April 1982

A.E. Res. 82-25

SCALE ECONOMIES IN RURAL WATER SUPPLY SYSTEMS AND WATER QUALITY STANDARDS OF THE U.S. ENVIRONMENTAL PROTECTION AGENCY

Ramesh Vaidya

David J. Allee

Department of Agricultural Economics
Cornell University Agricultural Experiment Station
New York State College of Agriculture and Life Sciences
A Statutory College of the State University
Cornell University, Ithaca, New York, 14853

TABLE OF CONTENTS

LIST OF TABLES

SCALE ECONOMIES IN RURAL WATER SUPPLY SYSTEMS AND WATER QUALITY STANDARDS OF THE U.S. ENVIRONMENTAL PROTECTION AGENCY

BY

RAMESH VAIDYA AND DAVID J. ALLEE*

I. INTRODUCTION

Managers of water supply systems in rural communities are faced with meeting the cost of making major alterations to physical facilities for improving the quality of water they supply. In the United States, there are more than 34,000 rural community systems, each serving a population of at least 25 people. All are legally required to meet the quality standards set by the Environmental Protection Agency's interim primary drinking water regulations. Typically, however, these rural systems lack funds to invest in capital construction after operating and maintenance costs. As a result, they rely heavily on external sources of funds from various government agencies and commercial markets. Yet, because of the federal cutbacks in funding and the increasing costs of long-term financing by issuing bonds, these sources of funding may disappear or be severely limited.

* The authors are respectively Research Associate and Professor at the Department of Agricultural Economics, Cornell University, Ithaca, New York 14853. They are indebted to Joe D. Francis for valuable suggestions and for permission to use the rural water survey data from the National Statistical Assessment of Rural Water Conditions. They are thankful to Bruce L. Brower for his help in transforming water quality data from the household level to the water system level. They are also grateful to Anne E. Johnson for editorial help. Financial support from Cornell University's College of Agriculture and Life Sciences is gratefully acknowledged.

The general purpose of this study was to explore ways of improving the quality of rural water without adding to its cost. In specific the study examined a number of rural community systems in operation to discover (1) whether the average cost of producing water tended to be less for large systems, (2) whether such scale effects were large enough to mitigate the scale diseconomies likely to occure because of decreasing customer density as service areas expand, and (3) whether sufficient margin will remain for improving the quality of the water produced without escalating the cost. The scope of this study was limited to systems serving more than 25 people or 14 customers since these are the ones for which the EPA regulations are legally binding.

The hypothesis to be tested in this study is that, by maneuvering policy variables concerning the implementation characteristics of the water systems, it is possible to improve the quality of water produced by rural water systems, both existing and planned, without escalating costs. The policy variables postulated are (1) the size of a system, (2) the population density of the service area, and (3) the type of system ownership.

This study was built upon a national level data base developed in 1978 by an EPA sponsored research project at Cornell University entitled, The National Statistical Assessment of Rural Water Conditions (NSA). In that data base, data were available for 800 of an estimated 34,014 rural community water supply systems. Production costs were estimated by using both an accounting and an economic method. In the former method, the production cost was estimated as the sum of the debt retirement expenditures and the operating and maintenance costs incurred in 1978. In the latter, production cost was estimated as the sum of the resource input costs standardized to the 1978 prices. Two sets of average production costs were computed from each of these estimates; one uses the annual volume of water produced as a base, the other uses the number of customers served as a base.

Following these estimation multiple regression techniques were used to clarify the relationships between the average cost and the policy variables consisting of the size of the system, the population density of the service area and the type of system ownership. Water quality improvement was also introduced as an explanatory variable. In addition, three other explanatory variables which have limited maneuverability but do influence production costs were included: the region of a water system's location, the source of water and the system utilization rate. The results of these analyses were then used to discuss (1) the relationship between the average cost and the water quality improvement, and (2) the role of the three policy variables in improving water quality without escalating cost.

The presentation of this research is organized along the following lines. The main objective in Section II is to develop an analytic framework for finding out how the process of providing drinking water services is carried out by rural water systems. Section III describes how such a process is undertaken in the U.S. by using that analytic framework. Then, Section IV presents an analysis for investigating if there are better alternatives to what is being done and what they are, if any. Next, in Section V, discussions are conducted regarding why these better alternatives can or cannot be used. Finally, the conclusions and policy implications of this study are presented in Section VI.

II. A FRAMEWORK FOR ANALYSIS

The Conceptual Framework

The conceptual framework used for analysis is presented in Figure 1 and described in the following paragraphs.

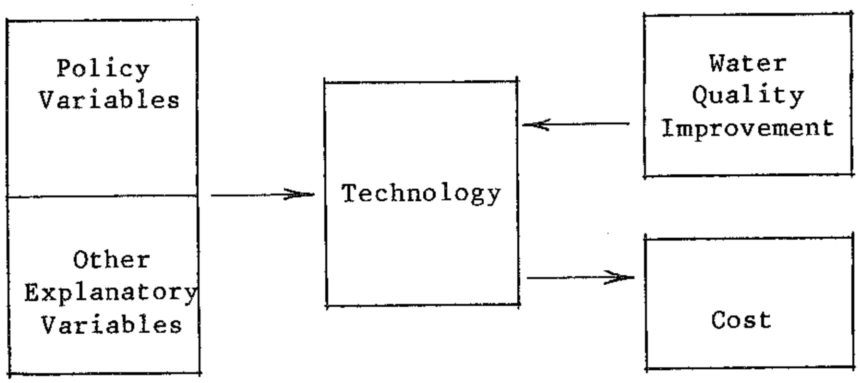

Figure 1. A Conceptual Framework for Cost-Quality Analysis

The primary goal of rural water systems is to provide drinking water services whose quality has been improved to the standards set by the EPA's interim primary drinking water regulations. The achievement of this goal is accomplished through various alternative modes or technologies of supplying drinking water which can be expressed in summary form as the mix of such resource inputs as capital investment requirements, labor, energy and chemicals. Because the choice of technology is sensitive to the characteristics of the system, e.g., the system size, there is room for discretion and for policy variables which can influence the mode or technology by which the system functions. In other words, these policy variables can influence the cost of production, because the cost is nothing but a sum of the resource inputs weighted by their prices. It may, therefore, be possible for systems to attain the level of water quality improvement required by the EPA without escalating the average cost of production if appropriate policy variables can be identified and maneuvered.

Three policy variables were identified for the analysis: (1) the size of water systems, (2) the population density of the service area, and (3) the type of system ownership. In addition, three other explanatory variables were identified which also influence the choice of technology and subsequently the cost of production, but maneuverability of these variables appeared to be marginal: (1) the region of a system's location, (2) the source of water (which explains differences in investment cost for storage facilities for different water sources such as ground or surface water), and (3) the utilization rate of the system's capacity measured as a ratio of the average daily production level to the design capacity (which explains short-run scale economies).

Measurement and Estimation of Production Cost

The crucial part of exploring interrelationships among the water quality improvement, the policy and other explanatory variables and the production cost lies in the measurement and estimation of the quality improvement and the cost.

Production cost has been estimated by two methods, the accounting and the economic.

In the accounting method, the production cost is measured by the annual expenditures of a system.[1] Such expenditures consist of operating, maintenance and debt retirement costs. In the National Statistical Assessment data base, data were available for estimating production cost by this method for 614 out of the 800 systems.

However, the production cost estimates prepared by this method had a number of limitations. First, the accounting method did not fully consider the replacement cost of the water system. The replacement cost is typically accounted for by including the depreciation cost in constant dollars in the estimates. It is expected that the accumulated depreciation reserve would equal the replacement cost of equipment and structures for the water system at the end of their productive life cycles. Second, the accounting method underestimated the capital cost for systems that funded their facilities partly by government grants or by the systems' internal reserve funds.

---

[1]For example, the accounting method in D. V. Bourcier and R. H. Forste, Economic Analysis of Public Water Supply in the Piscataqua River Watershed, (Durham: Water Resources Research Center, University of New Hampshire, March 1967).

Because of these limitations of the accounting method, in addition to this we have chosen to estimate cost by an alternative method--the economic method. In this method, the mix of the resource requirements for production were identified and production costs estimated by summing them and weighting them with their respective unit prices.[2] In this context, the resource inputs of concern are the (1) capital consisting of structures and equipment, (2) labor, (3) chemicals, (4) energy, and (5) administration and management.[3]

The capital input cost was calculated by multiplying together the construction cost of a system and the capital recovery factor, which is a function of the interest rate on capital loans and the depreciation period of structures and equipment. The capital cost thus included both the replacement and the interest cost components. The capital costs thus estimated were already standardized, because the construction costs used to estimate them were all standardized to 1978 by using price deflators for water supply facilities construction, and because the capital recovery factor using a constant depreciation period and interest rate was used for all systems.

---

[2]E.g., R. B. Johnson and W. P. Hobgood, Jr., A Cost Analysis of Rural Water Systems in Louisiana, DAE Research Report No. 483 (Baton Rouge, LA: Dept. of Agr. Econ. and Agribusiness, Louisiana State Univ. and Agr. and Mech. College, May 1975); and, US Environmental Protection Agency, Managing Small Water Systems: A Cost Study, 2 volumes (Cincinnati, Ohio: Municipal Environ. Res. Lab., September 1979.

[3]In US/EPA, op. cit., vol. 1, pp. 111-112, the cost proportions of these inputs have been estimated as the following for a typical water supply system with an average daily production of revenue producing water equal to 1.2 million gallons: (in percent) capital, 39, administrative and management cost, 25; labor, 22; energy, 11; and chemicals, 3.

This capital cost was added to the operating and maintenance cost in the NSA data base to obtain the production cost estimates by the economic method. The annual operating and maintenance (O&M) expenditure included the costs of labor, chemicals, energy and administrative inputs. The major limitation of the estimates developed by the economic method, however, was that this method required the construction cost data for systems which were available for only 99 out of the 800 systems. As a result, the number of observations was reduced from the 614, available for the accounting method estimates, to 99 for the economic analysis. Thus, separate analyses were conducted using cost estimates from both methods to disclose relative strengths and weaknesses of the results produced.

Measurement of Water Quality Improvement

Water quality improvement is commonly measured for each contaminant as the difference between the contamination level in the finished water and that of the untreated water.[4]

Unfortunately, the NSA base included data for only the finished water quality. Therefore for presenting the quality of raw water this study uses two surrogate variables: the source of raw water and the region of the system's location. As a result, we will be considering these two surrogate variables along with the contamination level of finished water for each applicable constituent.[5]

---

[4]EPA, op. cit., Vol. 1, pp. 142-147.

[5]For example, the contribution due to the water quality improvement in the total average production cost, c, will have three parts which can be demonstrated as follows:

In our data base, there were two alternative sets of data available for the bacteriological and chemical constituents of the finished water: (1) the data from the test performed on the water just before it is transmitted from the systems to the households, and (2) the data from the test of water samples collected from the households connected with these systems.

A choice of data set from these two alternative sets should consider the error that may be introduced when the water samples from different sources are tested in a large number of laboratories for the same constituent, e.g., lead. Recently, it was contended that the interlaboratory bias in measurement is large in the scientific testing of constituents

---

Let water quality improvement for the total coliform, q = the finished quality for that constituent, q_f - the raw quality, q_i.

Also, let q_i = f (source of water(s), region of location (r)).

Then, $\frac{\partial c}{\partial q} = \frac{\partial c}{\partial (q_f - q_i)}$.

Therefore, $$\frac{\partial c}{\partial q} \delta q = \frac{\partial c}{\partial (q_f - q_i)} \delta (q_f - f(s,r))$$

$$= \frac{\partial c}{\partial (q_f - q_i)} \delta q_f - \frac{\partial c}{\partial (q_f - q_i)} \delta f(s,r).$$

Since, $\delta f(s,r) = \frac{\partial f}{\partial s} \delta_s + \frac{\partial f}{\partial r} \delta r$, substituting for $\delta f(s,r)$,

$$\frac{\partial c}{\partial q} \delta q = \frac{\partial c}{\partial (q_f - q_i)} \delta q_f - \frac{\partial c}{\partial (q_f - q_i)} \frac{\partial f}{\partial s} \delta s - \frac{\partial c}{\partial (q_f - q_i)} \frac{\partial f}{\partial r} \delta r.$$

from the samples of air or water.[6] Such bias can be minimized by choosing a data set obtained from tests conducted in a minimum number of laboratories. We, therefore, chose the data set based on water samples collected from the households. For the preparation of this data set, the water samples were tested in not more than five laboratories and the test for a specific constituent, e.g., arsenic, was conducted in not more than two laboratories. In contrast to this, the former data set was based on the records of the community systems themselves and thus the number of laboratories that performed those tests may be as large as the number of relevant test laboratories in the country.

In the household data set, data were available for 1169 households connected to 800 systems, which includes the 614 systems for which cost estimates can be made. When viewed from the system end, there were 549 systems with data on one household each, 186 with two households each, 43 with three, and 22 with three to nine each. For the 549 systems with data on one household each, the household data were considered as a surrogate for the system data. For the other 251 systems with data on more than one household, averages of the household data were calculated and used as proxies for the system data.

Some error may occur from using household water quality as a surrogate for system quality because some additional contaminants may be mixed with the finished water because of bad plumbing within the households. If contaminated plumbing becomes an important factor, water quality of different households connected to the same community system will vary.

---

[6] J. S. Hunter, "The National System of Scientific Measurement," Science 210 (November 1980), pp. 869-874.

In its primary drinking water regulation, EPA identified twenty major constituents that are serious threats to health.[7] EPA established mandatory requirements that the maximum contaminant levels (MCL) of these constituents not exceed specified limits. These limits are legally enforceable for water supply systems serving populations greater than or equal to 25 persons or for systems with connections to 15 or more. In addition, in its secondary drinking water regulations EPA recommended limits on the levels of specific concentrations for five constituents considered to be of minor or uncertain health effects, but important from aesthetic and economic points of view.[8] Data were available for 1167 households connected with the 800 community systems for all the five secondary constituents and four out of the 20 primary constituents; those four constituents are: 1) total coliform bacteria, 2) nitrate-N, 3) lead, and 4) turbidity. In addition, data are available for a larger number of constituents for a limited number of households, about 10 percent of the 1167 households that are connected with 122 community systems. Such data are available for all the primary and secondary constituents except the following four: 1) endrin, 2) toxaphene, 3) 2,4-D, and 4) 2,4,5-TP.

Regarding the number of constituents to be included for analysis, two separate analyses have been carried out: one including the data for four

---

[7]1) Total coliform bacteria, 2) nitrate-N, 3) lead, 4) arsenic, 5) barium, 6) cadmium, 7) chromium, 8) mercury, 9) selenium, 10) silver, 11) fluoride, 12) turbidity, 13) endrin, 14) lindane, 15) methoxychlor, 16) toxaphene, 17) 2,4-D, 18) 2,4,5-TP, 19) gross alpha radioactivity, and 20) gross beta radioactivity.

[8]These five constituents are: sulfates, iron, manganese, color (due to dissolved organic matter), and total dissolved solids (consisting of inorganic salts).

primary and five secondary constituents for all the systems; and the other including the 16 primary and the five secondary ones for 122 systems.

Measurement of Explanatory Variables

Unlike the water quality improvement, the measurement of explanatory variables was straightforward and was divided into two groups: 1) the policy group which includes the variables that can be maneuvered by the systems, and 2) the other variables group which consists of the ones that cannot be easily maneuvered.

Policy Variables: These variables included the size of a system, the population density of the area served, and the type of system ownership. In the literature, the size of water supply systems has been measured by the volume of water produced and by the number of customers served. Some studies have used either of the two in their analyses,[9] while others have used both.[10] For examining scale economies, the quantity of water produced is the appropriate measure of system size. Because the data on the number of customers served typically exclude the distribution of customers between residential and nonresidential users, the number of customers may inaccurately reflect system size.

Closely linked to the size of a system is the size of the area over which a system provides its service. Such a size of a water supply system would have been reflected best by (a) the transmission distance between

---

[9]Bourcier and Forste, op. cit.

[10]A. B. Daugherty and J. D. Jansma, "Economies of Size Among Municipal Water Authorities in Pennsylvania," Southern Journal of Agricultural Economics 5:2(December 1973), pp. 1-6.

the site of acquisition and the centroid of the service area, and (b) the population served per square mile in that area. In our data base, however, neither of these types of data were available. Rather, the number of connections served was the proxy for the population served and the miles of pipeline laid was the proxy for the service area. The best approach would have used the miles of pipeline laid in two different ways: a) the miles of pipeline between the acquisition site and the centroid of the service area as a measure of the transmission distance, and b) the number of connections divided by the miles of pipeline in the distribution network for the service area as a measure of the customer density. Unfortunately, data were not available for attributing the miles of pipeline between the transmission length and the distribution network. Because of this lack of data, a modified measure of customer density was developed where the density was calculated as the number of connections per mile of pipeline and the length of pipeline used in the denominator included both the transmission and distribution stages.

The type of system ownership variable in this framework influences the average production cost because of differences in policy and management aspects rather than economic or technical aspects. In the case of rural water supply, in addition to the public and private profit systems, there is a third important group consisting of individual owners and cooperatives running not-for-profit systems. These were classified as private nonprofit systems.

Other Explanatory Variables: In addition to the three variables that are amenable to policy maneuvering, there are three other explanatory variables that also influence cost but are relatively difficult to maneuver. They are: 1) the region of a system's location, 2) the source of water, and 3) the capacity utilization rate of a system.

The region of a system's location was divided into four geographical categories: 1) Northeast, 2) Northcentral, 3) South and 4) West. Next, the source of water was classified as either ground or surface. Finally, the capacity utilization rate, a continuous variable unlike the other two, was measured as the percentage ratio of the average daily production level of a system to its design capacity.

III. PRELIMINARY ANALYSIS OF THE CONCERNED VARIABLES

In this section, preliminary analyses will be conducted for the independent and dependent variables identified during the preparation of the analytic framework in the preceding section. The primary objective here is to determine the status of rural water systems in regard to finished water quality, production cost and several implementation characteristics. This preliminary analysis will be followed by a multiple regression analysis in the subsequent section where the objective will be to examine whether better alternatives are available to the present mode of providing services by rural water systems.

The presentation in this section will start with analyses of several variables concerning the implementation characteristics of the systems by dividing them into two groups: (1) those which are amenable to policy maneuvering for controlling their influence on production cost, and (2) those which are not maneuverable. These will then be followed by examinations of the status of the quality of finished water and the production costs of water.

Policy Variables

The three policy variables of concern are: (1) the size of a system, (2) the customer density of the service area and (3) the type of system ownership.

The Size of Systems: Two ways of measuring size are by the number of connections, and by the average daily volume of water produced. We will present data for each of these.

The population served by a water supply system often influences the decision on the design capacity of the system and in turn on the production cost of water. Since data were not available on the population served by each system, the number of connections served was considered as its proxy. The limitation here is that the customer served by a connection may be an industrial, a commercial or a residential customer. According to the data gathered from cases representing a population of 34,014 systems, the median and mean number of connections are respectively 63 and 253. The number of connections at the 99th percentile is 2,500 and that at the 75th percentile rapidly drops down to 250. The mode is 20 connections which is very close to the minimum number of connections for qualifying to be a community system, 15.

Next, the average daily volume of water produced by a system is typically used as an indicator of the size of a water supply system. As discussed earlier, a gallon of water produced often is a better indicator of the system size than a connection served, because the customer served by that connection may consume gallons of water over a wide range depending upon its use, industrial, commercial, or residential. However, data on the average volume of water produced are more difficult to obtain compared with that on the number of connections. In our data base, such data were available for only 65 percent of the systems in the population. The median and mean average daily volume of water produced were 0.04 and 0.44 million gallons, respectively. The average volume was 11 million gallons per day (MGD) at the 99th percentile and quickly dropped to 0.14 mgd at the 75th percentile. The mode was 0.002 mgd, close to the 10th percentile average volume.

Connections per Mile Pipeline as Proxy for Population per Square Mile: Data were available to calculate the customer density for 74 percent of the systems. The median and mean values were 30 and 40 connections per mile, respectively. The 99th percentile was 186 connections per mile and the minimum value was one connection per mile. The mode, 25 connections per mile, was close to the median value. For most of the systems in our data base, the margin of error may be minimal for including both the transmission and distribution pipelines in the denominator while calculating the customer density. The transmission length is probably only a small proportion of the total pipeline length for systems because the number of connections per mile of total pipeline was as low as 25 to 40, and the source of water for such systems was probably ground water acquired from within the service area.

Type of Ownership: In our data base, the private nonprofit system was the largest ownership category, 43 percent. This type was closely followed by the public ownership category, 38 percent, with the remaining 19 percent under private profit ownership.

Other Explanatory Variables

Besides the policy variables discussed so far, there are three other explanatory variables. They are: (1) the capital utilization rate, 2) the source of water, and (3) the region of location.

The Capital Utilization Rate: Even the highest value of the average daily volume of water produced was typically 15 to 25 percent below the design capacity, the maximum daily volume producible. The margin was left mainly for fire protection purposes. In addition, the average daily volume produced was often much lower than the highest possible value after putting aside the fire protection margin. The capital utilization rate was one indicator of the level of utilization of a water supply system and was measured by the percentage ratio of the average daily volume of water produced to the maximum daily volume of water a system can produce. This utilization rate primarily reflects the short-run scale economies for a system. In an analysis involving a large number of systems operating at various capital utilization rates, it becomes necessary to consider both the average daily production volume and the utilization rate, because even though daily production volumes are similar for two systems the one operating at a high utilization rate is more likely to be incurring low average production cost.

The capital utilization rate was calculable only for 48 percent of the systems in the data base. As discussed earlier, the average daily volume was known for 65 percent of the systems, but the maximum daily volume was known only for 49.5 percent of the systems. Furthermore, 1.5 percent of the systems had to be weeded out for this computation because of data-reporting errors showing such capital utilization rates higher than 100 percent. The computations indicate median, mean and mode values of 31, 35 and 33 percent, respectively. Although some systems reported data showing a 100 percent capital utilization rate, the 95th percentile was 81, the rate systems operating at the highest capacity should have reported after leaving out the fire protection margin.

Source of Water: The average cost of water is also influenced by the source of water. The primary sources are surface and ground. Since the aquifer itself acts as a storage reservoir for ground water, the capital cost for construction is typically lower for ground water systems. The operational cost of pumping may, however, make the energy cost higher for ground water systems. Also, both the capital and operating costs are lower for ground water systems because they are typically not contaminated with bacteriological constituents and industrial waste. It seems, however, that the ground water systems may not be able to supply water for service areas with large populations. Nevertheless, ground water was the principal source of supply for the systems in our data base because they are mainly rural systems except for cases where rural systems have regionalized with nearby urban systems. According to the data base, 82 percent of the systems have chosen ground water supply and 11 percent have selected surface water. In addition, 10 percent of the systems have purchased treated water from other systems and managed only distribution. It is not known whether the purchased water comes from the ground or the surface. These percentages add up to more than 100, because some systems have used more than one of the three sources; for example, ground water supplemented by purchased water during the seasons of high water usage, generally summer.

Region of Location: In addition to the source, the quality of raw water depends on the geographic region from which it is acquired, and this quality ultimately influences the cost of treating it to produce finished water. In our data base, 12 percent of the systems are in the northeastern and 18 percent in the western regions. A much larger percentage of the systems are from the southern and north-central regions, 41 percent and 29 percent, respectively.

Water Quality Improvement

The quality-related factor that influences average production cost is the difference in water quality between the raw and finished water as a result of subjecting the raw water to various treatment processes. Since about 36 percent of the systems did not treat their water at all, our concern from the cost analysis point of view was with the remaining 64 percent of the systems that did treat water and accomplished various levels of quality improvement.

For analyzing the improvement attained, we needed to know the quality of both finished and raw water. However, quality data for raw water were not available. The collection of such data was not within the scope of the original purpose for which the data base was established. Assuming raw water quality to be determined mainly by the source of the water and the location of its acquisition, the level of quality improvement was analyzed using a cross-tabulation of the finished water quality by the source of water and acquisition location.

Such a cross-tabulation will be presented for each of the two separate measurements of the finished water quality, the laboratory test-based objective measurement and the users' subjective measurement.

Objective Quality: The concern will be mainly with the bacteriological and chemical constituents which have been identified by the Environmental Protection Agency as serious health threats and for which that agency has set up legally enforceable maximum contaminant levels (MCL). Also of concern will be a number of chemical constituents that have been judged to have minor or uncertain health effects but are of importance from aesthetic and economic points of view. For these constituents, EPA has recommended maximum contaminant levels. For the purposes of analysis here, we need to know if there is sufficient variation among the contamination levels of treated water by each constituent for systems within a specific region, and using a specific source of water. If such variation exists, the differences in the water quality improvements among the water supply systems could be used to analyze which factor explains the variations in the average production cost of water.

To recapitulate the status of the data base, data are available for the total coliform, nitrate-N, lead, turbidity, and all the secondary constituents for 21,979 systems that treated water out of a total population of 34,014 systems. But, for the remaining primary constituents, data are available only for a sample of 2,572 systems that treated water. However, this difference is not significant for discussions in this section, because the major criterion of classifying systems for each constituent is whether the contaminant level for the water produced by a system is less than, equal to or above the MCL. The divisions of systems between these two categories are expressed in percentages and are presented in Table 1.

According to this classification, the data on barium and chromium contamination levels will not be helpful in explaining production cost variations, because for the 2,572 sampled systems, the contaminant levels of these constituents are less than the MCL values. For both ground and surface water sources in each of the four regions, almost all the systems have contamination levels of selenium and fluoride below or equal to their MCL values. The constituents whose data may be the most helpful in explaining cost variations are total coliform, mercury and lead in the primary constituents group and iron, manganese and the specific conductance or the dissolved solids, in the secondary constituents group. The remaining constituents may be included in the regression equation, but they may be of limited help; these are nitrates, turbidity, cadmium and silver in the primary constituents group, and sulfates and color in the secondary constituents group.

Table 1. Crosstabulation of Finished Water Quality (Objective) for Rural Community Systems That Treated Their Water, by Source of Water and by Region, Percentages, 1978

		Ground Water				Surface Water			
		North-east	North Central	South	West	North-east	North Central	South	West
Primary Constituents (21,979 systems)									
Total Coliform	:0*	86	79	68	84	49	93	89	44
	:1	14	21	32	16	51	7	11	56
Nitrate-N	:0	100	96	99	98	100	77	100	100
	:1	0	4	1	2	0	23	0	0
Lead	:0	93	90	85	70	96	88	91	96
	:1	7	10	15	30	4	12	9	4
Turbidity	:0	99	90	100	100	100	96	81	95
	:1	1	10	0	0	0	4	19	5
Primary Constituents (2,572 systems)									
Arsenic	:0	100	33	100	100	100	100	100	100
	:1	0	67	0	0	0	0	0	0
Cadmium	:0	100	96	72	100	100	88	100	100
	:1	0	4	28	0	0	12	0	0
Mercury	:0	31	33	71	96	68	88	72	0
	:1	69	67	29	4	32	12	28	100
Selenium	:0	100	96	100	96	100	100	100	100
	:1	0	4	0	4	0	0	0	0
Silver	:0	100	100	100	24	100	100	66	100
	:1	0	0	0	76	0	0	34	0
Fluoride	:0	100	93	98	100	100	100	100	100
	:1	0	7	2	0	0	0	0	0
Secondary Constituents (21,979 systems)									
Sulfate	:0	100	85	91	87	100	100	100	97
	:1	0	15	9	13	0	0	0	3
Iron	:0	93	74	85	100	57	96	71	95
	:1	7	26	15	0	43	4	29	5
Manganese	:0	92	66	89	93	86	84	68	100
	:1	8	34	11	7	14	16	31	0
Color	:0	100	97	99	100	67	100	82	95
	:1	0	3	1	0	32	0	18	5
Specific Conductance	:0	61	18	50	69	98	92	82	28
	:1	39	82	50	31	2	8	18	72

* 0: less than or equal to the maximum contaminant level allowable (primary) or recommended (secondary).

1: greater than the maximum contaminant level allowable (primary) or recommended (secondary).

The data presented in Table 1 also demonstrate the importance of using the source of water and the location of region as dummy variables in the same equation to serve as the proxy for raw water quality. The importance of the source of water in influencing the finished water quality is clearly seen in the cases of the total coliform, mercury and the total dissolved solids. For these three constituents, the percentage of systems with finished water quality above, below or equal to MCL levels changes drastically depending on the source of water. For example, 86 percent of the ground water systems in the Northeast have the total coliform count less than or equal to MCL while only 49 percent of the surface water systems in the same region have similar results. For mercury, the results appear even more prominent; for all the regions except the South, the percentage of systems that satisfy the MCL requirements changes from relatively low to high values or vice versa as the source of water changes from ground to surface. Similarly, the percentage of systems meeting MCL levels becomes high or low for the total dissolved solids as the source of water changes for the North Central and the Southern regions.

The location of the system appears to be a less important influence on finished water quality. There are, however, a number of cases where its importance is seen. For example, the number of ground water systems that met the MCL requirements for mercury is 71 percent of the sampled systems for the South compared to 31 and 33 percentages in the Northeast and the North Central, respectively. The percentage is even higher for the West, 96 percent, but it is difficult to say how much of this is due to the source and how much due to the location, because such a percentage changes to zero for the surface water systems.

The systems serving purchased water were not included because their original source of raw water, ground or surface, was unknown.

Subjective Quality: As an alternative to the laboratory test data for finished water quality, the users' perceptions may also be useful for explaining production cost variations due to water quality improvement. The assumption here is that the better the users' perceptions of finished water quality within a specific region for a certain source of water, the higher will be the treatment cost and thus the production cost. This assumption, however, has two limitations which should be considered while interpreting this data. First, users typically judge water in terms of palatability. Yet the purest water is not necessarily the most palatable. Distilled water, for example, is generally unpalatable. Second, treatment of water often produces certain unpalatable characteristics during the process of eliminating contaminants. Chlorine treatment, for example, produces an unpleasant odor.

Five characteristics have been identified in the NSA data base for subjectively evaluating finished water quality: taste, smell, clarity, color and sediment. These characteristics are classified into six categories that range from complete absence of that characteristic to constant presence. For example, for the taste characteristic, they are: never any taste (coded as 0); generally no taste (1); occasional slight taste (2); prevalent taste (3); generally strong taste (4); and constant strong taste (5). Data for classifying finished water quality into one of these categories are generated from various combinations of the users' responses to two questions dealing with the presence or absence of the characteristic and its frequency of occurrence for a period extending up to one year from the date the interview was taken.

These six categories have been regrouped into two classes for the specific purpose of determining whether the data on any of these six characteristics will be helpful for explaining production cost variations. The two categories are: complete absence of a characteristic coded 0 and presence of the characteristic with varying frequencies of occurrence, coded 1. Table 2 presents a crosstabulation for these characteristics based on the two-class method by the source of water and by the region of location.

Table 2. Crosstabulation of Finished Water Quality (User Perceived) for Rural Systems That Treated Their Raw Water, by Source of Water and by Region of Location, Percentages, 1978

		Ground Water				Surface Water			
		North-east	North Central	South	West	North-east	North Central	South	West
Taste	:0*	33	53	43	36	29	41	68	80
	:1	67	47	57	64	71	59	32	20
Smell	:0	66	76	68	59	28	50	59	18
	:1	34	24	32	41	72	50	41	82
Clarity	:0	84	61	58	59	29	74	54	90
	:1	16	39	42	41	71	26	46	10
Color	:0	74	65	70	97	37	70	78	30
	:1	26	35	30	3	63	30	22	70
Sediment	:0	84	75	58	95	66	57	82	24
	:1	16	25	42	5	34	43	18	76

* 0: less than or equal to the maximum contaminant level allowable (primary) or recommended (secondary).
1: greater than the maximum contaminant level allowable (primary) or recommended (secondary).

All five quality characteristics have water supply systems distributed between the two classes. They are also close to one-half each in most cases. This is quite different from the distribution observed for the laboratory tested quality data. It seems, therefore, that the perceived data may be helpful in explaining cost variations within the limitations discussed earlier.

The data also reinforce the fact that variation in finished water quality depends on the source of water. For smell, clarity, color and sediment in the Northeast, the percentage of the total number of systems below or equal to MCL in each class changes from relatively high to low values as the source of water changed from one to another. Similarly, in the West for all the characteristics except clarity, such percentages changed directions from high to low values or vice versa.

Across the regions, the change in the percentages within each class appear relatively less prominent. The Northeast had less percentage in the absence of taste class compared with the other regions both for the ground and surface water sources. But in other cases it is difficult to attribute the percentage differences between the raw water source and the region of location. Because of this problem, we cannot say that the region of location is not important. Therefore, during the regression analysis the region variable may be a useful one to include as one of the proxies for raw water quality.

Average Cost of Production

Average cost of production was computed from the data available in the data set by using two different bases, the volume of water produced and the number of connections served by the water supply system. The total cost of production was estimated by two types of techniques, the accounting method and the economic method. We, therefore, have two estimates each for the average cost with each basis.

Volume-Based Average Cost: Table 3 presents summary statistics for the volume-based average cost estimates computed by using the two techniques. For the estimate prepared using the accounting method, the results were based on the data for 12,936 out of 34,014 systems; for the remaining systems, data were missing for one or more items necessary for the estimation. Similarly, for the economic method estimate, the results were based on the 2,964 (out of 6,788) systems that were able to provide original construction cost information.

For comparison between the cost estimates made by the two techniques, estimates have been prepared using the accounting method on the same 6,788 systems that were able to provide the original construction cost information necessary for the estimation by the economic method (Table 3). The mean value of the cost estimate prepared by the economic method is 2.65 times that prepared by the accounting method. Similarly, the median and mode also have values larger by 2.39 and 2.89, respectively, for the economic method.

There are a number of reasons why the cost estimates of the economic method are larger than those of the accounting method. First, the accounting method does not include the depreciation cost of the water system in constant dollars. In the economic method, such a cost has been included by assuming a depreciation period of 50 years.[11] This

---

[11] The depreciation period suggested for water supply systems varies from 40 to 67 years. In Johnson and Hobgood's study of rural water systems in Louisiana, a depreciation period of 40 years has been used. In

depreciation cost is included to provide for the replacement of worn-out equipment and structures. As a result, the depreciation cost item of the capital cost component provides for funds for an ongoing water supply service.

Table 3. Average Cost of Water Production for Rural Community Systems, 1978

	Accounting Method		Economic Method
Total Number of Systems	34,014	6,788	6,788
Number of Systems Used	12,936	2,884	2,964
	(1978 dollars per thousand gallons)		
Mean	1.40	1.62	4.30
Median	0.89	1.14	2.73
Mode	0.33	0.59	1.59
Standard Deviation	2.20	1.68	4.66
Coefficient of Variation	1.57	1.04	1.08

In the accounting method, the debt retirement cost for the funds obtained as loans does typically include a component for the retirement of the principal sum. Nevertheless, the accumulated amount of such a principal retirement component is much less than the replacement cost, especially during the periods of high inflation rates, because this

---

U.S. Treasury Department/Bureau of Internal Revenues, Bulletin "F": Income Tax Depreciation and Obsolescence, Estimated Useful Lives and Depreciation Rates (Washington, DC: Government Printing Office, 1948), p. 67, 67 years has been suggested as the average life of typical plant and system. Finally, 50 years has been recommended as system life in U.S. Treasury Department/Internal Revenue Service, Depreciation: Guidelines and Rules (Washington, DC: Government Printing Office, 1962), p. 21. We chose a period of 50 years because rural systems have less sophisticated structures compared with typical water supply systems which include a population of both urban and rural utilities.

component is not adjusted for price increases between the year of water system construction and each year of debt retirement. In contrast to this, in the economic method, both the depreciation cost and the interest cost components are adjusted for such price increases.

Next, the accounting method does not include the opportunity cost of the capital investment in structures and equipment from government grants and from the system's internal reserve funds, where such cost is the return foregone from alternative investment. For the purposes of calculation, the long-term government bond yield for 1978, 7.9 percent, has been chosen as the interest rate. This rate appropriately represents the rate of return the systems could have received by investing in an alternative, long-term investment instead of investing in the water systems. Moreover, it represents the rate the government pays to borrow from the public and is therefore an appropriate measure of the cost of its grants.

Furthermore, the 7.9 percent rate has been used to calculate the interest component of the capital cost for public and private interest-bearing sources. First, to make the capital costs comparable with each other, it was necessary to choose a standard rate for all the systems instead of using the various rates actually paid. (When standardized for 1978 by using the long-term government bond yields, the mean values of borrowing costs for all systems were 6.4 percent for EDA loans, 7.2 percent for FmHA loans, 7.4 percent for municipal bonds and 9.8 percent for commercial loans.) Second, the alternative cost foregone for the capital loan available is not the cost systems pay, because these costs are subsidized by the government either directly through low interest loans or indirectly through federally guaranteed bonds. It is the highest alternative cost foregone for long-term capital loans issued in 1978 that should be used for our calculations. The appropriate rate is thus 7.9 percent.

Number Of Connections-Based Average Cost: Table 4 presents summary statistics for the connections-based average cost estimates prepared by using both the accounting and the economic methods. For the accounting-method estimate, the results were based on the data for 18,731 out of 34,014 systems; data were missing for one or more items necessary for this estimation for the remaining systems. In addition, estimates were prepared by using both the economic and accounting methods on the basis of the data for 4,412 and 4,317 systems, respectively. These are the systems for which there were no missing values in the data subset of 6,788 systems for which data were available for their original construction cost. As mentioned earlier, the estimation was made for the same data subset by using both the methods mainly for comparative purposes.

Table 4. Average Cost of Water Production for Rural Community Systems, 1978

	Accounting Method		Economic Method
Total Number of Systems	34,014	6,788	6,788
Number of Systems Used	18,731	4,317	4,412
	(1978 dollars per connection per year)		
Mean	116	136	370
Median	73	67	209
Mode	75	239	401
Standard Deviation	145	147	510
Coefficient of Variation	1.25	1.08	1.38

The mean value of the estimates prepared by the economic method is 2.72 times that by the accounting method. Also, the median and mode are 3.12 and 1.68 times larger, respectively. The reasons behind these differences are the same as those discussed earlier. The relative ratios here, however, are different from that for the volume-based average cost estimates presented in Table 3, because the numbers of systems included for computing summary statistics were different because of the larger number of missing values in the volume-based estimation.

IV. MULTIPLE REGRESSION ANALYSIS

In this section, we will examine the intrerrelationships between the dependent and independent variables by using multiple regression techniques. The average cost is a dependent variable and the policy and other variables are the independent ones. Although the status of the water quality improvement variable is less clear, it does qualify as an independent variable if we recapitulate the fact that water systems are legally required to attain a certain minimum level of contaminant levels set by the EPA. This certainly is not a policy variable because, unlike the size of a system, the requirements of the levels of finished water quality are not endogenously set by the water system's management body.[12] It is, however, an independent variable in the sense that the exogenously set EPA requirements will influence the average cost of producing drinking water. Therefore, during the multiple regression analyses in this section, the average cost will be treated as the dependent variable and the water quality improvement together with the policy and other explanatory variables as independent variables.

The presentation of this analysis section is organized along the following lines. An introductory subsection to the analysis will be presented where the specific data used for the analysis, the model specifications and the prior expectations of relationships between the dependent variable and the independent variables will be discussed. This will be

---

[12]It should be noted that the water quality improvement has some policy variable type of characteristics in the sense that water systems may often choose to attain water quality higher than that required by EPA. But within the scope of this study, we are concerned mainly with the attainment of the EPA requirements.

followed by the analysis subsection where multiple regression equations will be presented.

Introduction to the Regression Analysis

Data Subsetting: Only a subset of the data base available will be used for analysis. The concerned factor related to water quality is the quality improvement, i.e., the difference in quality between the raw and the finished water for each system.

But in our data base, some of the water supply systems did not treat their water at all; there were no quality improvements in the water they produced. We have therefore chosen to include only those water systems that treated their water.

This subset gets narrowed further for the regression estimates using the cost estimates based on the economic method. In those cases, the systems included are the ones that had no major alterations done since their original construction. This subsetting became necessary, as discussed earlier, because the economic method of estimation requires data on the original construction, and during the 1978 NSA survey, data on original construction were gathered only if the systems had no major alteration conducted.

These layers of subsetting changed the types of systems included in the analysis from the 800 systems in the NSA survey to 664 systems. We have therefore presented summary statistics for the continuous and categorical variables of our concern in Tables 5.1 and 5.2, respectively, for the original data set and its two subsets--1) the systems that treated water, and 2) the systems that treated water and had no major alterations since their original construction. In Table 5.1, the statistics are presented for the continuous variables and in Table 5.2 for the

Table 5.1 Summary Statistics for Selected Continuous Variables for the Set of All Water Systems and Two Subsets

			All Systems				Systems That Treated Water				Systems That Treated Water and had no Major Alt. Since Const.			
	Unit	MCL	Median	Mean	Std. Dev.	Coeff. of Var.	Median	Mean	Std. Dev.	Coeff. of Var.	Median	Mean	Std. Dev.	Coeff. of Var.
Dependent Variables:											2.80	4.57	5.87	128%
Volume-based average production cost	$/1000gal.		0.89	1.40	2.20	157%	0.89	1.44	2.38	165%	1.14	2.00	2.08	104%
Number-of-connections-based avg. prod. cost	$/conn/yr.		72.54	115.73	144.75	125%	82.09	128.51	156.29	122%	84.39	111.27	111.58	100%
											184.34	331.30	514.84	155%
Independent Variables:														
Number of connections	# conn.		63	253	1,134	448%	105	330	1,387	420%	150	256	552	216%
Annual volume of water produced	thou. gals.		14,600	162,333	1,736,829	1,070%	18,250	205,987	1,995,706	969%	12,118	97,295	697,490	717%
Capital utilization rate	%		31	35	25	71%	34	39	22	56%	22	31	20	65%
Number of connections per mile of pipeline	# conn/mile		30	40	41	103%	30	40	43	108%	37	46	75	163%
Laboratory-test-based finished water quality:														
Total coliform	bacterium/100ml	1	0	4,180	48,633	1,163%	0	3,781	48,278	1,277%	0	26	117	450%
Lead	mg/l	0.05	0.009	0.028	0.046	164%	0.009	0.025	0.038	152%	0.010	0.030	0.046	153%
Nitrate-N	mg/l	10	0.3	1.07	2.35	220%	0.3	0.96	2.02	210%	0.4	1.13	2.31	204%
Turbidity	NTU	1-5	0.4	0.96	4.13	430%	0.35	0.72	1.40	194%	0.3	0.44	0.52	118%
Sulfate	mg/l	250	20.5	71.4	121.9	171%	30.6	85.9	140.3	163%	15	40.7	46.4	114%
Iron	mg/l	0.3	0.1	0.22	0.50	227%	0.1	0.26	0.61	235%	0.1	0.15	0.53	353%
Manganese	mg/l	0.05	0.03	0.05	0.09	180%	0.03	0.05	0.10	200%	0.03	0.05	0.14	280%
Color	color units	15	5	5.3	6.4	121%	4	4.9	5.0	102%	4	4.2	2.2	52%
Specific conductance	micromhos	500	484	545	446	82%	510	563	445	79%	510	421	243	58%
User-perception-based finished water quality:														
Taste	number		0	1.10	1.52	138%	0	1.29	1.52	118%	0	1.07	1.38	129%
Smell	number		0	0.76	1.29	170%	0	0.94	1.36	145%	0	0.54	0.99	183%
Clarity	number		0	0.78	1.20	154%	0	0.89	1.28	144%	0	0.77	1.20	156%
Color	number		0	0.54	1.03	191%	0	0.61	1.09	179%	0	0.39	0.91	233%
Sediment	number		0	0.57	1.00	175%	0	0.66	1.05	159%	0	0.59	0.94	159%

Note: MCL stands for Maximum Contaminant level.

Table 5.2 Summary Statistics for Selected Categorical Variables for the Set of All Water Systems and Two Subsets

	Unit	All Systems Number	All Systems Percentage	Systems That Treated Water Number	Systems That Treated Water Percentage	Systems That Treated Water and Had No Major Alteration Since Construction Number	Systems That Treated Water and Had No Major Alteration Since Construction Percentage
Independent Variables:							
Type of Ownership:							
Public	# of systems	13,028	39%	10,432	48%	1,458	54%
Private	# of systems	20,619	61%	11,327	52%	1,263	46%
Private nonprofit	# of systems	14,567	43%	8,103	37%	805	29%
Private profit	# of systems	6,052	18%	3,224	15%	458	17%
Source of Water:							
Primary	# of systems	31,594	90%	19,686	85%	2,275	80%
Ground	# of systems	27,841	79%	16,479	71%	1,886	66%
Surface	# of systems	3,753	11%	3,207	14%	389	14%
Secondary	# of systems	3,501	10%	3,330	15%	559	20%
Region of Location:							
Northeast	# of systems	4,018	12%	2,666	12%	434	16%
North Central	# of systems	9,823	29%	6,341	29%	1,019	37%
South	# of systems	14,057	41%	10,249	47%	1,167	43%
West	# of systems	6,116	18%	2,723	12%	114	4%
Users' Perceptions Based Finished Water Quality:							
Taste	0		61		52		54
	1		3		5		7
	2		12		19		27
	3		16		16		5
	4		3		5		1
	5		4		4		5
Smell	0		72		64		69
	1		4		5		17
	2		5		7		5
	3		17		20		6
	4		2		3		1
	5		1		.4		6
Clarity	0		64		59		66
	1		11		12		13
	2		14		15		2
	3		8		8		20
	4		2		3		.3
	5		2		2		0
Color	0		75		73		80
	1		6		5		10
	2		10		13		2
	3		8		8		6
	4		1		2		2
	5		1		.1		0
Sediment	0		74		70		71
	1		1		1		1
	2		20		24		26
	3		4		4		2
	4		.5		1		0
	5		.2		.4		0

categorical variables. Data on subjective water quality are presented in both the tables--once, considering it continuous between its values from 0 to 5 and then considering it categorically for the six values of 0, 1, 2, 3, 4 and 5.

For a comparison of these summary statistics using a single measure, the coefficient of variation has been computed for the continuous variables in Table 5.1. It is observed that the value of this coefficient does not change by more than 10 percent by subsetting from all systems to the systems that treated their water; the exceptions are the coefficients for the capital utilization rate, turbidity, color, taste and smell. However, when the data base is subsetted further to the systems that had treated their water and had no major alterations undertaken, the coefficient of variation changes by more than 10 percent for all the continuous variables except the capital utilization rate, lead, nitrate-N and the perceived quality variables. For the total coliform, the ratio of the two coefficients is as high as 2.6. More about the effects of this subsetting on the results of the regression analysis will be discussed when the equations are presented.

Model Specifications

During the multiple regression analyses, the average costs of production, both for volume and number-of-customers-based, are introduced in logarithmic forms. Also, logarithmic forms are introduced for four of the explanatory variables: annual volume of water produced, number of connections served, number of connections per mile of pipeline laid and total coliform contamination level. The purpose behind presenting these variables in such a form is to reduce the number of large residual errors that may arise because of the wide range of their numerical values. In case of the total coliform variable, a relatively small number, 0.5, was added to its value before calculating the logarithms because a large number of its observations has zero values.

In addition, three of the explanatory variables are introduced as dummy variables: source of water, region and type of ownership. In the case of the water source, purchased water was considered the reference category, and two dummy variables were created: ground source and surface source. The value for the ground source variable was set at '1' for ground water and at '0' for surface and purchased water. Similarly, the value for the surface source variable was set at '1' for surface water and at '0' for ground and purchased water.

The Western region was used as the reference category for the region variable and three categorical variables were created: Northeastern, North Central and Southern. For each of these, values were set at '1' if the water supply system was located in that region and at '0' if not.

The type of ownership is the last of the three explanatory variables introduced as a dummy variable. The reference category in this case was the private nonprofit class of ownership. The two variables created were public ownership and private ownership. The values were again set at '1' if a system fell under its type of ownership and at '0' if not.

Prior Expectations of the Relationships

So far we have described the data subsets that have been used for regression runs and the forms in which the dependent and the independent variables are introduced. In this subsection, we will discuss the types of relationships we expect between the dependent variable and each of the independent variables. The speculative discussion regarding the prior expectations is meant primarily for establishing a basis for technical discussions about the signs of coefficients during the presentation of the regression results. The a priori expectations have been used for deciding to include variables in the model only for the case of the finished quality variables because, for the coefficients of those variables, theoretical and intuitive judgments about their signs are strong enough to refute unexpected results. The basis for such judgments will be presented during the discussion of a priori expectations for finished water quality variables.

An increase in the log values for annual volume produced and for number of connections is expected to produce a decrease in the log values for average cost. This behavior is expected because the percentage change in the fixed cost component would typically be smaller than the percentage change in the water volume produced or the number of connections. The implication in terms of the regression equations is that these variables would have coefficients whose values are less than zero. That is, the log values for water volume and for the number of connections will be negatively related to the log values for average cost.

For the log values for the number of connections per mile of pipeline, an increase in customer density would generate an increase in the level of output demanded without increasing the fixed cost component of the total cost. The implications of this relationship for the log of average cost is that the log of number of connections per mile variable would have a coefficient with a negative sign in the regression equations.

Finally, for the type of ownership factor, the prior expectation based on the discussions during the preliminary analysis is that the private nonprofit ownership category is likely to be associated with an average cost less than that for both the public ownership category, and the public form, in turn, is likely to be less than the private profit class.

For the source of water variable with the three classes of ground, surface and purchased water, the prior expectation is that ground water would have low average cost relative to surface water. We cannot, however, predict similar behavior for purchased water because data were unavailable to trace the source of that water. An alternative model design could have been selected where two variables would have related to the source of water: one that classified the source by primary (or produced) and secondary (or purchased) and another that categorized by ground and surface. In such cases, purchased water could be expected to have lower average cost than the produced water. But by classifying the source of water in this fashion, the degrees of freedom in the equations would have been reduced because all the systems using purchased water would have missing values for the ground-surface water variable. The source of water variable was therefore introduced by using the purchased source as the reference category and by creating two variables. The implications of these prior expectations on the ground-source and surface-source dummy variables will be such that the ground-source variable will have a coefficient less than that for the surface-source variable.

For introducing the region variable, three categorical variables have been created using the Western region as the reference category. On the basis of the NSA data on water prices, the prior expectations of relationships are that the Western region would have the highest average cost followed by the North Central, the Northeastern and the Southern regions. The implication of this expectation on the regression equations is that all the three created variables would have negative coefficients. Also, the coefficients will have smaller values for the North Central relative to the Northeastern and the Northeastern smaller relative to the Southern region variable. But prior expectation based on the NSA data may not be directly applicable in this context because their calculations included all water supply facilities with two or more connections.

Finally, the capital utilization rate variable is expected to have a coefficient with a negative sign. Any increase in the ratio of the average daily water produced to the design capacity would typically reduce the average cost, because the fixed cost gets distributed over a larger level of output.

In addition to the policy and other explanatory variables, there are twenty finished quality variables as independent variables, fifteen of which are based on the laboratory test data and five on the users' perception data.

The expected relationship for water quality improvement is that the value of the log of average cost will increase as the water quality improves, because the water quality improvement typically involves additional treatment facilities which in turn will increase the average cost. That is, a lower level of quality improvement will leave a higher level of contaminants in the finished water but will incur lower average cost.

Intuitive judgment here needs to be reinforced by a theoretical argument to ensure that for this kind of model the coefficient of the finished quality variables should have negative signs. During the presentation on measurements, p.8, we contended that the source of water and the region variables can act as surrogates for raw water quality. If these surrogate variables are included in a regression equation, it may not be unreasonable to expect the log of average cost to go down for increasing values of the contaminants in the finished water.

Such an expectation has a theoretical support from the mathematical note on pages 8–9. If the source of water and the region are introduced as surrogate variables for raw water quality along with the finished water quality variables, the coefficient of the finished water quality variable would be $\partial c/ \partial(q_f - q_i)$ instead of $\partial c/ \partial q_f$. This reflects the sensitivity of the log of average cost to the water quality improvement, not to the finished water quality. Therefore, the coefficients of the finished quality variables in our equations should have negative signs.

Multiple Regression Equations

The multiple regression equations will be presented in two subsections, one each for the two dependent variables--the volume-based average cost and the number-of-connections-based average cost. Within each subsection, separate equations will be presented for various combinations of the two methods of estimating production cost, the accounting and the economic methods, and of the two schemes of measuring finished water quality, the laboratory test-based and the users' perception-based measurements. Also, the relative importance of the independent variables in each equation for explaining the log of average cost will be analyzed by using stepwise regression techniques.

Volume-Based Average Cost of Production: The first group of models presents the relationship between the volume-based average cost of production and the explanatory variables using different methods of estimating total cost and of measuring finished water quality (Table 6).

The first three equations in Table 6 were estimated by using the cost estimates based on the accounting method and the remaining two by using those based on the economic method. Within the first three equations, there are differences in the data subset used and in the measure or extent of coverage of the finished water quality. The first and the third equations were estimated by using a data subset of all the systems that treated their water, while the second was estimated by using an exactly similar subset but from a different sample of systems. The systems in that sample were sampled in the NSA survey for collecting data on all the contaminants included in EPA's interim primary drinking water regulations. Furthermore, the first and the third equations were estimated by using different quality measures, the laboratory test-based and the user perception-based measurements, respectively. Similarly, the last two equations in Table 6 were estimated by employing different measures of quality.

In the first and second equations, the coefficients of the annual water volume have a value less than zero as expected. Also, the capital utilization rate and the customer density variables are negatively related to the average cost, also according to the a priori expectations.

Regarding the dummy variables, both equations have coefficients that indicate that ground water has a lower average cost than that for surface

Table 6. Multiple Regression Equations Explaining Log (Volume-Based Average Cost of Production)

	Equations				
	No. 1	No. 2	No. 3	No. 4	No. 5
Degrees of freedom	288	33	282	40	32
Multiple corr. coefficient, R^2	0.35	0.64	0.28	0.61	0.39
Intercept	1.7569	2.0301	1.3843	2.5155	1.5750
Log (Annual Volume of Water Produced)	-0.2833 (7.07)	-0.3108 (2.18)	-0.2400 (5.81)	-0.4603 (4.48)	-0.3690 (2.69)
Capital Utilization Rate	-0.0015 (1.72)	-0.0038 (1.48)			
Log (connections per pipeline mile)	-0.1236 (1.84)	-0.3317 (2.71)			-0.0485 (0.21)
Source of water: ground = 1	-0.1596 (2.48)	0.0417 (0.27)	-0.2327 (3.17)	0.2976 (1.83)	0.2384 (1.13)
Source of water: surface = 1	0.0609 (0.86)	0.3522 (1.45)	0.0196 (0.25)	0.2958 (1.59)	0.3500 (1.43)
Region: Northeast = 1	-0.2932 (3.07)	-0.0623 (0.11)	-0.0934 (0.87)	0.3700 (1.71)	-0.0718 (0.16)
Region: North Central = 1	-0.1988 (2.12)	-0.2970 (0.51)	-0.0323 (0.31)	-0.1023 (0.66)	-0.2622 (1.53)
Region: South = 1	-0.0977 (1.14)	-0.0635 (0.11)	0.1781 (1.82)		
Type of ownership: Public = 1	-0.1310 (1.93)		-0.2154 (3.19)	0.2750 (1.68)	0.2568 (1.08)
Type of ownership: Private profit = 1	-0.1603 (1.80)	-0.2855 (0.44)	-0.2052 (2.23)	0.2559 (0.64)	0.1765 (0.32)
Log (total coliform + 0.5)	-0.0668 (2.59)			-0.1134 (1.37)	
Lead				-3.7384 (2.36)	
Nitrate-N	-0.0263 (2.99)			-0.0436 (2.34)	
Iron	-0.0402 (1.05)				
Manganese	-0.2538 (1.37)	-0.3432 (0.41)			
Color	-0.0148 (1.92)	-0.0368 (1.00)		-0.0598 (2.09)	
Specific conductance				-0.0004 (1.56)	
Cadmium		-6.0335 (0.63)			
Smell					-0.1136 (1.58)
Clarity			-0.0469 (2.85)		

Note: Numbers in parentheses are t-ratios.

water. But the first one reveals that ground water is cheaper than the purchased water while the second one implies the opposite. Nevertheless, both equations appear to be accurate from the source of water point of view because this equation should not be used to compare ground and surface water sources with the purchased water source.

Next, regarding the region dummy variable, the first equation shows that the Northeastern region has the lowest average cost followed by the North Central, Southern and Western regions, in that order. But the second equation shows that the North Central region has the lowest average cost followed by the Southern, Northeastern and Western regions. This discrepancy may have occurred because the second equation was estimated on the basis of a sampled data subset.

Furthermore, regarding the ownership dummy variable, both equations reveal that private profit ownership is cheaper than the private nonprofit ownership. In addition, the first equation indicates that public ownership is cheaper than the private nonprofit ownership but expensive compared with the private profit ownership. These observations are contrary to expectations. We will discuss this discrepancy in detail when we compare the first equation, based on the accounting method, with the fourth one, based on the economic method.

Finally, regarding the laboratory test based quality variables, in the first equation total coliform, nitrate-N, iron, manganese and color variables were included and have coefficient signs as expected. In the second equation only the color variable was included because it was highly correlated with iron. The cadmium and manganese variables were included in this equation, but they have t-ratios less than unity. On the whole, it seems we have not been able to gain much in the second equation regarding the number of quality variables that could be included to explore their relationships with the average cost.

The third equation is quite similar to the first except that the perceived clarity variable substitutes for the laboratory test based variables in the first one, and not in the second. It seems the appropriate criterion to compare these two equations will, therefore, be the multiple correlation coefficient. Since the first equation has a higher coefficient, the quality measure based on the laboratory test appears to be preferable to that on the users' perception. This preference, however, is based purely on the consideration of how well the explanatory variables indicate the variations on the average production cost. From the point of view of explaining other considerations such as how interested a community will be to improve its system's water quality, the equation based on the user perception quality variables may be the appropriate one. This issue will be addressed in a subsequent publication.

Next, we will compare the estimates in the first equation with the better of the two equations developed by using the cost estimates based on the economic method, that is, the fourth and the fifth equations. Because these equations are similar in many respects, the multiple correlation coefficient again appears to be a useful criterion to judge the usefulness of quality measure. The fourth has a higher R^2 value and the laboratory test based quality measure therefore appears to be superior.

Finally, compare the first and fourth equation. The main difference between these two equations is in terms of the data subset used to estimate them. Because the fourth equation was prepared by using average costs estimated by the economic method, only those systems that treated their water and had not had any major alteration since their original construction were included in its data subset. These were the only systems for which data on the total capital investment were available. As a result of this restriction on the creation of a data subset, the coefficients of the created variables related to the source of water and to the region do not reveal much about their relationships with the average cost in the fourth equation. Nevertheless, they were included in the equation to reflect water quality improvement along with the finished water quality variables. The Southern region variable was the only one excluded, because the t-ratios of the other two region variables became very small in its presence.

As far as the laboratory test based quality variables are concerned, both equations include five variables each and their coefficients have expected signs and t-ratios greater than unity.

The main strength of the fourth equation, however, lies in the estimation of coefficients for the two variables created for discovering the relationship of the type of ownership variable with the average cost variable. There are differences in the ways systems under various ownership categories compute their debt retirement costs and in the costs of capital they have to pay. The systems under private profit ownership typically do not include the principal component in their debt retirement costs and, during the NSA survey, they may not have included the depreciation cost because questions were asked about revenue and percentages expended for debt retirement and about operating and maintenance costs. In the fourth equation, the average cost variations due to such computational differences were standardized by the economic method. The revealed effects are therefore more likely to reflect the differences due to the ownership-related management aspects. This equation indicates that the private nonprofit systems had the lowest average costs and that the private profit and the public systems produced water at similar average costs.

The results of the stepwise regressions (Table 7) indicate that the annual volume of water produced is the most important explanatory variable for both the equations. This variable is then followed by the total coliform in the first equation and by lead in the fourth. Total coliform drops from the second-most important position in the first equation to the seventh in the fourth equation. The reason for this change in relative importance is that the coefficient of variation changes from 1163 percent for the total coliform data subset used for the first equation to 450 percent for that utilized for the fourth one, (see Table 5.1).

To sum up, both the first and the fourth equations are useful for exploring relationships between the explanatory variables and the average cost.

Number-of-Connections Based Average Cost of Production

The multiple regression equations presented so far have been estimated by using the logarithmic form of the water production cost per one thousand gallons as the dependent variable. From the point of view of examining relationships with independent variables such as the size of systems, the volume of water produced is definitely a superior basis compared with the number of connections, especially in cases where data are not available for differentiating between residential and nonresidential users. During the preliminary stages of planning for new rural water systems, decision makers may be concerned with the total annual production cost of water using a specific technique and the number of families served in the community. In this case, the average cost per family per year will be the criteria for comparing alternative techniques of producing water. We have, therefore, chosen to conduct and include additional regression analyses similar to the earlier ones by substituting the volume-based average cost with the number-of-connections based average cost.

Table 7. Relative Importance of Independent Variables in the Multiple Regression Equations Explaining the Log (Volume-Based Average Cost)†

Independent Variables	Equation 1* Rank	Equation 4** Rank
Log (Annual volume of water produced)	1	1
Log (Total coliform + 0.5)	2	7
Color	3	4
Nitrate-N	4	-
Log (No. connections per pipeline mile)	5	3
Manganese	6	-
Capital utilization rate	7	-
Type of ownership: Public = 1	8	6
Type of ownership: Private profit = 1	9	8
Iron	10	-
Lead	--	2
Specific conductance	--	5

† The source of water and the region variables were included in all the steps during stepwise regression.

* Estimate includes only the systems that treated water.

** Estimate includes only the systems that treated water and had no major alterations since original construction.

The final form of the regression equations are presented in Table 8. The procedure used for obtaining these five equations are the same as those used for obtaining those in Table 6. Also, using a similar comparison among the equations indicates that the first and the fourth equations are superior to the others and the appropriateness of the one as opposed to the other of these two equations is largely a matter of the variable of concern.

The main difference between the results of Tables 6 and 8 concerns the relationship between the average cost and the system size. When the average cost is volume-based, there are clear indications of scale economies as demonstrated by the negative signs for the coefficients of the log of annual water volume variable in Table 6. However, as the dependent variable is changed to the number-of-connections-based average cost in Table 8, for equation 1 the size variable, log (number of connections), has coefficients with negative signs but with a t-ratio much less than unity, 0.27. Furthermore, for equation 4, the size variable is positively related with log (average cost) indicating scale diseconomies.

The use of the log of number of connections instead of the log of annual water volume as the independent variable representing system size has little to do with the differences in results. Although the number of connections and the annual water volume variables have relatively low correlation coefficients (0.25 for the data subset in equations 1 and 3, 0.17 for that in equations 4 and 5, and 0.60 for that in equation 2), the log of number of connections and the log of annual water volume have high correlation coefficients: 0.72 for the data subset in equations 1 and 3, 0.80 for that in equations 4 and 5, and 0.90 for that in equation 2.

Table 8. Multiple Regression Equations Explaining Log (Number-of-Connections-Based Average Cost of Production)

	Equations				
	No. 1	No. 2	No. 3	No. 4	No. 5
Degrees of freedom	368	46	455	51	35
Multiple corr. coefficient, R^2	0.21	0.50	0.12	0.55	0.26
Intercept	1.9687	2.8993	1.7097	1.8753	1.8717
Log (Number of connections)	-0.0140 (0.27)	-0.1831 (0.98)	0.0808 (1.85)	0.06694 (0.48)	0.1663 (0.96)
Log (No. of connections per pipeline mile)	-0.0792 (1.16)	-0.6083 (3.91)			-0.3411 (1.85)
Source of water: ground = 1	0.0113 (0.20)	-0.4031 (2.26)	-0.0090 (0.18)	0.2865 (2.08)	0.2375 (1.42)
Source of water: surface = 1	0.0036 (0.05)	0.0193 (0.07)	0.0069 (0.12)	0.2671 (1.70)	0.2043 (1.04)
Region of location: Northeast = 1	0.0570 (0.66)	0.1943 (0.63)	0.0185 (0.25)	0.0458 (0.23)	0.0676 (0.20)
Region of location: North Central = 1	0.0424 (0.50)	0.2104 (0.63)	-0.0159 (0.23)	-0.4419 (3.83)	-0.2010 (1.44)
Region of location: South = 1	0.2576 (3.43)	0.1421 (0.49)	0.2430 (3.81)		
Type of ownership: public = 1	-0.1532 (2.23)	0.4464 (2.30)	-0.0860 (1.63)	0.4944 (4.48)	0.3437 (1.77)
Type of ownership: private profit = 1	-0.2348 (2.85)		-0.1573 (2.25)	0.6710 (2.07)	0.2671 (0.72)
Log (total coliform + 0.5)	-0.0525 (2.40)			-0.0898 (1.34)	
Lead				-2.4749 (1.72)	
Nitrate-N	-0.0146 (1.75)			-0.0262 (1.44)	
Iron	-0.0100 (0.26)				
Manganese		-1.3799 (1.88)			
Color	-0.0177 (3.06)				
Cadmium		-11.3362 (1.08)			
Smell					-0.0940 (1.72)

Note: Numbers in parentheses are t-ratios.

As a result, the size variable used in Tables 6 and 8 represents both measures regardless of which one is used during the estimation.

The relative importance of the independent variables were examined for equations 1 and 4 by following the stepwise regression techniques similar to that used for analyses in the earlier subsection. The results are presented in Table 9. There are two interesting points that are observed from the differences in the rank of variables between Tables 7 and 9. First, the type of ownership stands out as a relatively more important variable for explaining the production cost per connection per year. Second, the relative importance of the size variable drops from first place for both equations 1 and 4 in Table 7 to last place for those equations in Table 9.

The latter observation may be used to explain the discrepancy in the results in Tables 6 and 8. Although the size variable for equations 1 and 4 in Table 8 indicated scale economies with a coefficient of low t-ratio or scale diseconomies, respectively, in both the cases the relative importance of the size variable dropped from first place in Table 7 to last place in Table 9. This indicates that the scale economies strongly observed in Table 6 do exist and that the results in Table 8 can be attributed to the nonavailability of information about the users with very different levels of water usage, that is, residential and nonresidential customers.

Because of these weaknesses in the regression analyses using the number-of-connections-based average cost and because of the confusion that might arise while using both sets of results, we have chosen to use only the results of analysis for the volume-based average cost for discussing policy implications of the results.

Table 9. Relative Importance of Independent Variables in the Multiple Regression Equations Explaining the Log (Number-of-Connections-Based Average Cost)†

Independent Variables	Equation 1* Rank	Equation 4** Rank
Log (number of connections)	7	6
Log (Total coliform + 0.5)	1	5
Color	4	3
Nitrate-N	6	-
Log (No. connections per pipeline mile)	2	-
Manganese	-	-
Capital utilization rate	-	-
Type of ownership: Public = 1	3	1
Type of ownership: Private profit = 1	5	4
Iron	8	-
Lead	-	2
Specific conductance	-	-

† The source of water and the region variables were included in all the steps during stepwise regression.

* Estimate includes only the systems that treated water.

** Estimate includes only the systems that treated water and had no major alterations since original construction.

SUMMARY OF RESULTS

The interrelationship between the average production and the various independent variables were examined using multiple regression techniques. Two measures of average cost were used: the volume-based and the number of connections-based. The results of the regression analysis indicated that among the factors influencing the volume-based average cost, the annual volume of water produced is relatively the most important followed by three finished water quality variables and customer density measured by the number of connections per mile of pipeline. It is also demonstrated that among the factors influencing the number-of-connections-based average cost, the total coliform, a finished water quality variable, is relatively the most important followed by the customer density variable. In all the equations tried, the source of water and the region variables were included during the stepwise regression so that they could serve as proxies for raw water quality. There were other variables, but they are relatively less important. While analyses using the number-of-connections based average cost have also been conducted only the volume-based average cost will be used for policy discussions.

V. RESULTS AND POLICY IMPLICATIONS

In this section, we will discuss the implications of the results of the regression analyses conducted in the earlier section. We will mainly be concerned with the implications regarding the choice of alternatives for a number of policy variables. We have chosen to restrict our discussions to three policy issues for which the regression equations are most appropriate for choosing alternatives. They are the following:

1. Cost-Quality Relationships: Will it be possible to add quality to the water produced without adding cost if small rural systems opt for regionalization of their existing facilities?[13]

2. Cost-Size-Density Relationships: Will it be feasible for the new rural systems to choose larger sizes and then meet EPA's quality standards without excessive production costs, or is this size option infeasible for rural areas because they typically have low customer density?[14]

---

[13]Regionalization of the existing facilities has been encouraged in the Safe Drinking Water Act. The Act provides for an extension of the time allowed for compliance among systems which enter into "an enforceable agreement to become a part of a regional public water system." See The U.S. Senate, Committee on Environment and Public Works, The Safe Drinking Water Act--As Amended Through November 1977, Serial No. 95-10 (Washington, D.C.: Government Printing Office, 1977), p. 14. In addition, EPA, the regulatory agency responsible for implementing the provisions of this Act, views regionalization more specifically in the context of small systems. The Agency considers it as one possible way for joining together for benefitting from the scale economies due to centralized treatment processes and then to meet the provisions of the Act. See U.S. Environmental Protection Agency, National Safe Drinking Water Strategy: One Step at a Time, Washington, D.C., February 1977, pp. 39-41.

[14]While accepting the scale economy effects in water systems, the literature has been concerned also about typical low customer density in large systems and consequently the scale diseconomies due to higher water transmission and distribution costs. See R. M. Clark, "Water Supply Regionalization: A Critical Evaluation," Journal of the Water Resources Planning and Management Division, ASCE, 105 (September 1979), pp. 279-294.

3. Cost-Ownership Relationships: Assuming the hypothesis that private systems will be better able to handle capital shortage problems is true also for the rural systems, will the private systems be able to compete with the public systems in efficiency of production?[15]

Cost-Quality Relationships

Can the existing water supply systems in the rural areas add quality to the water they produce without adding cost by taking advantage of the scale economies effect? Consider a hypothetical case of two water systems that are regionalized. As a result of this, the annual volume of water produced by the combined system will be higher than either of the two water systems. The customer density of the new regionalized system, however, may be smaller than either of the two original water systems. Nevertheless, the regionalization may have opened up an avenue for improving water quality further. Assume that the regionalized system has the same characteristics as the two systems regarding water supply source, location of region, ownership, capital utilization rate, and all the finished water quality variables except the total coliform. Then, starting with Equation 1 in Table 6, for sensitivity analysis, we can develop an equation that relates the percentage changes in the annual water produced, the total coliform and the customer density with the percentage change in the average cost. The new equation can be written as follows:

---

[15]For details on the hypothesis, see S. H. Hanke, "Crisis-Ridden Water Systems Should Go Private," <u>Wall Street Journal</u> (3 September 1981), p. 22.

$$\frac{\Delta(\text{Average Cost})}{\text{Average Cost}} = -0.2833\ \frac{\Delta(\text{Annual Water Volume})}{\text{Annual Water Volume}}$$

$$-0.1236\ \frac{\Delta(\text{No. of Connections per Mile})}{\text{No. of Connections per Mile}}$$

$$-0.0668\ \frac{\Delta(\text{Total Coliform} + 0.5)}{\text{Total Coliform} + 0.5}$$

We then assume that the percentage change in the average cost of producing water is zero. That is, the unit cost remains unchanged before and after regionalization. We can therefore write this equation as:

$$0.2833\ \frac{\Delta(\text{Annual Water Volume})}{\text{Annual Water Volume}} = -0.1236\ \frac{\Delta(\text{No. of Connections per Mile})}{\text{No. of Connections per Mile}}$$

$$-0.0668\ \frac{\Delta(\text{Total Coliform} + 0.5)}{\text{Total Coliform} + 0.5}$$

From the perspective of one of the two systems, if the water volume produced increases by 10% ex post-regionalization, we observe from this equation that there would be no margin left for improving water quality only if the customer density decreases by 23 percent due to regionalization. Assuming the customer density decreases by only 10 percent and the total coliform pre-regionalization is 1.5 bacteria/100 ml of water, it would be possible to bring the total coliform count down to EPA's maximum contaminant level of 1 bacterium/100 ml without incurring additional cost per unit of production.

Cost-Size-Density Relationships

Because of low customer density, is it infeasible in rural areas to improve water quality standards by building new water systems of large size. Again, look at Equation 1 of Table 6. Using that equation, we have presented a matrix in Table 10 showing the percentage change in unit cost for various hypothetical combinations of percentage increases in the annual water volume produced and of percentage decreases in the customer density. The percentage values in the cells indicate the percentage changes in the average cost due to the changes in the two axes. These values indicate that the size option may or may not be feasible depending on the customer density of the service area.

Table 10. Sensitivity of the Percentage Change in the Average Cost to the Percentage Changes in the System Size and the Customer Density

		Percentage Change In System Size				
		10	25	50	75	100
Percentage Change in Customer Density	-5	-2.2	-6.5	-13.5	-20.6	-27.7
	-10	-1.6	-5.8	-12.9	-20.0	-27.1
	-25	+0.3	-4.0	-11.1	-18.2	-25.2
	-50	+3.3	-0.9	- 8.0	-15.1	-22.1

Cost-Ownership Relationships

Finally, what is the relationship between the average cost and the type of system ownership. Equation 4 in Table 6 is appropriate because the cost calculations in this equation have already been standardized for differences in the cost-accounting procedures for systems under different types of ownership. According to that equation, using private nonprofit ownership as the reference group, publicly owned systems produce water at 11 percent higher cost, and private profit systems do so at 10 percent higher cost. This implies that firms operating as private nonprofit cooperatives or operated by a private individual can produce water at a cheaper cost. The firms run by private profit and public firms produce water at nearly the same cost. In other words, cooperatives may be a better form of organizational ownership where it is feasible. But in a choice between private profit and public firms, the choice is not affected by average production cost. The actual rate private profit systems charge to the users may, however, be higher because between the average production cost and the unit price charged, they would need a margin for paying returns to investors and for paying taxes.

VI. SUMMARY AND CONCLUSIONS

The initiating hypothesis for this study was that if a number of policy variables concerning the implementation characteristics of the water systems were maneuvered it would be possible to improve the quality of water produced by rural water systems without escalating the cost. An analytic framework was suggested and multiple regression techniques were employed to test the relationships indicated in this framework. Detailed data from a national level data base for 800 rural water systems were used. The results of the analyses indicated that, depending on the choice made from the alternatives within each of the three identified policy variables, it is possible to reduce the average cost of producing water. First, the average production cost of water decreased as the system size increased. Second, the cost decreased as the customer density increased, with density measured by the number of customers per mile of pipeline. Third, the systems under private nonprofit ownership can produce water at a low cost relative to those under private profit and public ownership.

The implication for policy centers on increasing system size. It is possible to develop a margin for improving the quality of the finished water without escalating its average production cost only if system size can be increased sufficiently to mitigate the effects of customer density. Additional margin may be developed by choosing private nonprofit ownership in the form of local cooperatives, but this option may be feasible only for the water systems in the lower end of the system size.

This study has remedied some omissions in the literature by developing a framework and by analyzing the interrelationships between production cost on the one hand and water quality and independent variables on the other hand. Time and resource constraints have made it possible to make only a modest effort in this direction. The economic modeling pursued here is a first step toward the development of models for practical use. Further efforts are needed before the results developed here can be used for management decision making. First, in this study, technology has been considered only implicitly. It would be necessary to deal with water production and especially with water treatment technology more explicitly by analyzing the cost and quality implications of various add-on treatment technologies available for small water systems. Second, it has not been possible to include various geological considerations because of the national scope of the data base and the large number of systems considered. For the management decision-making models, however, it would be essential that those factors be considered. If these research steps are undertaken, then the information developed in this study of rural water systems should prove useful for specific decision making related to water quality improvement.

REFERENCES

Bourcier, D.V. and R.H. Forste. Economic Analysis of Public Water Supply in the Piscataqua River Watershed. Durham, N.H.: Water Resources Center, University of New Hampshire, March 1967.

Clark, R.M. "Water Supply Regionalization: A Critical Evaluation," Journal of the Water Resources Planning and Management Division, American Society of Civil Engineers, 105 (September 1979), 279-294.

Daugherty, A.B. and J.D. Jansma. "Economies of Size Among Municipal Water Authorities in Pennsylvania," Southern Journal of Agricultural Economics, 5:2 (December 1973), 1-6.

Fox, W.F. Size Economies In Local Government Services: A Review. Rural Development Research Report No. 22. Washington, D.C.: Economics, Statistics and Cooperative Services, U.S. Department of Agriculture, August 1980.

Francis, J.D. et al. National Statistical Assessment of Rural Water Conditions. 5 vols. Report Prepared for the Office of Drinking Water, U.S. Environmental Protection Agency. Ithaca, N.Y.: Department of Rural Sociology, Cornell University, 1982.

Hanke, S.H. "Crisis-Ridden Water Systems Should Go Private," Wall Street Journal (3 September 1981), 22.

Hirshleifer, J., J.C. DeHaven and J.W Milliman. Water Supply: Economics, Technology and Policy. Chicago, IL: The University of Chicago Press, 1960.

Hunter, J.S. "The National System of Scientific Measurement," Science, 210 (November 1980), 869-874.

Johnson, R.B. and W.P. Hobgood, Jr. A Cost Analysis of Rural Water Systems in Louisiana. DAE Research Report No. 483. Baton Rouge, LA: Department of Agricultural Economics, Louisiana State University, May 1975.

U.S. Environmental Protection Agency. Managing Small Water Systems: A Cost Study. 2 vols. Cincinnati, OH: Municipal Environmental Research Laboratory, US/EPA, September 1979.

U.S. Environmental Protection Agency. National Safe Drinking Water Strategy: One Step at a Time. Washington, D.C., February 1977.

U.S. Senate, Committee on Environment and Public Works. The Safe Drinking Water Act--As Amended through November 1977. Serial No. 95- 10. Washington, D.C.: Government Printing Office, 1977.

U.S. Treasury Department, Bureau of Internal Revenues. Bulletin F: Income Tax Depreciation and Obsolescence. Washington, D.C.: Government Printing Office, 1948.

U.S. Treasury Department, Internal Revenue Service. Depreciation: Guidelines and Rules. Washington, D.C.: Government Printing Office, 1962.

ERRATUM

A.E. Res. 82-25, p. 25, Table 2, footnote:

Reads: "0: less than or equal to the maximum contaminant level allowable (primary) or recommended (secondary).
1: greater than the maximum contaminant level allowable (primary) or recommended (secondary)."

Should Read: "0: complete absence of the characteristic.
1: presence of the characteristic with varying frequencies of occurrence."