annual price increase of 3 percent or more per customer, demand may become negative for forecasts in the later part of the projection period. Alternative sets of projections for total municipio commercial water use are presented in Appendix B. The projections were first based on the estimated water demand per customer (using the model relating to total commercial income) and on assumptions relating to commercial income and prices. The individual customer demands are then totaled assuming different growth rates of the number of customers and different metering policies. Because the individual firm demands were quite sensitive to pricing, the variations in pricing policies had to be somewhat restricted.

Industrial Water Demand

Nature of Industrial Water Demand and Previous Analyses

Industrial water use is defined by the Puerto Rico Aqueduct and Sewer Authority as water purchased by manufacturing establishments, Industrial or manufacturing processes require water for cooling, washing and transportation of waste materials. For detailed sensitivity analysis, estimation of water demand should be carried out on a plant and process basis. Such information is generally not available to the planner and even if it were, one would have to aggregate individual plant demands over many establishments. Industrial water users frequently, by the purchase of water rights, have the option of developing their own sources of water or using municipal supplies. The effect of self-supplied firms on demand projections will vary according to the nature and extent of water use for the industry. However, the firm's use of rational-internal accounting procedures would suggest that self-supplied firms assign values to their inputs at the going price of purchased water and make the same cost calculations. Adjustments can therefore be made in the proposed estimation techniques which would normalize the data to account for self-supplied sources. The following discussion first considers existing techniques of forecasting industrial water use. Special emphasis is placed on input-output forecasting techniques because of the availability of a recently constructed regional input-output model for the Island. This discussion provides a detailed description of how one might employ the input-output model for industrial water use projections. Because one of the purposes of this report was to provide procedures for projecting water demands as input to the Puerto Rico Planning Model, input-output models could not be used becuase the regional delineation therein was inappropriate from the perspective of the natural hydrology of the Island.

At the firm level the quantity and quality of water for manufacturing processing are determined by input demand relations. Water use at the plant level has been studied for several major industries. These include steam electric generation (Cootner and Lof, 1965), the beet sugar industry (Lof and Kneese, 1968), the paper pulp industry (Bower, et al, 1971) and petroleum refining (Russell, 1973). While such studies are important in defining the economically feasible region of production for water as an input, their usefulness to regional planning at present is limited. Not only must individual firm demands be aggregated but plant-level data must be developed for many other industries.

Other studies which have considered industrial water demand include those of Tate and Robichaud (1973), De Rooy (1970), and the North Atlantic Regional Study of the U. S. Corps of Engineers (1972). Tate and Robichaud (1973) propose that industrial water demand projections for Canada be generated by an industrial dynamics simulation model. Proceeding on a somewhat ad hoc basis a set of equations are specified for aggregate industrial (manufacturing) water demand. Parameters of the model are subjectively estimated and model predictions are compared to historical data for 1970, the only year in which the water use data were available. De Rooy (1970) explicitly considered water as an input of production.

The North Atlantic Regional Study (1972) employed an input-output approach to project industrial water demand for that region. A regionalized input-output matrix computed from the National U.S. Table was used as a basis for economic projections. While this study could not tie the projection procedure to rigid regional delimitation, the following discussion is designed to suggest how the Commonwealth's regional input-output model (Planning Board, 1970) could be usefully applied to water resource planning.[2]

Suppose the matrix of interindustry transactions is represented

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1j} & \cdots & x_{1n} \\ \vdots & & \vdots & & \vdots \\ x_{i1} & \cdots & x_{ij} & \cdots & x_{in} \\ \vdots & & \vdots & & \vdots \\ x_{n1} & & x_{nj} & & x_{nn} \end{bmatrix}$$

where a row denotes the selling industry and a column represent a buying industry. Direct input or technical coefficients per dollar of output is given by

$$a_{ij} = \frac{x_{ij}}{X_j} \qquad i, j = 1, \ldots, n$$

where $X_j = \sum_{i=1}^{n} x_{ij} \qquad j = 1, \ldots, n.$ Then

[1] Although Turnovsky (1969) estimated an industrial water demand relation, his purpose was to examine the responsiveness of firms to uncertain supplies during a specific drought period.

[2] While the following discussion is highly abbreviated, an elementary explanation of the economic basis and application of input-output analysis can be found in Yan (1969)

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & & a_{ij} & & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & & a_{nj} & & a_{nn} \end{bmatrix}$$

Vectors denoting the final demand Y and total output X for the system are

$$Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_i \\ \vdots \\ Y_n \end{bmatrix} \qquad X = \begin{bmatrix} X_1 \\ \vdots \\ X_i \\ \vdots \\ X_n \end{bmatrix}$$

Net output is computed as gross output minus intermediate use or

$$X - AX = (I-A)\,X = Y\,.$$

The solution of the system is

$$X = Y\,(I-A)^{-1}$$

where $(I-A)^{-1}$, the Leontief inverse matrix, provides the direct plus indirect input requirements of each industry per dollar of output of final demand. Suppose one defines a diagonal matrix of water use coefficients whose elements are interpreted as the quantity of water withdrawn by individual economic sectors per dollar value of the sector's gross output. Premultiplication of the Leontief inverse matrix by the water use coefficient matrix provides the unit of water volume of output of industry j required by industry i for its delivery of one monetary unit of production to the final demand sector, that is,

$$\left[I - A\right]^{-1} \begin{bmatrix} b_{1.} & & & & & \\ & \cdot & & & 0 & \\ & & \cdot & & & \\ & 0 & & b_{i.} & & \\ & & & & \cdot & \\ & & & & & b_{n} \end{bmatrix} = \begin{bmatrix} w_{11} & \cdots & w_{1n} \\ \vdots & & \vdots \\ w_{i1} & \cdots & w_{in} \\ \vdots & & \vdots \\ w_{n1} & & w_{nn} \end{bmatrix}$$

The row -sums of the table forms a vector of direct plus indirect freshwater requirements per unit output. By projecting future final demands and assuming technical coefficients remain the same, the analyst can work backwards to obtain projected water requirements (Lofting and Davis, 1968). Generalized technical water use coefficients are available for selected industries in U.S. Bureau of Census (U.S. Bureau of Census, 1966) and U.S. Bureau of Mines (Kaufman and Nadler, 1966) publications.

Two significant objections to this technique are: the implicit assumption that technical conditions are fixed over the entire planning period and that pricing policies do not influence water demand. While these objections remain valid, the availability of state, regional or regionalized versions of the published national input-output model make this method of projection easy to apply.

Empirical Analysis and Forecasting Procedure

Analysis of input demand at the plant or firm level entails examination of various factores including input substitution possibilities,[1] input price levels (price of the input itself, and prices of complementary and competitive inputs) and factors determining market demand for the firm's output. The firm's input demand function may be derived from its production function given that this function was known and a set of market conditions was assumed. Even if this was known for individual firms, aggregating these demands to a regional level might introduce substantial error into the input demand functions. Furthermore, traditional formulations of input production relationships frequently consider only the capital and labor inputs. Although there is a substantial amount of research underway attempting to make conceptual production models more realistic, little of this has been applied to empirical situation.

---

1 In technical terms input substitution is characterized by the elasticity of substitution between inputs.

Productive inputs are often classified according to whether they are flow or stock quantities. Stock inputs include real capital goods such as plants and machinery while flow inputs include labor services, raw material and energy. Data relating to even the basic economic inputs of capital, labor, and other flow inputs such as water, materials, and energy are normally available only on a highly aggregated basis and usually include substantial measurement errors. Such problems have, of course, been encountered in (other) studies related to the demand for electricity (Fisher and Kaysen, 1962) and natural gas (Balestra and Nerlove, 1966). Perhaps the most useful way to proceed is to discuss several intuitive and somewhat simplistic model formulations.

At a given point in time and for a fixed technology, it might be argued that water as a production input is directly proportional to output and in the short-run is characterized by limited substitution possibilities. Then for the ith good of the jth firm

$$q(t)_{ij} = A_{ij} + BQ(t)_{ij} + u(t)$$

where $q(t)_{ij}$ is the rate of water used by the jth firm for production of good i, and $Q(t)_{ij}$ is a measure of the rate of output. Aggregation must then be carried out over products for a multiproduct firm and over establishments. Under fixed technology an increase in the price of water will be transmitted to the output sector through increased output prices and reduced output, that is, the relative price of water intensive goods will increase relative to non-water intensive goods. In the case of the multiproduct firm, outputs may be adjusted in the direction of the less water intensive goods. For firms which develop their own supplies, it would be rational to value the inputs at the going market price[1] (Fisher and Kaysen, 1962). Thus, the quantity of water demanded by the firm will be inversely related to the market price. For $q(t)_i$ the rate of total firm water demand then is

$$q(t)_i = A_i + BQ(t)_i + \pi_i p(t) + u(t) \tag{10}$$

where $\pi_i < 0$ and p_i is price.

Suppose individual plants were classified according to capacity, production process and nature of output. For a particular class, I, equation (10) would then become

$$q(t) = NA + BQ(t) + \pi(t) + u(t) \tag{11}$$

where $q(t) = \sum_i q(t)_i$, $Q(t) = \sum_i Q(t)_i$, $p(t) = \frac{1}{n} \sum_i p(t)$ for all i in I

and where B and π are appropriately adjusted to account for heterogenous production processes. These individual classes may, of course, be aggre-

[1] If cost conditions for self-supplied water are much lower than prices of public water supply, it might be rational to sell water to other manufacturing establishments.

gated further to the industry level by using weighted averages from the individual classes. While it would not be difficult to estimate the above equations on an industry basis, historical-industrial water use data for industrial industries were not available in a machine-readable form. Therefore, one should only expect that relationships discussed are valid in an approximate fashion. The problem of output measurement is complicated by the fact that the value of output may not reflect the value added by the individual firms. Seasonal fluctuations in plant employment levels and the intensive use of part-time labor suggests that average plant employment may not be an appropriate measure of value added. Data on average plant production man-hours (AMH) were used to measure the value added (Q) by the manufacturing establishment.

With the short-time series available, the data could not be used to examine the influence of technical change on industrial water demand. With this restriction, the analysis was carried out under the assumption of fixed technology. In order to capture part of the short-run dynamic adjustments which result from changes in output and water prices, a dynamic model was employed. Because average plant capital measures were not available, the local wage rate was included to reflect the degree of regional plant mechnization. Short-run dynamic plant water demand was assumed to be characterized by the following expansion of (11)

$$q_t = B_o + B_1 q_{t-1} + B_2 \Delta AMH + B_3 AMH_{t-1} + B_4 \Delta p_t + B_5 p_{t-1} + B_6 w_{t-1} \qquad (12)$$

Because the full sample parameter estimates were somewhat ambiguous, the sample was split into two groups. Group I[1] included areas that were characterized by labor intensive and light manufacturing processes while the second area appeared to be characterized by heavy manufacturing and greater capital intensity in plant processes. Table 10 indicates the

Table 10.--Characteristics of Group I and Group II

	Average monthly water use (thousand gals)	Weekly average wages (dollars)	Average plant man-hours
Group I	100,247	$36.03	1648
Group II	317,876	42.89	1453

[1]Group I includes the following municipios: Caguas, Fajardo, Loíza, Luquillo, Río Grande, Toa Baja, Trujillo Alto; and Group II includes: Bayamón, Carolina, Cataño, Guaynabo, and San Juan.

extent of the variation in the two groups. Municipios in Group I are characterized by lower water using industries with greater labor intensity while the municipios in Group II are characterized by industries which are more water and, probably, capital intensive. Separate regressions were estimated for the two regions and the parameter estimates presented in table 11 and 12. In order to increase efficiency of the parameter estimates, the regression coefficient associated with average plant man-hours was restricted to be zero. Parameters of the regression equations indicate substantial differences in the industrial water demand relationship.

Calculated short-term price and output elasticities are presented in table 13. As expected, the price and output elasticities for Group II (High Water Using) are greater than those for Group I (Low Water Using). While the grouping of the municipios was somewhat arbitrary and might have been further refined, the regression results suggested that the groups are substantially different.

The forecast performance of the demand equation was investigated using the regression estimated from a shorter sample period and generated predictions were compared to the 1970 data. The Thiel U coefficient ranged from 0.117 to 0.296 implying reasonable performance. Experimentation with other models and alternative sample periods indicated that forecast performance is rather sensitive to the length of the sample period.

Projections for industrial water demand are presented in tables (B-5) and (B-7) in Appendix B. These were based on alternative assumptions relating to growth in prices, output, wages and number of customers. The specific assumptions regarding these variables are presented with the projections. For example, by comparing projection set (1) and (2) the influence of alternative growth rates in prices may be observed. Because projections are based on a short-term model, caution must be used in interpretation. The short-run model does not capture long-run adjustments in capital; thus, projections are conditioned on constant technology and capital equipment.

Because the Puerto Rico Aqueduct and Sewer Authority supplies only potable water to industrial users, industrial firms may develop their own lower quality water at a cheaper rate since industrial process water generally does not have to be potable. Therefore, the projections are limited by the range of choice for water quality within the sample information. Finally, the projections do not take into account the rather substantial influence on the amount of water intake which effluent standards and charges are likely to have in the future. With these qualifications, models can really not be used for projections for more than 5 years in the future. For longer periods, procedures for developing technological forecasts need to be included. However, model updating does provide a means whereby long-term adjustments can be taken into account by the model through better sample information. The importance of periodic model updating and refinement is discussed.

Table 11.--Estimated coefficients for industrial water use for high-water intensive municipio of San Juan Region

	B_o Constant	B_1 (q_{t-1})	B_2 (ΔAMH)	B_3 (AMH_{t-1})	B_4 (Δp_t)	B_5 (p_{t-1})	B_6 (W_{t-1})	R^2	SEE	β	γ	η	K_γ	K_η
OLS	122000 (121000)	0.818 (.104)	44.3 (48.6)	15.5 (38.4)	-2870000 (518000)	-578000 (520000)	1210 (1580)	0.669	67,263	0.225 (1.201)	40.238 (51.957	-283.405 (48.768)	85.428	17344
EC	169000 (62000)	.591 (.126)	65.4 (46.4)	11.7 (47.0)	-3350000 (487000)	-2370000 (674000)	4900 (1880)	.749	86,221	-.3169 (8337)	74.800 (53.07)	-272.202 (48.986)	28.624	-210.994

Table 12.--Estimated coefficients for industrial water use for low-water intensive municipio of San Juan Region

	B_o Constant	B_1 q_{t-1}	B_2 ΔAMH	B_4 Δp_t	B_5 p_{t-1}	B_6 (W_{t-1})	R^2	SEE	β	γ	η	K_γ	K_η
OLS	103000 (28400)	0.511 (.109)	506 (7.90)	-317000 (86800)	-352000 (83600)	1530 (562)	-	25468	-0.647 (.375)	6.699 (10.477)	-18.743 (10.827)	-	-14.865
EC	67500 (15300)	.313 (.114)	5.40 (7.23)	-324000 (82100)	-436000 (87600)	1990 (567)	.576	27688	-1.046 (.391)	8.218 (11.055)	-16.152 (11.455)	-	-12.876

OLS = ordinary least squares estimation method
EC = error components estimation method
q_t = quantity of manufacturing plant water demand
AMH = average plant man hours
p_t = per unit water price
W_t = wages (weekly) in dollars
SEE = standard error of estimate (same units as q_t)
Numbers in parenthesis are the standard deviations of the regression coefficients
Note: B_3 was restricted to zero

Table 13.--Price and output elasticities for industrial water use

		Price	Output
Group I	Low water using region		
	OLS	-0.229	0.0568
	EC	-.34	.012
Group II	High water using region		
	OLS	-1.33	.134
	EC	-1.06	.206

Updating the Demand Models

Along with the acceptance and implementation of the preceeding methods for projecting water demand, provision should be made for updating the models. In particular, as new sample information becomes available this information can be utilized to update the estimate of the model coefficients and thereby provide a basis for improving the model forecasts. Moreover, because the dynamic models are recursive in nature, forthcoming information would eliminate forecast errors that result from employing predicted values of the endogenous variables for making forecasts for future periods. Previous experience[1] both in this study and elsewhere, with split sampling and the error components technique of combining cross-sectional and time-series information suggests that additional sample information significantly improves the model fits and efficiency of the parameter estimates.

Econometricians have only very recently become aware of the potential advantages of using forthcoming sample information to update or revise the parameter estimates of econometric models. As evidence of this activity, a new body of literature is developing which addresses the question of finding an optimal way of combining revised estimates with the previous results and sample information. Theoretical questions related to the applicability and power of classical statistical tests with revised estimates are addressed in several recent papers by Cooley and Prescott (1973a, 1973b, 1973c) and Duncan and Horn (1972). Although there are still many unanswered questions, this flurry of activity will eventually result in the formalizing of decision rules and estimates for combining new sample information with previous estimates. However, for the present, additional sample data in the context of the proceeding demand models may be incorporated by simply estimating parameters of the equation with the larger sample. The benefits of these updated estimates include improved statistical properties of the estimates providing a more accurate and reliable basis for hypothesis testing.

Practical steps which would significantly facilitate the updating of the models includes keeping all billing records in computer-readable form (a great deal of time, approximately 15 man-years, was spent developing the present data base). Secondly, industrial and commercial customers of Puerto Rico Aqueduct and Sewer Authority could be classified by SIC code. Industrial water use information which is classified

[1] More specifically, the residential and commercial water demand models used to test the forecast performance omitted the 1970 observations. When these models were compared with the models estimated from the full samples, the full sample models were significantly better in terms of fit and the efficiency of the parameter estimates.

by SIC code would prove valuable to local planners involved in local zoning decisions by providing a gross estimate of local water system demands if municipal supplies are used for processing. Thirdly, the Commonwealth could attempt to improve their estimates of the amount of self-supplied water use by location and use. Finally, estimates if system leakage could be made and associated with specific geographical areas.

ECONOMIC IMPACT OF WATER-RESOURCES INVESTMENTS

FOR PUERTO RICO 1960-68

Perspective of Analysis

Public works water resources investments frequently have been advocated as stimulants for regional industrialization and development. It has been argued that such investments may induce long term growth by injecting large public expenditures into the region and by increasing the total water available. Because empirical studies presented in the literature relate principally to the United States, it was felt that the recent historical experience of Puerto Rico should be examined before recommendation for development of additional analytical tools (such as regional econometric models) is made. Because future water demand is dependent on economic growth, it is of interest to examine the influence of water resources investments on growth. If there are substantial growth effects, the additional models would aid in reducing the uncertainty of such effects. However, if the growth effects are limited, then there is little likelihood such investments will induce substantial changes in municipal water demands. On the basis of this investigation, it would appear that the large scale modeling efforts aimed at assessing the secondary benefits of water resource investments may at this time be of limited value to planners.

This section of the report initially describes the economic setting and nature of the data. After presenting a theoretical framework, several sets of empirical results are presented and interpreted.

Economic Setting

Puerto Rico has developed rapidly since World War II and now exhibits a rather diversified economy. However, growth patterns and present levels of economic development are quite disparate across the Island. The San Juan metropolitan area along with the cities of Mayagüez on the west coast and Ponce on the south coast represent the major urban areas on the Island. Incomes in these areas are comparable to those of the rural United States. However, in 1968 per capita income for the majority of the interior rural municipios was less than $500 per year. This discrepancy reflects variation in the economic base of the respective areas, resulting from the process of industrialization which generally proceeded from coastal port areas toward the interior and from Government subsidies to industries that were initially concentrated in the seaport cities.

During the period covered by this study, from 1960 to 1968, a substantial part of the Island's public investment was devoted to the development of water resources and electrical power supply. There have been considerable differences in the type and quantity of water available, particularly in several drier areas along the south coast. Municipal water supply is the responsibility of one agency (Puerto Rico Aqueduct and Sewer Authority) while irrigation water and electricity are administered by another single agency (Puerto Rico Water Resources Authority)[1] for the entire Island.

For the purpose of this study, the Island is divided into four regions: the San Juan metropolitan area, the Southwest Region, which includes the major cities of Mayagüez and Ponce, and the relatively[2] undeveloped areas of the Northwest Region and Southwest Region. All data for the analysis were on a municipio basis (which are analogous to counties in the continental United States) with the above regions containing 13, 20, 21 and 20 cross-sectional units (municipios), respectively.[3] The object of the exploratory regression analysis was to investigate the impact of alternative public investment patterns on regional growth within particular regions and over the Island. In the next section the theoretical framework of this study is discussed.

Theoretical Framework

The process and pattern of regional growth and industrialization in a developing area is intimately associated with new plant and industrial location decisions. For industries with water-intensive production processes,

---

1 In general, there is no differential or peak demand pricing of water over the Island. The water and electric supply authorities are Government controlled public corporations.

2 Table 14 indicates the exact regional delimitation employed. It might also be noted that these regions roughly correspond to regions established by the Planning Board in the construction of their regional input-output model (Puerto Rico, Planning Board - 1970). Because of extremely poor data resulting from a low level of economic activity, the municipios of Culebra and Vieques were omitted from the analysis.

3 In general, there is no differential or peak demand pricing of water over the Island. The water and electric supply authorities are Government controlled public corporations.

Table 14.--Regional delimitation for investment impact analysis

Region I

Bayamón
Caguas
Carolina
Cataño
Ceiba
Fajardo
Guaynabo
Loíza
Luquillo
Río Grande
San Juan
Toa Baja
Trujillo Alto

Region II

Adjuntas
Añasco
Cabo Rojo
Coamo
Guánica
Guayanilla
Hormigueros
Jayuya
Juana Díaz
Lajas
Las Marías
Maricao
Mayagüez
Peñuelas
Ponce
Sabana Grande
San Germán
Santa Isabel
Villalba
Yauco

Region III

Aguada
Aguadilla
Arecibo
Barceloneta
Camuy
Ciales
Corozal
Dorado
Hatillo
Isabela
Lares
Manatí
Moca
Morovis
Quebradillas
Rincón
San Sebastián
Toa Alta
Utuado
Vega Alta
Vega Baja

Region IV

Aguas Buenas
Aibonito
Arroyo
Barranquitas
Cayey
Cidra
Comerío
Guayama
Gurabo
Humacao
Juncos
Las Piedras
Maunabo
Naguabo
Naranjito
Orocovis
Patillas
Salinas
San Lorenzo
Yabucoa

available water may be an important factor in determining plant location, particularly if there are substantial temporal and spatial variations in the amount of water available for processing and waste handling. However, there are obviously other considerations involved in the plant location decision. Because one of the functions of economic theory is to identify relevant economic variables entering the decision process, it would be instructive to analytically characterize the long-run output position of the profit maximizing firm as formulated by Dhrymes (1964).[1]

Consider a profit maximizing firm which sells in monopolistic markets and produces in several locations. Let the demand function for the firm's products in the ith market be $p_i = f_i(Q_i)$ where p_i is the price and Q_i is the quantity offered with revenue R_i where $i = 1, \ldots, k$. Suppose that the firm produces in m locations with the plant cost function for the jth location described by $C_j = C_j(Q_j)$. The cost of transporting Q units to market location i may be written as $T_{ij} = t_{ij}(Q_{ij})$. Also let V_i be the quantity of goods sold in the ith market and π, the total firm profit. For the profit maximizing firm the objective function has the following form (Dhrymes, 1964).

$$\pi = \sum_{i=1}^{k} R_i(V_i) - \sum_{j=1}^{m} C_j(Q_j) - \sum_{j=1}^{m}\sum_{i=1}^{k} t_{ij}(Q_{ij}) \qquad (13)$$

subject to the constraints

$$V_i = \sum_{j=1}^{m} Q_{ij}, \quad Q_j = \sum_{i=1}^{k} Q_{ij}$$

Using Lagrangean methods, Dhrymes derives the following equation for the profit maximizing output from the first order conditions:

$$\frac{dR_i(V_i)}{dV_i} = \frac{dt_{ij}}{dQ_{ij}} + \frac{dC_j}{dQ_j} \qquad (14)$$

Equation (14) is given the following interpretation: for some output to be produced in the jth location and marketed in the ith market it is necessary that marginal revenue in the ith market be equal to marginal production costs for location j plus the marginal transport cost to the ith market. This expression (equation 14) is the long-run equilibrium condition with all factors variable in every location (Dhrymes, 1964). For each output

---

1 An alternative theoretical discussion of plant location decisions in relation to water resource development may be found in "Public Investment Impacts and Regional Growth" (Lewis, 1973).

level, costs at the jth location may be expressed as a function of input levels and input prices. Thus, the cost function for location j is

$$C_j = h\,(x_{1j},\ x_{2j},\ \ldots,\ x_{nj};\ p_{1j},\ \ldots,\ p_{nj}) \qquad (15)$$

where x_{ij}, $i=1, \ldots n$ represents input usage levels and p_{ij} $i=1, \ldots n$ input prices at the jth location. If an explicit production function was assumed, then presumably a cost function in terms of output levels, inputs and input prices can be specified. Parametric analyses of such a function within the framework of equation (14) permits investigation of the alternative output-location tradeoffs.

As suggested by equation (15) long-run locational decisions are rather complex, with production, transportation and market considerations being important factors. Public water resource investments may tend to reduce particular operating costs associated with developing and pumping water for processing and waste handling. Although many major industrial-water users may still develop their own supplies, the availability and pricing of public water supplies and public wastewater handling limits the internal price a firm would be willing to incur. Another factor not readily measurable is the reliability of the public system. In some cases an industrial operation will develop its own supply, even at greater cost, simply because the output from the public supply cannot be counted on, either in quantity or quality. Obviously, other factors also influence the cost of producing at location j, such as, the relative wage rate and relative prices of other production factors. In fact, Fuchs (1962) suggested that dominant factors inducing the redistribution of United States manufacturing facilities from 1929-62 were labor wage differentials and the availability of specific natural resources, although market demand considerations are probably the single most important variable.

Empirical Approach

Observations on the economic location problem for firms would suggest that investigations of regional development patterns which are primarily the result of industrial location decisions should incorporate the influences of labor availability, market locations, and prices related to other factors of production. In the following analysis, hypotheses are tested which relate alternative measures of regional economic development to water resource investment, non-public works investment, regional wage rates and regional location relative to specific markets.

Measures of municipio growth (for a given time period) include the change in income generated by the local manufacturing sector and the change in the municipio's share of annual Island income. While the first of these variables may be taken as a measure of the pattern of industrialization, the second indicates a change in the distribution of Island income. These growth measures were assumed to be explained by variables representing public investment, regional wage rates, and location relative to major markets. Hypotheses which were tested relate to the statistical significance of the explanatory variables and whether variables relate to growth measure from a regional or an Island-wide basis. In particular, several sets of results were generated whereby conditions were imposed on variable coefficients. These results suggested interpretations consistent with alternative hypotheses as to the influence of specific variables. The specific procedures for performing these tests are described in the next section. Proposed tests are somewhat different from previous empirical studies (for examples see Howe, 1967; Garrison and Paulson, 1971) in that the pattern of regional development is examined in an area where growth results principally from locational decisions. Moreover, these studies did not consider other factor prices such as transportation or labor costs in location analysis.

The model proposed to investigate the impact of alternative public investments takes the following form:

$$y_i = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + u$$

Where y_1 = change in annual municipio manufacturing income over the period of 1960 to 1968

y_2 = change in municipio share of Island income from 1960 to 1968[1]

x_1 = intercept dummy variables

x_2 = water supply and waste facility investment for 1960 to 1968[1]

x_3 = investment in electrical supply for 1960 to 1968[1]

x_4 = non-public works Government investment for 1960 to 1968

x_5 = municipio manufacturing wage rate for 1960[2]

x_6 = distance by road to nearest cargo port (San Juan, Ponce and Mayagüez)

u = random component

$\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6$ are regression coefficients.

---

[1] Includes investment in resource development facilities and distribution system.

[2] This variable was designed to reflect relative wage rate differentials and was somewhat arbitrarily chosen at this point to avoid the possibility that public investment might induce short-term fluctuations.

Variables x_2, x_3, x_5, and x_6 relate to costs of inputs and are rationalized on the basis of the previous discussion. Non-public works investment, x_4, is principally directed toward social overhead projects such as hospitals, schools and Government buildings along with a fraction going directly to investment in industry. This direct part frequently takes the form of building plants in certain locations and leasing them at relatively low rates in order to induce industrial firms to locate in an area. This variable was therefore included because, for reasons previously indicated, it represents a production externality to firms in the respective area. The empirical analysis must be considered as an exploratory investigation of the relative impact of the specific variables on the growth measures considered. In particular, the model specification is somewhat ad hoc because it does not explicitly consider the exact functional relationship between production inputs, multiproduct firm operations or the various reasons for immobility of local firms. Preliminary results (not shown) suggested that the immediate and short-term influences of the public investment variables were not significant in terms of explaining the measures of industrialization or growth. Periods initially examined include cross sections for a 1-year lag and 3-year time period. On the basis of these results, it was decided to investigate the hypothesis over a 9-year period, (1960-1968).

Along with ordinary least squares, regional regression equations were estimated jointly with subsets of parameter coefficients restricted across all regions. Unrestricted coefficients were interpreted to reflect differences resulting from distinct regional influences. The estimation procedure is described in Johnston (1972) and Rausser and Johnson (1971).

Results

Initially, separate least squares regressions were estimated for each of the four regions under consideration. Results of these regressions are presented in table 15 where A and B are identified as the change in manufacturing income and the change in share of Island income, respectively. Estimated regression coefficients are taken to indicate the relative response of the growth measures to the explanatory variables within each region. This perhaps explains why neither the wage rate at the beginning of the period (x_5) nor the location variable (x_6) were tested to be statistically significant. Because wage rates are relatively homogenous within a given region and the subregional units are geographically contiguous, wage rates and location variables could prove more important when compared across regions. Perhaps the principal comment that may be made with respect to table 15 is that it suggests that there are probably substantial variations in the influence of water-resources investments among regions.

Coefficient estimates of restricted regressions are presented in

Table 15.--Separate regional regression coefficients

Dependent variable		Intercept (A = thousands of dollars) (B = percent)	Water investment X_2 (thousands of dollars)	Electrical supply investment X_3 (thousands of dollars)	Non-public works investment X_4 (thousands of dollars)	Wage rate X_5 (dollars per week)	Distance to urban center X_6 (miles)	R^2 **
Change in manufacturing income	1A	-1157.137	1.483*	-1.4210*	0.55711*	143.512	- 83.06	0.9412
			(.7023)	(.3981)	(.18585)	(186.849)	(133.439)	
	2A	- 38.199	-4.4268*	.20108*	1.31189*	35.9414	- 25.4766	.9347
			(1.2907)	(.07410)	(.19656)	(66.7440)	(83.0370)	
	3A	657.369	.76735	- .02789	.33610*	19.7949	- 29.3621	.6126
			(.78031)	(.13519)	(.09087)	(39.6439)	(30.2034)	
	4A	90.911	3.16703*	4.6682*	- .73128*	51.886	- 61.6933	.6422
			(1.29770)	(1.1966)	(.31054)	(65.242)	(41.0324)	
Change in municipio share of island income (percent)	1B	113.799	- .00512	.00193	.00266	3.0335	- .44021	.8702
			(.01006)	(.00570)	(.00266)	(2.6768)	(1.91167)	
	2B	- 7.925	.00289	.00134*	- .00628*	- .6589	.56604	.7992
			(.00879)	(.00050)	(.00134)	(.4546)	(.56561)	
	3B	- 18.401	- .02756*	.00074	- .00060	- .08242	- .04511	.4733
			(.00894)	(.00123)	(.00083)	(.36163)	(.27551)	
	4B	2.642	- .01688	.02794	.00001	- .2568	- .12313	.1824
			(.01326)	(.01222)	(.00317)	(.6665)	(.41917)	

*Denote statistically significant variables at 95 percent level.
Numbers 1, 2, 3, and 4 refer to San Juan, the Southwest, Northwest and Southeast Regions, respectively.

**Coefficient corrected for degrees of freedom.
Numbers in parentheses are standard deviation of regression coefficient.

tables 16 and 17. In both tables, coefficients for Region 1 include the average and restricted coefficients while coefficients for Regions 2, 3 and 4 are left in their differential form in order to statistically test for structural shifts in coefficient values across the set of regional equations.[1] Using an F-test, all regressions were tested to be statistically significant at the 90 percent level. In table 16 the impact of the explanatory variables was investigated from an overall (Island-wide) perspective by restricting the entire set of parameters to be equal across regions but letting the intercept terms vary. For changes in manufacturing income water-resources investment was not associated with increases in manufacturing income. However, there appears to have been an overall redistribution of income in favor of municipios with such investments. Because rather substantial water resource investments were for irrigation projects, agricultural land values for selected areas would have been enhanced, thus tending to discourage growth in manufacturing and perhaps biasing the overall results of the equation for manufacturing in table 16.

In table 17, the relative impact of public investment within regional units was investigated by restricting subsets of coefficients across regions and permitting unrestricted coefficients on public investment to reflect systematic variations resulting from characteristics of the region and the nature of investments. It should be observed immediately that the explanatory power (as measured by R^2) of both sets of regressions in table 17 was increased substantially over table 16. Examining the regressions relating to manufacturing activity, the overall coefficients (Region 1) imply a positive relationship between the change in manufacturing income and water resource investment as expected. Additions of the respective incremental coefficients to those of Region 1 indicate that only in the Southeast Region (Region 4) was water investment associated with a statistically significant negative regression coefficient. For several municipios of the southeast the Toa Vaca irrigation accounted for large water resources expenditures which apparently had no effect on the area's industrial development, which would explain the result of the statistically negative coefficient associated with water supply investment. Industrial development, primarily from location of the petrochemical factories would probably not have taken place without the anticipated development of reliable electrical energy which resulted from the electrical supply investments. This would explain the large positive coefficient associated with electrical power investment. Results relating to changes in regional income distribution indicate substantial differences in the responsiveness to public investment within regions.

---

1 Full regional coefficients are found by adding the average to regional incremental coefficients and variances are the sum of the respective variances minus twice the covariances. Covariance terms were quite small; not significantly changing the variance of the full coefficients.

Table 16.--Restricted regression coefficients with regional variations in intercepts

Dependent variable	Region	Intercept	Water investment X_2 (thousands of dollars)	Electrical supply investment X_3 (thousands of dollars)	Non-public works investment X_4 (thousands of dollars)	Wage rate X_5 (dollars per week)	Distance to urban center X_6 (miles)	R^{2}**
Change in manufacturing income	1	26.416* (.158)	-0.00811* (.00325)	0.000732 (.000455)	0.00497* (.00061)	0.9434* (.2847)	-0.6509* (.2058)	0.8110
	2	-23.195* (2.243)						
	3	-21.929 (12.193)						
	4	-15.12 (12.44)						
Change in municipio's share of Island income (percent)	1	36.50* (16.95)	.00863* (.00334)	.01184* (.00467)	.003704* (.000619)	- .5225 (.2756)	.0920 (.1966)	.2749
	2	-49.07* (14.19)						
	3	-39.76* (13.86)						
	4	36.788* (13.983)						

* Denotes statistically significant variables at 95 percent level. Numbers 2, 3, and 4 refer to Southwest, Northwest and Southeast Regions, respectively and coefficients are in differential form. Number 1 refers to the average of the regional coefficients and includes the San Juan Region. Although another could have been chosen, it was felt investigation of the regional responses of less developed areas would be more revealing.

** In the seemingly unrelated regression context, R^2 would not have the same interpretation as in a single equation estimation method. However, it is employed here as a measure of goodness of fit and calculated in the usual manner, that is, by treating the set of regressions as a single regression,

$$R^2 = 1 - \frac{\sum e_i^2}{\sum y_i^2} = 1 - \frac{\text{sum of squared residuals,}}{\text{total sum of squares}}$$

NOTE: Numbers in parentheses represent standard error of coefficient estimate.

Table 17. Estimated coefficients with subsets of parameters restricted

Dependent variable	Region	Intercept	Water investment X_2 (thousands of dollars)	Electrical supply investment X_3 (thousands of dollars)	non-public works investment X_4 (thousands of dollars)	Wage rate X_5 (dollars per week)	Distance to urban center X_6 (miles)	R^{2**}
Change in manufacturing income	1	18.977	1.7975*	-1.3444*	0.46202*	-50.276	-57.271*	0.9308
			(.5803)	(.3014)	(.07191)	(29.460)	(22.640)	
	2	-26.453	- .6702	1.5272*				
			(.7179)	(.3069)				
	3	-10.588	-1.7100	1.2845*				
			(.9676)	(.3340)				
	4	-28.299	-2.7127*	7.4425*				
			(.8943)	(1.1671)				
Change in municipio's share of Island income (percent)	1	- 4.039	.00067	- .00728	.02167	- .2627	.4254*	.7362
			(.00810)	(.00415)	(.0202)	(.2859)	(.2157)	
	2	- 9.0523	.04885*	.00803*	-.0139*			
			(.01565)	(.00420)	(.0027)			
	3	.25722	- .032287*	.00658	-.00320			
			(.012329)	(.00469)	(.00220)			
	4	3.9986	- .01757	.02408*	-.003142			
			(.01377)	(.01111)	(.00338)			

* Denote statistically significant variables at 95 percent level. Numbers 2, 3, and 4 refer to Southwest, Northwest, and Southeast Regions, respectively and coefficients are in differential form. Number 1 refers to the average of the regional coefficients and includes the San Juan Region. Although another could have been chosen, it was felt investigation of the regional responses of less developed areas would be more revealing.

** In the seemingly unrelated regression context, R^2 would not have the same interpretation as in a single equation estimation method. However, it is employed here as a measure of goodness of fit and calculated in the usual manner, that is, by treating the set of regressions as a single regression,

$$R^2 = 1 - \frac{\Sigma e_i^2}{\Sigma y_i^2} = 1 - \frac{\text{sum of squared residuals}}{\text{total sum of squares}}$$

NOTE: Numbers in parentheses represent standard error of coefficient estimate.

For the rapidly developing Southwest Region (Region 2) income was redistributed in favor of areas having water and electrical supply investments. The results for the Northwest Region (Region 3), one of the two most undeveloped regions, imply that public investments did not induce a redistribution of Island income within the region for the duration of the period considered.

CONCLUSIONS

Results presented here have several general implications. First, there appears to be substantial variations in regional responses to water resource investment, when responses are measured in terms of industrialization and changes in regional income distribution. Second, the results suggested that the nature of the water resource investment determines whether to expect increases in regional industrialization.

This latter point may be significant when examining longer time periods. Because such regions receiving heavy irrigation investments are unlikely to attract manufacturing activities, except perhaps in food processing industries, there will be no industrial base to sustain the income growth and to sustain the trends in the redistribution of income resulting from initial investment. Finally, the effectiveness of achieving income redistribution by such investments appears to depend crucially on the level of development within the region, as the substantially less developed areas indicated the smallest responses to such investments.

SUMMARY AND FUTURE DATA NEEDS

The purpose of this study was to develop operational techniques for water demand analysis and making forecasts. In the report, determinants of industrial, commercial and residential water demand for the San Juan metropolitan test area were examined. The discussion included both an analysis of the nature of water demand for Puerto Rico and also the presentation of procedures for making demand forecasts. Furthermore, the data base that was developed, particularly of municipio water use, could provide a means of developing demand models for other areas of the Island and for updating the models presented here. There are several steps which would aid in the eventual implementation and improvements of the models, the most important of which is keeping all billing records in computer-readable form.

The present data base (1960-70) for areas outside the test region will require additional refining and general clean-up. Industrial and commercial customers could be classified according to four digit SIC code and identified by SIC code on the billing records of Puerto Rico Aqueduct and Sewer Authority. A major reason for the poor results of the commercial demand estimates resulted from the aggregation of the diverse water. Present Government records relating to the amount of water self-supplied could also be improved. Information relating to the water system leakage could be developed for various geographical areas. In order to make optimal use of the procedures and new sample information, individual demand models for different areas and classes of users should be re-estimated perhaps annually and new projections could be developed which include the most recent sample information. Finally, the water resources planning agencies could establish a data base for plant sewer and effluent charges similar to the water use data base. Such a data base would enable the planner to estimate the responsiveness of firms to changes in effluent charges.

The second part of this study examined the impact of water resources investment on Puerto Rico from 1960 to 1968. Results of this analysis were presented because future water demand is dependent on economic growth and it was of interest to examine possible feedbacks of increased water demands induced by increased supplies (from investment and increased water availability). It was concluded that the impact of such investments were highly dependent on the nature of the investment and level of development of the area.

An evaluation of the degree of success and applicability of the models is called for. After a considerable amount of experimentation with static demand models, the dynamic models were chosen because they best represented the changes in water demand which took place during this period

Of rapid economic development. As discussed previously, dynamic models of these types have a forecast range (six periods) whereby the forecasts' variances become quite large. Therefore, for these models a considerable degree of uncertainty should be attached to projections which are made, for example, for 10 years after the last year actual data were available. Because of a comparative degree of disaggregation and reliability of water use and economic data the residential models seemed to perform best followed by those related to industrial water and finally commercial water use.[1] The models, methods, programs and data generated by this project represent a beginning for the Commonwealth in terms of the development of accurate water demand forecasts.

No attempt was made to provide a cookbook or manual of procedures for use by untrained personnel because the present (1975) state of economic demand forecasting particularly for water is not well developed, and because improvements in data along with further disaggregation will result in substantial improvements in the models. Application of these techniques, which we believe are operational, will be an improvement over present projection methods for the Commonwealth but need to be done by trained professionals.

[1] The development of the present data base represented an investment of at least 3-man years.

REFERENCES

Armstrong, J.W., and Graham, M.C., 1972, A comparative study of methods for long-range market forecasting: Management Sci., v. 19, no. 2, p. 211-221.

Bain, J., Caves, R., and Margolis, S., 1966, Northern California's water industry: Baltimore, Johns Hopkins Univ. Press, p. 766.

Balestra, P., and Nerlove, M., 1966, Pooling cross section and time series data in the estimation of a dynamic model-- the demand for natural gas: Econometrica, v. 34, no. 2, p. 585-612.

Bower, B.T., Löf, G.O., and Hearon, W.M., 1971, Residuals generation in the pulp and paper industry: Natural Resources Jour., v. 11, no. 4, p.605-623.

Buck, Seifert, and Jost, 1971, Memorandum on improvements to Naguabo water supply: San Juan, Puerto Rico, consultants' rept.

Conely, B.C., 1967, Price elasticity of the demand for water in southern California: Annals of Regional Sci., Dec., p. 180-189.

Cooley, T.F., and Prescott, E.C., 1973a, Tests of an adaptive regression model: Rev. Economics and Statistics, v. 55, May, p. 248-256.

_______, 1973b, The adaptive regression model: Internat. Economic Rev., v. 14, June, p. 364-371.

_______, 1973c, Varying parameter regression--a theory and some applications: Annals Economic and Social Measurement, v. 2, Oct.

Cootner, P.H., and Löf, G.O., 1965, Water demand for steam electric generation-- an economic projection model: Baltimore, Johns Hopkins Univ. Press, 144 p.

Cordero, I.A., 1969, The management and control of water in Puerto Rico: Univ. Puerto Rico, Water Resources Research Inst., Tech. Completion Rept. A-010-PR, 223 p., (PB 189-157).

De Rooy, J., 1970, The industrial demand for water resources --an econometric analysis: Ph.D. dissertation, Rutgers Univ., 292 p.

Dhrymes, P., 1964, On the theory of the monopolistic multiproduct firm under uncertainty: Internat. Econ. Rev., v. 5, no. 3, p. 239-257.

_______, 1970, Econometrics: New York, Harper and Row, 592 p.

Duncan, D.B., and Horn, S.D., 1972, Linear dynamic recursive estimation from the viewpoint of regression analysis: Statistical Assoc. Jour., v. 67, Dec., p. 815-821.

Fisher, F.M., and Kaysen, C., 1962, A study in econometrics-- the demand for electricity in the United States: Amsterdam, North Holland, 190 p.

Frieden, A., A program for the estimation of dynamic economic relations from a time series of cross-sections: Annals of Econ. and Social Measurement, v. 2, no. 3, p. 347-350.

Fuchs, V.R., 1962, The determinants of the redistribution of manufacturing in the United States since 1929: Rev. Econ. and Statistics, v. 44, no. 2, p. 1967-1977.

Garrison, C.G., and Paulson, A.S., 1971, Effect of water availability on manufacturing employment in the Tennessee Valley region: Water Resources Research, v. 44, p. 301-316.

Goldberger, A.S., 1964, Econometric theory: New York, John Wiley and Sons, 398 p.

Hanke, S., 1970, Demand for water under dynamic conditions: Water Resources Research, v. 6, no. 5, p. 1253-1261.

Headly, J.C., 1963, The relation of family income and use of water for residential and commercial purposes in the San Francisco-Oakland metropolitan area: Land Econ., v. 39, no. 6, p. 441-449.

Houthakker, H., and Taylor, L.D., 1971, Consumer demand in the United States: 2nd ed., Cambridge, Mass., Harvard Univ. Press, 321 p.

Houthakker, H., Verleger, P., and Sheehan, P., 1973, Dynamic demand analysis for gasoline and residential electricity: paper presented at Am. Econ. Assoc. Meeting, Dec.

Howe, C.W., 1967, Water resources and regional economic growth in the United States, 1950-1960: Southern Econ. Jour., v. 34, p. 477-489.

Howe, C.W., and Linaweaver, F., 1967, The impact of price on residential water demand and its relation to system design and price structure: Water Resources Research, v. 3, no. 1, p. 13-37.

Howe, C.W., Russell, and Young, R.A., 1971, Future water demands, the impacts of technological change, public policies and changing market conditions on water use patterns in selected sectors in the United States economy, 1970-1990: Rep. EES 71,001 Natl. Water Comm., 102 p.

Johnston, J., 1972, Econometric methods: 2nd ed., New York, McGraw-Hill Book Co., 300 p.

Kaufman, A., and Nadler, M., 1966, Water use in the mineral industry: Inf. Circ. 8285, Washington, D.C., U.S. Bur. of Mines, 46 p.

Kmenta, J., 1971, Elements of econometrics: New York, Macmillan Co., 651 p.

Lewis, W.C., 1973, Public investment impacts and regional growth: Water Resources Research, v. 9, no. 4, p. 851-860.

Löf, G.O., and Kneese, A., 1968, The economics of water utilization in the beet sugar industry: Baltimore, Johns Hopkins Univ. Press, 125 p.

Lofting, E.M., and Davis, H.C., 1968, The inter-industry water content matrix--applications on a multiregional basis: Water Resources Research, v. 4, no. 4, p. 689-695.

Moody, D.W., Attanasi, E.D., Close, E.R., López, M. A., and Maddock, T., III, 1973, Puerto Rico water resources planning model study: U.S. Geol. Survey open-file rept., 113 p.

Nerlove, M., 1971a, A note on error component models: Econometrica, v. 39, no. 2, p. 383-396.

_______, 1971b, Further evidence on the estimation of dynamic economic relations from a time series of cross sections: Econometrica, v. 39, no. 2, p. 357-382.

Puerto Rico Planning Board, 1970, Regional growth model for Puerto Rico: San Juan, Puerto Rico, 115 p.

Puerto Rico Aqueduct and Sewer Authority, 1969, Regional domestic water demand estimates, 1970-1990: Commonwealth of Puerto Rico Aqueduct and Sewer Authority, 36 p.

Rausser, G.C., and Johnson, S.R., Structural shifts and equality constraints across equations: Metroeconomica (in press).

Russell, C.S., 1973, Residuals management in industry--a case study of petroleum refining: Baltimore, Johns Hopkins Univ. Press, 193 p.

Sewell, R.D., and Bower, B.T., eds., 1968, Forecasting the demand for water: Dept. of Energy, Mines and Resources, Ottawa, Canada, 261 p.

Tate, D.M., and Robichaud, R., 1973, Industrial water demand forecasting: Water Planning and Management Branch, Environment Canada, Ottawa, Canada, 18 p.

Thiel, H., 1961, Economic forecasts and policy: Amsterdam, North-Holland Publishing Co., 567 p.

_______, 1971, Principles of econometrics: New York, John Wiley and Sons, 736 p.

Turnovsky, S.J., 1969, The demand for water--some empirical evidence on consumers; response to a commodity uncertain in supply: Water Resources Research, v. 5, no. 2, p. 350-361.

_______, 1973, Optimum government investment in systems yielding uncertain outputs: in Models for Managing Regional Water Quality, ed. by Robert Dorfman, Henry Jacoby, and Harold Thomas, Jr., Harvard Univ. Press, Cambridge, p. 233-264.

U.S. Bureau of Census, 1966, Water use in manufacturing: Bull. MC 63 (1)-10, Washington, D.C., U.S. Govt. Printing Office, 174 p.

U.S. Corps of Engineers, 1972, North Atlantic regional water resources study: App. T, May 1970, North Atlantic Regional Water Resources Study Coordinating Comm., p. 335.

Verleger, P.K., Jr., 1972, Models of the demand for air transportation: Bell Jour. of Econ. and Management Sci., v. 3, no. 2, p. 437-457.

Wong, S.T., 1972, A model on municipal water demand--a case study of northeastern Illinois: Land Economics, v. 48, no. 1, p. 34-44.

Yan, C., 1969, Introduction to input-output economics: New York, Holt, Rinehart and Winston, 134 p.

Young, R.A., 1973, Price elasticity of demand for municipal water--a case study of Tucson, Arizona: Water Resources Research, v. 8, no. 4, p. 1068-1072.

APPENDIX A

Estimation From A Time Series of Cross-Sections By Error Components Models

Perhaps the most widely applied method of efficiently combining time series and cross-sectional information is by application of error components estimation techniques. Further, the particular technique used here, originally developed by Nerlove and Balestra (1966) also considers statistical problems which result when the set of explanatory variables includes a lagged endogenous variable. The estimation procedure described here is described in Nerlove (1971). The specific form of the model is set up first according to individual cross-sectional units then according to time periods

$$y_{it} = \alpha y_{it-1} + \beta x_{it} + u_{it} \qquad \begin{matrix} i = 1,\dots,N \\ t = 1,\dots,T \end{matrix}$$

where

$$y = (y_{11},\dots,y_{1T},\dots,y_{N1},\dots,y_{NT})'$$

$$y_{t-1} = (y_{10},\dots,y_{1T-1},\dots,y_{N0},\dots,y_{NT-1})'$$

$$x = (x_{11},\dots,x_{1T},\dots,x_{N1},\dots,x_{NT})'$$

$$u = (u_{11},\dots,u_{1T},\dots,u_{N1},\dots,u_{NT})'$$

The components of the error term have the following specification

$$u_{it} = \mu_i + \nu_{it}$$

$$E\mu_i = E\nu_{it} = 0 \qquad \text{all } i \text{ and } t$$

$$E\mu_i\mu_{i'} = \begin{matrix}\sigma^2_\mu \\ 0\end{matrix} \qquad \begin{matrix} i \neq i' \\ \end{matrix}$$

$$E\nu_{it}\nu_{i't'} = \begin{matrix}\sigma^2_\nu \\ 0\end{matrix} \qquad \begin{matrix} i = i', \ t = t' \\ \text{otherwise.} \end{matrix}$$

The error term is composed of a component associated with variation over individuals and time. The variance-covariance matrix of the error terms is

$$E_{uu'} = \sigma^2 \begin{bmatrix} A & 0 & \dots & 0 \\ 0 & A & \dots & 0 \\ \dots & & & \\ \dots & & & \\ \dots & & & \\ 0 & 0 & \dots & A \end{bmatrix}, \qquad A = \begin{bmatrix} 1 & p & \dots & p \\ p & 1 & \dots & p \\ \dots & & & \\ \dots & & & \\ \dots & & & \\ p & p & \dots & 1 \end{bmatrix},$$

and $p = \sigma^2_\mu / \sigma^2$.

This particular procedure involved iteratively estimating parameters of equation 1A in order to eliminate the presumed autocorrelation in the error term. That is, first slopes are estimated by a least squares regression of deviations of the dependent variable from individual means, that is $(y_{it} - y_i)$ on deviation of the independent variables. With β_k^* the estimates of these slopes and e^2 the sum of the squared residuals. The estimate of ρ is derived from the following procedure:

$$\hat{\rho} = \frac{\sigma_\mu^2}{\sigma_\mu^2 + \frac{e^2}{NT}}$$

where

$$\hat{\sigma}_\mu^2 = \frac{1}{N} \sum_{i=1}^{N} \{(\bar{y}_i - \bar{\bar{y}}) - \sum_k \beta_k^* [\bar{x}_i(k) - \bar{\bar{x}}(k)]\}$$

The estimate of ρ is then used to compute weights

$$\theta = (1-\hat{\rho}) + T\hat{\rho}$$

$$\eta = (1-\hat{\rho})$$

which are used to transform the variables

$$y_{it}^* = \frac{y_{it} - \bar{y}_i}{\sqrt{\eta}} + \frac{\bar{y}_i}{\sqrt{\xi}}$$

upon which the second round of least squares regression parameters are estimated for the equation:

$$y_{it}^* = \hat{\alpha}\, y_{it-1} + \hat{\beta} x_{it} + u_{it} .$$

Nerlove (1971) has shown that the least squares estimates from the final round of regressions on the transformed variables is equivalent to the application of generalized least squares, that is, Aitkens estimators, for the original model specification. Monte Carlo experiments of alternative estimation procedures suggest that the proposed procedure compares favorably with nonlinear maximum likelihood methods of estimation.

APPENDIX B

Illustrative Water Demand Projections for Residential, Commercial and Industrial Water Use

Table B-1.--Illustrative projection of total residential water demand

Municipio	1975			1980			1985			1990		
	1	2	3	1	2	3	1	2	3	1	2	3
Bayamón	287	291	297	348	359	380	419	444	495	504	553	665
Caguas	132	134	138	162	167	181	197	212	242	241	270	335
Carolina	135	138	142	176	185	200	224	244	284	284	323	413
Cataño	39.3	40.0	40.8	49.3	51.1	54.4	60.9	65.0	73.3	75.0	82.9	101
Ceiba	10.7	10.8	11.1	14.0	14.6	15.5	18.0	19.1	21.6	22.7	25.1	30.7
Fajardo	35.2	35.8	36.8	43.2	45.0	48.5	52.7	56.7	65.1	64.3	72.0	90.4
Guaynabo	147	149	154	180	188	205	220	239	278	269	305	391
Loíza	46.8	47.1	47.7	58.9	60.0	61.8	72.0	74.4	79.1	87.6	92.0	103
Luquillo	12.6	12.8	13.1	15.2	15.7	16.6	18.3	19.4	21.6	22.0	24.0	28.9
Río Grande	18.7	19.1	19.7	23.6	24.7	26.8	29.6	32.2	37.5	37.0	42.2	54.0
San Juan	990	1022	1076	1241	1340	1522	1553	1770	2220	1948	2365	3344
Toa Baja	79.9	80.4	81.1	96.7	98.1	101	115	118	125	137	143	157
Trujillo Alto	34.9	35.1	35.5	44.2	44.9	46.4	54.9	56.6	60.4	67.6	71.2	79.5

Projections are in million gallons per month.

Table B-1.--Illustrative projection of total residential water demand--Continued

Municipio	1975			1980			1985			1990		
	4	5	6	4	5	6	4	5	6	4	5	6
Bayamón	336	288	280	479	354	340	683	435	412	981	540	503
Caguas	155	132	128	225	165	157	327	207	193	479	262	239
Carolina	159	136	133	246	182	175	375	239	227	574	316	295
Cataño	46.1	39.4	38.5	68.1	50.3	48.3	99.9	63.6	60.1	147	80.9	75.1
Ceiba	12.5	10.7	10.4	19.4	14.3	13.7	29.5	18.7	17.5	44.6	24.4	22.5
Fajardo	41.3	35.3	34.2	60.1	44.2	42.0	87.3	55.3	51.7	128	70.0	64.1
Guaynabo	172	148	145	251	186	180	367	235	226	542	300	285
Loíza	54.4	46.4	44.9	79.9	58.6	55.5	114	72.3	66.8	163	88.9	79.9
Luquillo	14.8	12.6	12.1	20.9	15.3	14.3	29.8	18.7	17.0	42.7	23.0	20.3
Río Grande	22.1	18.7	17.9	32.9	24.0	22.3	49.5	31.0	28.0	74.8	40.4	35.4
San Juan	1181	1013	993	1786	1323	1282	2722	1743	1673	4201	2326	2214
Toa Baja	92.8	79.5	77.5	131	96.5	92.6	182	116	109	255	140	129
Trujillo Alto	40.5	34.7	33.6	59.9	44.0	42.0	87.2	55.2	51.5	126	69.0	62.7

Projections are in million gallons per month.

Table B-1.--Illustrative projections of total residential water demand--Continued

Municipio	1975		1980		1985		1990	
	7	8	7	8	7	8	7	8
Bayamón	358	292	440	361	541	446	671	554
Caguas	166	136	207	172	259	216	327	274
Carolina	170	147	227	202	299	270	395	357
Cataño	46.1	41.1	58.9	53.6	74.7	67.3	95.2	83.3
Ceiba	13.3	11.7	17.7	16.2	23.2	21.4	30.3	28.1
Fajardo	43.5	35.9	54.5	45.5	68.3	57.3	86.4	72.6
Guaynabo	184	151	232	191	292	242	373	310
Loíza	57.9	48.6	73.1	62.9	90.1	78.3	111	96.6
Luquillo	15.4	12.7	18.6	15.5	22.8	18.8	28.1	23.0
Río Grande	23.0	19.8	29.4	26.0	38.1	34.0	49.6	44.3
San Juan	1239	1030	1619	1358	2133	1791	2848	2363
Toa Baja	96.0	81.0	117	99.5	140	118	169	140
Trujillo Alto	43.1	37.2	54.8	48.8	68.8	62.0	85.9	77.6

Projections are in million gallons per month.

Table B-2.--Assumptions to table B-1 illustrative residential projections

Assumptions for forecast (number)	Income growth (percent)	Population growth (percent)	Growth in prices (percent)	Growth in metering (percent)	Growth in percentage serviced
1	3	3	0	0	3
2	5	3	0	0	3
3	8	3	0	0	3
4	5	6	0	0	3
5	5	3	1	0	3
6	5	3	3	0	3
7	5	3	1	25	3
8	5	3	1	0	6

Note: Growth rates are presented on an annual basis.

Table B-3.--Comparison for San Juan metropolitan area residential water use

Projections *	1980	1985	1990
No. 5-table B-2	1650.4	1852.2	2079.5
No. 8-table B-2	2139.3	2770.8	3631.9
PRASA (1969) projection	2202.9	2935.7	3803.77

* In millions of gallons per month.

PRASA = Puerto Rico Aqueduct and Sewer Authority

Table B-4.--Illustrative commercial water use projections

Municipio	1975			1980			1985			1990		
	1	2	3	1	2	3	1	2	3	1	2	3
Bayamón	37.5	35.875	41.18	38.3	34.635	46.111	38.7	32.94	51.17	39.2	31.270	56.91
Caguas	13.2	11.680	14.52	13.3	10.082	16.210	13.4	8.399	18.07	13.6	6.677	20.22
Carolina	9.5	9.161	10.528	9.7	8.778	11.729	9.9	8.349	13.041	10.0	7.939	14.55
Cataño	2.68	2.266	2.788	2.58	1.955	3.071	2.59	1.238	3.371	2.59	1.238	3.682
Ceiba	15.26	14.845	16.782	15.46	14.573	18.713	15.49	14.104	20.640	15.52	13.610	22,76
Fajardo	28.80	29.935	31.764	29.12	27.198	35.424	29.34	26.317	39.372	29.61	25.448	43.832
Guaynabo	6.82	6.481	7.521	6.89	6.181	8.402	6.2	5.786	9.287	6.90	5.367	10.260
Luquillo	1.505	1.299	1.658	1.501	1.057	1.823	1.502	.806	2.009	1.503	.543	2.215
Río Grande	1.785	1.414	1.963	1.780	.981	2.154	1.782	.532	2.373	1.786	.061	2.628
San Juan	313.87	313.74	345.66	330.71	327.46	401.191	351.784	346.41	470.255	378.470	374.69	557.767
Toa Baja	3.853	3.567	4.152	3.761	3.146	4.364	3.761	2.797	4.713	3.768	2.439	5.109
Trujillo Alto	3.878	3.589	4.282	3.914	3.291	4.771	3.920	2.943	5.275	3.924	2.578	5.831

Projections are in million gallons per month.

Table B-5.--Assumptions for commercial water use projections.

	Growth in income (percent)	Growth in prices (percent)	Growth in number of customers (percent)
1	3	0	0
2	7	1	0
3	3	0	1

Growth rates are on an annual basis. Growth in the number on new customers refers to metered customers only.

Table B-6.--Illustration of projections for water intensive regions

Municipio	1975			1980			1985		
	1	2	3	1	2	3	1	2	3
Bayamón	47.4	43.19	38.87	54.84	45.24	35.12	63.15	47.92	31.30
Carolina	34.00	32.53	31.04	36.74	33.49	29.93	39.48	34.22	28.46
Cataño	11.85	9.71	7.54	14.6	9.83	4.73	17.92	10.28	19.24
Guaynabo	28.37	26.29	24.17	32.18	27.47	22.51	36.45	28.89	20.74
Toa Baja	12.80	10.71	8.57	15.24	10.53	5.53	17.87	10.39	2.18

Projections are in million gallons per month.

Table B-7.--Assumptions for projections in water intensive regions

	1	2	3
Rate of plant output growth. in percent	5	7	5
Rate of growth of customers, in percent	0	0	0
Rate of growth of prices, in percent	0	.75	1.5
Growth in local wage rate, in percent	2	2	2

Table B-8.--Illustration of predictions for industrial water use for low water intensive regions

Municipio	1975				1980				1985			
	1	2	3	4	1	2	3	4	1	2	3	4
Caguas	12.4	11.7	11.4	14.6	14.1	12.34	11.81	19.23	15.85	13.23	12.30	25.57
Fajardo	1.15	1.011	.965	1.231	1.36	1.06	.954	1.55	1.60	1.12	.954	1.984
Loíza	1.41	1.36	1.35	1.718	1.50	1.40	1.36	2.22	1.60	1.44	1.38	2.87
Luquillo	.619	.551	.528	.674	.750	.600	.55	.894	.895	.661	.577	1.200
Río Grande	1.55	1.462	1.431	1.826	1.774	1.571	1.501	2.445	2.017	1.70	1.586	3.30
San Juan	69.65	67.42	66.67	85.09	76.29	71.38	69.68	113.50	83.66	75.97	73.23	152.25
Trujillo Alto	1.08	1.032	1.015	1.295	1.185	1.076	1.038	1.691	1.301	1.130	1.069	2.222

Projections are in million gallons per month.

Table B-9.--Assumptions for alternative projections in low water intensive regions

	1	2	3	4
Rate of plant output growth, in percent	5	5	5	5
Rate of growth of number of customers in percent	0	0	0	0
Rate of growth of prices, in percent	0	.75	1.5	1
Rate of local wage increase, in percent	2	2	2	2

Appendix C

Description of Principal Components Analysis

Although the principal components technique was not used here for projecting water demand for the test area, the technique may prove valuable when data from other regions of the island are employed. The purpose of the principal components analysis is to define a number of mutually uncorrelated independent variables exhibiting in some sense the maximal variance. The approach is particularly appropriate in situations when the number of explanatory variables is large relative to the number of observations; thus, limiting the available degrees of freedom. The technique is frequently applied to reduce the dimensionality of the set of explanatory variables, particularly when the original variables are highly collinear. Because of this reduction, the technique has been used for index construction. In the following discussion, the essential points of the technique are presented and additional references are provided.

Initially, its description is carried out with reference to the collinearity problem, whereby a set of variates X behave nearly proportionally. That is, if the variables are proportional, their behaviour may be described by

$$X = a Z \qquad \text{(C 1)}$$

where a is row vector and X is an m by n matrix B1. Suppose that a'a = 1. The first principal component is defined to reflect the maximum variation of X, suggesting the sum of the squared discrepencies are minimized

$$\text{tr } (X-Za)\ (X-Za)^{1} = \text{tr } (X'X) - 2Z'\,Xa + a'a \qquad \text{(C 2)}$$

where

$$\text{tr } aZ'Xa = Z'Xa \text{ and tr } Z'Z\ a'a = a'a$$

Differentiating equation (C 2) with respect to a for a given Z and setting it equal to zero provides

$$a = X'Z \qquad \text{(C 3)}$$

which gives the coefficient vector a in terms of Z. In order to interpret these relationships observe that by substituting (C 3) into (C 2) then tr X'X - Z'X'XZ indicates that the problem is equivalent to maximizing Z'X'XZ for variations in Z subject to Z'Z = 1. If the Lagrangian expression is formed and differentiated the condition for solution of Z is

$$(XX' - \lambda I)\,Z = 0 \qquad \text{(C 4)}$$

---

[1] The trace of A'A is (where A is a m x n matrix) the sum of the squared residuals.

where λ is the Lagrangian multiplier. Therefore, Z is a characteristic vector of the n x n positive semidefinite matrix X'X corresponding to the largest root, although it is not normalized to have unit length. If $\lambda Z = X(X'Z) = Xa$ then $Z = \frac{1}{\lambda} Xa$ where the Z provides the best linear description of the X columns, that is, exhibiting maximum variations and is identified as the first principal component. Second, third and higher order components may be derived in a similar fashion, when the following expression is minimized

$$tr\,(X - Z_1 a_1 - \ldots Z_k a_k')' \; (X - Z_1 a_1 \ldots Z_k a_k)$$

where

$$Z_i Z_k = 1 \qquad i = k$$

$$Z_i Z_k = 0 \qquad i \neq k$$

Additional references relating to physical component analysis includes Thiel (1971), and Dhrymes (1970).

Appendix D

List of Data Developed or Used in Study

Data Covers 76 Municipios

1. Puerto Rico Planning Board data file on economic and social statistics (PLANØI).

2. Supplemental data on personal income and wages (Puerto Rico Planning Board).

3. Quantities supplied customers, and revenues generated from municipal water supplies (Puerto Rico Aqueduct and Sewer Authority).

4. Puerto Rico Census of Manufacturing (1965-71).