PENN LOH

SETTING URBAN WATER RATES FOR EFFICIENCY AND CONSERVATION

A DISCUSSION OF ISSUES

A Report Sponsored by the

CALIFORNIA URBAN WATER CONSERVATION COUNCIL

M.CUBED

OAKLAND, CALIFORNIA

JUNE 1994

SETTING URBAN WATER RATES
FOR
EFFICIENCY AND CONSERVATION

A DISCUSSION OF ISSUES

A Report For

THE CALIFORNIA URBAN WATER CONSERVATION COUNCIL

By

David L. Mitchell

And

Michael Hanemann

With Assistance From

Miranda Du

Final Report

JUNE 1994

M.CUBED ■ OAKLAND ■ CALIFORNIA ■ (510) 547- 4369

TABLE OF CONTENTS

PREFACE

THE CALIFORNIA URBAN WATER CONSERVATION COUNCIL URBAN RETAIL WATER RATES PROJECT

This report completes the first of a two phase project sponsored by the **California Urban Water Conservation Council** to develop information on innovative ways to improve how water is priced. The goal of the project's first phase is to provide an introduction to rate designs that are oriented more towards promoting efficient use of water resources than traditional designs have been; and to lend insight into the plethora of social and technical constraints that make ratemaking such a challenging exercise. It is intended for a broad audience. A familiarity with basic economic principles and water resource issues will more than suffice to derive benefit from the material presented in this report. If you do not have such a background but wish to develop one, this report might serve as a useful introduction.

The second phase of the Urban Retail Water Rates Project will develop a handbook providing practical assistance to utilities and their rate consultants implementing rate structures that promote more efficient use of water while at the same time recognizing and accounting for the many other functions a rate structure must fulfill. It will be a technical document designed for utility staff and professionals responsible for developing rate proposals. While a substantial amount of research on efficiency enhancing water rates has already been conducted over the last two decades, it is generally scattered throughout a broad literature, often oriented more towards theory than actual application. The purpose of the second phase of the Urban Retail Water Rates project is to take existing knowledge about developing rates, add to it new insights and experiences, and present it in a unified, well organized, and easily accessible way.

The Memorandum of Understanding Regarding Urban Water Conservation in California (MOU) signed by more than 300 water agencies, public interest groups, and other interested parties provided the primary impetus for this project. The MOU specifies 16 Best Management Practices (BMPs) for urban water conservation. BMP 11 states that member agencies will make good faith efforts to implement conservation pricing for water and/or sewer services. The Urban Retail Water Rates Project is designed to assist agencies with this effort. Even without the impetus of the MOU, there remains the need for this project. With demand for urban water service outpacing supply at an increasing rate, urban water agencies face a new reality where the provision of reliable, affordable service will depend as much on how they manage demand as on how they manage supply. One manifestation of this new reality has been the growing interest in finding new, innovative ways to price water service to encourage more efficient use and the adoption of cost-effective conservation measures. The Urban Retail Water Rates Project is an offshoot of this growing interest.

ACKNOWLEDGMENTS

This report is not the product of a single effort. The hours spent in writing, reviewing, commenting, and suggesting were shared by many. In particular, the authors extend special thanks to the following individuals and organizations: the members of the Rate Structure Subcommittee of the California Urban Water Conservation Council for their dedication to this project; Wendy Illingworth, Tom Chesnutt, John Christianson, Richard West, Tim Blair, Betsy Reifsnider, Polly Smith, Don Kuhl, Joe Young, Debra Braver, and John Farnkopf for their comments and insights; and the California Urban Water Conservation Council and California Urban Water Agencies for financial support.

CHAPTER 1 SOME INTRODUCTORY OBSERVATIONS ON WATER AND PRICE

Always make water when you can.

– Duke of Wellington

1.1 Introduction

Most of us -- even those that work in the water industry -- pay our bill for water each month without much thought. Very few of us understand the relationship between how much we pay for water and how much we use, or even how to interpret the bill we receive. The bill comes, we glance at the amount, we pay it. We are, it seems, careless consumers of water, treating it almost as though it were a free good. But how much we pay for water and how we pay for it have important consequences for the way we employ our economic and natural resources as a community, as a region, and even as a nation. Particularly for those of us living west of the 100th meridian, the price of water matters, even if we do not at first glance see why.

Utilities are now coming to realize that the price of water matters, not simply because it generates revenue needed to pay their bills, but also because it can shape the way water is used by their customers. In our economy, prices and efficiency are intimately connected. Prices that do not fully reflect the costs of goods and services will lead to inefficient levels of consumption. Prices that more fully reflect the costs of goods and services will lead to more efficient consumption. When resources are scarce (as they always seem to be), efficiency matters.

Water has not fully reflected the cost of its provision. This is true on both the social level -- the price of water historically has not reflected the social costs of environmental degradation associated with water development -- and the utility level -- the price of water historically has not reflected the opportunity costs of plant and equipment.[1] Many water utilities are now seeking ways to change this, to price their service in a way that more fully reflects its cost and promotes a more efficient use of water.

Change is never easy, of course. There are always good reasons for doing things the way they have always been done. There are always uncertainties associated with moving from a familiar to an unfamiliar practice. But change is necessary if we are to continue to improve the

---

[1] Opportunity cost is an economic concept we will introduce more formally later on. It refers to the value of a resource if it were employed in its highest yielding alternative use. If this alternative use can yield a greater economic benefit than the current use, it would be inefficient to maintain the resource in its current use.

way we do things, to seek better solutions. Change, however, should not be pursued recklessly. A change for the better requires an understanding of where we are as well as where we wish to go; it requires an understanding of the data before us; it requires careful planning.

This is a report about changing the way urban retail water service is priced. It endeavors to provide an understanding of the principles of rate design and an introduction to a variety of rate structures; it tries to present the facts (though we recognize that on this subject one person's facts often are another person's fiction); and throughout, it hopefully stresses the need for well laid plans. Specifically, this report discusses:

(1) ratemaking criteria for revenue collection, cost allocation, and customer incentive provision;

(2) conflicts that emerge between these criteria in practical ratemaking situations;

(3) the approach to ratemaking traditionally used by the water utility industry and its effectiveness in terms of revenue collection, cost allocation, and customer incentive provision;

(4) other approaches to ratemaking that may improve resource allocation and incentives to conserve water without jeopardizing revenue sufficiency or stability;

(5) legal constraints to rate design and cost recovery; and

(6) public involvement in the ratemaking process.

This discussion is organized into six chapters. Chapter 2 presents criteria commonly used for evaluating the efficacy of a particular rate design and introduces the practical considerations that emerge when applying them in practice. Chapter 3 reviews the traditional approach to ratemaking in the water utility industry in terms of the criteria from Chapter 2; it addresses the weaknesses of the traditional approach in terms of incentive provision and cost allocation, and considers some efficiency enhancing modifications. Chapter 4 explores other less traditional approaches to ratemaking, including marginal cost pricing, peak load pricing, increasing block rates, and priority pricing. Chapter 4 also considers the relationship between rate structure and utility finance, and how alternative rate structures might be used to enhance a utility's financial position during periods of cost or demand volatility. Chapter 5 presents a review of the legal and regulatory codes, as well as decisions from case law, that impact ratemaking in California. While this chapter relies on specific examples from California, the issues it addresses are likely to be relevant anywhere in the nation. Chapter 6 concludes the report with a discussion of the public's involvement in ratemaking.

To begin, however, the stage must be set for what is to come. Therefore, we devote the remaining two sections of this chapter to some introductory observations on water delivery, cost, and price. In sections 1.2 and 1.3, we consider two issues of central importance to water utility ratemaking. The first is the current rate of increase in water service costs and the consequences this may hold for urban retail agencies and their customers. The second is the relationship between how much people pay for water and how much they use. Supply represents only half of the market for water. Designing an efficient rate structure requires an understanding of demand and its associated characteristics, as well as an understanding of supply and its associated costs.

1.2 Trends in the Cost of Water Service

While the price of water historically has been relatively low, perhaps too low from a social point of view, available evidence suggests that this will not continue into the future (Hanemann 1992; Mann 1993). In fact, there good reasons to believe that urban water may become substantially more expensive in the near future. One reason is higher drinking water quality standards will require new investments in treatment facilities and push operating costs higher. According to a recent study by the National Regulatory Research Institute, the 1986 Clean Water Act amendments could add as much as $1.3 billion to the water supply industry's annual operating costs and another $2.0 billion to its annual capital costs (Beecher, Mann et al. 1990). Moreover, California urban water utilities will soon have to comply with the EPA's 1989 surface water treatment rule, which requires suppliers to filter all surface water. As one indication of how these cost increases could translate into higher water prices, Metropolitan Water District of Southern California (Metropolitan) has indicated that it may need to raise its surcharge for treated water by more than 150 percent between now and the end of the decade. In addition, increasingly stringent environmental regulations affecting the Delta and other California waterways will continue to push costs higher. Tougher water quality standards for the San Francisco Bay/Delta and endangered species recovery will effectively reduce the quantity and raise the cost of urban water supply throughout the state. Finally, many utilities are having to cope with burgeoning demands in their service areas caused by growth, which may require the development of new, more expensive sources of supply.

As the electricity and gas supply industries learned more than a decade ago, rapid escalation in costs and uncertain supply reliability pose serious challenges to the traditional methods of utility pricing. Unless ratemaking explicitly considers the demand side of the supply-demand relationship, the result is likely to be rates that are incompatible with realized demand and revenues that are insufficient to cover costs. Conversely, by factoring the demand-side into the ratemaking process, the utility gains a potentially powerful new tool for demand-side management at a time when supply shortages are an increasingly common occurrence.

1.3 Aspects of Urban Water Demand Behavior Important to Ratemaking

According to a United Nations report, the minimum amount of water a person needs for a "healthy living standard" is about 23 gallons per day (United Nations, 1976). Any water use beyond that level -- whether in support of people's livelihood or their lifestyles -- can arguably be termed discretionary, in that it reflects economic and social choices about industrial structure and patterns of urban living.

Gaining an understanding of these choices is one of the first steps a water utility has to take when developing a demand-side-management program. This section reviews the characteristics of demand typical to most urban water utilities in California. Most of the information presented is about residential demand. This is not out of neglect for the other customer classes, but rather because the majority of empirical research has focused on the residential sector.

1.3.1 General Patterns of Demand and System Loading

Urban demand for water exhibits distinct temporal patterns. The most important pattern to consider for ratemaking is the seasonal pattern of demand over the annual production cycle. Water use rises dramatically in the summer months, particularly in the arid west, when outside uses increase, and falls off in the winter months. Weekly and daily patterns also are prevalent to most systems. During the winter, demand typically falls off over the weekend as a result of business inactivity. During the summer, the opposite occurs, when weekends often register the highest demands as a result of lawn and garden watering. Daily demands often are bimodal, with a peak in the morning and again in the early evening.

These temporal patterns are fundamental occurrences that determine to a great extent the design and cost of urban water systems. The need for fire protection also is an important determinant of system design. Fire-flow requirements are important to the design of the distribution and storage network in the immediate vicinity of customers, while system-wide seasonal peaks in demand determine the capacity of most major storage facilities, transmission lines, pumping stations, and treatment plants.

Seasonal patterns of demand are relevant to ratemaking because, as we have just seen, they affect in important ways the size and cost of a utility's delivery system. To accurately reflect costs, rates need to account for this relationship. If they do not, they will be less efficient than they otherwise could be. Traditionally, utilities have recovered costs driven by seasonal peaks in demand with year-around charges that do not signal to customers that the cost of their consumption is higher in one period than in another. The traditional approach to recovering costs of peak usage will be examined in detail in Chapter 3. Peak period pricing presents an alternative to the traditional year-around cost recovery approach by recovering costs associated with peak period consumption only in that period. As is discussed more fully in Chapter 4, this approach signals to customers the true cost of providing peak service, and can avoid the problem of off-peak uses subsidizing the costs of peak period uses.[2]

One reason that the traditional approach to setting rates did not differentiate prices by season is that until recently many in the water industry did not believe that a significant relationship existed between the price of water and its use. The typical utility saw its demand mainly as a function of the size of the population it served. Rates were not considered to encourage or discourage water use; in that regard, they were considered to be neutral. More than two decades of research, however, has shown that price also matters.

1.3.2 Price Responsiveness of Demand for Water

That rates influence demand for water has been shown repeatedly by empirical research. A telling measure of this relationship is called the *price elasticity* of demand, which gauges the expected response in demand given a small change in price. It answers the question "If price were to change by x%, by what percent would demand change?". The water utility industry had for a long time assumed that the answer to this question was, for all practical purposes, zero. Empirical research, however, has not supported this assertion. In a recent review of the literature, for example, Dziegiewski, et al. (1991) summarize the empirical findings for residential demand as shown in Table 1-1.

---

2 For instance, with year-around capacity charges, customers living in densely populated urban centers wind up paying for part of the capacity required to meet seasonal demands driven by the outdoor usage of their suburban counterparts.

Table 1-1
Summary of Empirical Elasticity Estimates

Single-family Residential Customers	Range of Elast.
Winter season	-0.10 to -0.30
Summer season	-0.20 to -0.50
Multifamily Residential Customers	
Winter season	-0.00 to -0.15
Summer season	-0.05 to -0.20

Source: Dziegiewski, et al. (1991)

It is customary to present elasticity estimates in terms of the percentage change in demand given a 1% increase in price. Thus, the above summary can be interpreted in the following way: if the price for water in Anytown, USA, were to increase by 10%, we might expect demand by single-family residential customers to fall off by 1-3% during the winter, and by 2-5% during the summer. Similarly, we might expect demand by multifamily residential customers to decline by 0-1.5% in the winter, and by 0.5-2% in the summer.

There are several important features of this summary. First, demand is more elastic in the peak summer season than in the off-peak winter season. The most plausible reason for this difference is that outdoor use comprises a substantial portion of demand during the summer season but a much smaller portion during the winter season. Outdoor use tends to be much more discretionary compared to indoor use. As price changes, people are more able or more willing to adjust outdoor uses.

One of the best sources of information on the differences between indoor and outdoor water use is a study conducted some 25 years ago – the John Hopkins University Residential Water Use Research Project – for the Federal Housing Administration [Linaweaver et al., 1966; Howe and Linaweaver, 1967].[3] The study monitored residential water use in 41 study areas across the country, ranging in size from 34 to 2,373 dwelling units.[4] Both separate residences and multi-

---

[3] To our knowledge, this remains the only "real-time" residential water demand study in existence, where master meters were used to monitor daily and hourly water use across a broad spectrum of dwelling units and regions.

[4] Several of the study areas were in California. These were as follows:

<u>Metered public water and public sewer</u>

East Bay MUD - San Lorenzo, Creekside Acres, Burton Valley, and Chabot Park Estates

City of San Diego - Rancho Hills, Ruffin Rd. Residential, and Muirlands family complexes were included. Each area was master-metered with devices that recorded the accumulated flow every fifteen minutes for a period of two to three years. These flows were then aggregated into hourly and daily figures and expressed as averages per dwelling unit. Over the 41 study areas, average annual use per capita showed significant variation, ranging from 50 gallons per capita per day (gpcd) to 530 gpcd, with an overall average use of 115 gpcd. To isolate causes for this variation, water use estimates were broken down into estimates of indoor and outdoor residential use. This indicated that most of the *variation* in residential water usage was caused by outdoor use.

1.3.3 Indoor Residential Water Use

Indoor residential exhibited little variation across study areas, regardless of whether the dwelling unit was single or multi-family. Additionally, a comparison of metered and non-metered residences reveals very small differences in use. The study summarized its findings for indoor use as follows:

Table 1-2
Indoor Residential Use Measured by John Hopkins Study

	Low 10th percentile	Avg. Ann gpcd	High 90th percentile
Metered public water and public sewers			
Western areas	54	67	80
Eastern areas	41	51	61
Metered public water and septic tanks			
Eastern areas	44	47	51
Unmetered public water and public sewers	50	66	75
Apartment complexes	45	59	73

Source: Linaweaver et al., (1966)

These figures are probably still fairly accurate. The main factor affecting the growth of indoor residential use has been the adoption of domestic water-using appliances such as automatic clothes washers, dishwashers, and garbage disposal units, but these were already

---

California Water & Tel. Co. - Minot Avenue

Unmetered public water and public sewer

City of Sacramento - Golf Course Terrace

Apartments

City of San Diego - Apartment area

fairly common in the 1960's at the time of the Johns Hopkins study. According to data supplied by East Bay MUD, indoor residential use within its service area, which encompassed four of the study areas in the John Hopkins Study, is about 65 gpcd, just 2 gpcd less than that reported during the John Hopkins Study.

The above table suggests that indoor residential use is not very responsive to price. For example, a comparison of metered versus non-metered residences reveals little difference in indoor usage. Other research has also shown indoor use to be very price inelastic. For instance, Howe (1982) estimated that a 10% increase in price would reduce indoor residential demand by only 0.6%, while the data used by Morris and Jones (1980) indicated that demand might decline by 0.9%. Several studies have used winter usage as a proxy for indoor use. These estimates are typically higher than those of Howe or Morris and Jones since they include some outdoor usage, but still lie at the low end of the elasticity spectrum. To summarize, indoor residential use probably varies relatively little among regions or among different seasons of the year. Moreover, within the current range of price, it is relatively unresponsive to metering and other pricing policies.

1.3.4 Outdoor Residential Use

Outdoor residential demand, on the other hand, appears to be much more responsive to price. In the Johns Hopkins study, outdoor usage ranged from a low of 4 gpcd to a high of 768 gpcd. The following results were recorded for the study areas in the western part the United States:

Table 1-3
Outdoor Residential Use Measured by John Hopkins Study

	Low 10th percentile	Avg. Summer gpcd	High 90th percentile
Western Areas			
Metered public water and public sewers	67	102	135
Unmetered public water and public sewers	115	285	768
Apartment complexes	4	19	34

Source: Linaweaver et al., (1966)

While much of the variation in the above data can be explained by non-price factors, such as income, residential lot size (which is positively correlated with income), type of landscaping, climate (precipitation and temperature), and water use norms, the large differences between metered and unmetered residences suggests that price is also an important factor.[5] Unmetered summer usage in the western areas was, on average, 180 percent greater than metered usage -- an absolute difference of 183 gpcd. The John Hopkins Study also compared actual lawn sprinkling to agronomically derived lawn sprinkling requirements and found that the tendency for unmetered residences was to overwater while the tendency for metered residences was to underwater. On average, applied water by unmetered residences in the study was 238 percent of the watering requirement, while applied water by metered residences was 64 percent of the watering requirement.

Other empirical studies have also found outdoor use to be much more responsive to price than indoor use. For instance, Howe (1982) estimated that while a 10% increase in price would, on average, reduce indoor demand by less than 1%, outdoor demand would decline by more than 4%. Morris and Jones (1976), in a study of the Denver area, estimated that a 10% increase in price would reduce outdoor demand by more than 7%; Ben-Zvi (1980), studying demand in the Red River Basin, estimated that outdoor demand would fall by more than 8%.

These differences are important to a discussion of pricing because outdoor demand is a primary cause of peak demand during the summer season. And peak demand, as already discussed, is an important determinant of the size and cost of a utility's water storage and delivery system. If prices during the peak period do not fully reflect the cost of service during the peak period, an inefficient level of consumption will result -- a point to be emphasized throughout this report. Moreover, the higher the elasticity of demand during this period, the more inefficient the level of consumption will be. This same point can be turned around and put in more positive light. Because demand during the peak period is most elastic, the greatest gain in usage efficiency may be got from a peak period price that accurately reflects system costs.

1.3.5 Inelastic Demand

A second observation concerning the elasticity estimates presented in Table 1-1 is that, while not zero, neither are they very large -- a 10% increase in price results in less than a 5% decrease in demand. Economists distinguish between inelastic and elastic demands. An elastic demand has an elasticity that lies in the range between minus infinity and minus one. In this case,

---

[5] We refer the reader to appendix A in Dziegielewski, et al. (1991) for a review of empirical estimates of the relative importance of non-price factors as well as price in explaining the residential demand for water.

demand would fall by more than 10% given a 10% increase in price. An inelastic demand has an elasticity that falls within the range zero to minus one. This means demand would fall by less than 10% given a 10% increase in price. An elasticity close to zero implies that demand is not very responsive to price. In this case, only very large increases in price will result in significant reductions in demand.

From the viewpoint of economics, there are three interrelated reasons to expect an inelastic residential demand for water. One has to do with a household's ability to substitute other goods for water when its price rises. Water occupies a fairly unique position on the commodity spectrum in that there are really no close substitutes for it. To see the affect this has on price elasticity first consider the demand for a good that does have ready substitutes. Consider broccoli, for instance. If the price of broccoli rises a household can easily buy less broccoli and more of some other less expensive vegetable. The ability to substitute out of broccoli is virtually costless and depends almost solely on the household's preferences for vegetables. A household does not enjoy this relative ease of substitution when it comes to water. For a household to substitute away from water usually requires up-front investments in water saving devices or landscapes. Because these investments can require large initial sums of money and because the pay-back of the investment may not be obvious to households, they frequently meet with household resistance.

Another reason to expect an inelastic demand for water has to do the household's cost of monitoring its usage. Water use pervades daily household activities and it generally has not been a high priority to keep close tabs on use. Reinforcing this tendency is the fact that water use precedes payment, often by a significant amount of time. The disconnection between consumption and payment makes monitoring water usage seem less imperative. If, for instance, a utility could dole out water in the same fashion as a service station sells gasoline, demand would probably be much more responsive to price.

A final reason to expect an inelastic demand for water has to do with its low price. For the most part, a household's monthly expenditure for water service has not been very burdensome. This fact reinforces the previous two causes for an inelastic demand. If water service is inexpensive relative to the cost of substitution, for instance, then households will be reluctant to undertake costly investments to substitute away from water. If, however, the relative costs reverse, as occurred with energy in the 1970's then households will be more likely to make the necessary investments to substitute away from water.[6] Similarly, if the cost of water significantly increases relative to the "cost" of monitoring usage, a household may become more careful about how and when it uses water. Again, the experience with energy in the 1970's provides a ready analogy. A household consumes electricity in much the same way as water; its use pervades household activities, it is difficult to monitor, and consumption precedes payment. As the price for electricity increased throughout the 1970's, households began to be much more careful about their electricity consumption. Turning off lights and turning down the heat became routine habits for most households. The same possibility exists for water. As price rises above the historic range, demand may become more responsive to price.

1.3.6 Residential Demand - Short-run vs. Long-run

Another important aspect of the elasticity of demand for water is that its magnitude varies depending on whether it is viewed in the short-run or the long-run. One can think of water use as being determined by three factors: the stock of water using equipment, the efficiency of this equipment, and its utilization rate. Each of these factors may respond to a change in price, but at different rates. In the short-run, the stock of equipment is relatively fixed and its efficiency is, for the most part, also given. In the short-run, the most adjustable of the three components is the rate of utilization. Over the long-run, however, all three components can adjust. This added flexibility over the long-run usually results in long-run demand being more elastic than short-run demand. For example, Agthe, et al. (1986) estimated Tucson, Arizona's residential long-run demand to be almost twice the short-run. A study of residential demand in the Washington, D.C. area by Carver and Boland (1980) also found a large difference in long-run versus short-run demand. If consumer purchases of water using appliances, fixtures, and landscapes are based in part on the cost of water, as these results would suggest, then a price that does not reflect the full cost of service may encourage inefficient investments.

1.3.7 A Final Note About Elasticity

Before leaving elasticity, two cautionary notes about interpretation of elasticity are offered. The first concerns the degree to which elasticity estimates can be relied upon to predict demand adjustments in response to a change in price. Elasticity measures a change in demand given a change in price assuming *all other factors influencing demand are held constant*. In the real world,

---

[6]For instance, consider the massive shift into more fuel efficient automobiles following the rise in the real price of oil during the 1970s and early 1980s. Now that the real price of oil has fallen back to a relatively low level, consumers appear to be moving back towards larger, less fuel efficient cars.

of course, all other factors are not held constant. Variability in weather, income, and other factors will often be coincident with price changes. As a result, demand response to the price change can be very different from that suggested by an elasticity estimate, particularly if the price change is relatively modest. Martin (1988), for example, examining Tucson, Arizona's water consumption history between 1965 and 1987 found that in three out of the nine years in which price rose demand also rose; and that in five out of the eleven years in which price fell demand also fell. Non-price influences on demand can be controlled for in a demand forecast, but it is important not to naively interpret elasticity estimates as simple predictors of demand response to a price change.

The second cautionary note also has to do with the amount of relevant information contained in an elasticity estimate. Empirical studies, such as those cited above, most often provide point estimates of price elasticity. A point estimate measures the elasticity of demand at a particular point on the demand curve. That is, they characterize the behavior of demand only within the range of historically observed prices. As prices rise above or fall below this range, behavior of demand becomes much more uncertain. The further price ventures outside its historical range, the greater the error involved in any forecasted demand response. Indeed, many economists have suggested that OPEC's inability to accurately gauge the demand response to unprecedented increases in the real price of oil was an important factor leading to its demise as an effective cartel. It is certainly true that the degree of substitution and conservation that occurred following the sharp increases in the real price for energy during the 1970s surprised many rate analysts in the energy utility sector.

1.3.8 Elasticity -- Summary

Until recently, many in the water industry did not believe there to be a causal relationship between the demand for water and its price. More than 20 years of empirical research does not support this belief. Most research suggests that a 10% increase in price would, on average, result in a 2-5% decrease in demand. The relationship between price and demand is stronger for some types of water demand than others. Outdoor demand is typically more responsive to price than indoor demand; as is summer demand compared to winter. Some studies have indicated that demand for water is more responsive to price in the long-run than in the short-run. These results suggest that the price of water determines, in part, the choice of relatively long-lived water using appliances, fixtures, or landscapes. The behavior of demand if price were to rise outside its historical range is much more uncertain. While demand within the current range of price is fairly inelastic, this could change if the relative price of water were to increase substantially.

CHAPTER 2 A DISCUSSION OF PRINCIPLES FOR SETTING WATER RATES

I don't like principles ... I prefer prejudice

– Oscar Wilde

2.1 Introduction

There are many different rate structures that, in principle, could generate the same total revenue for a utility. The question, therefore, is which particular rate structure should a utility select? Moreover, since there are different groups of customers -- e.g., residential, commercial, industrial, etc. -- with different needs and placing different demands on its system, a utility also must address the questions of whether the same rate structure should be applied to all customer classes, or whether there is an economic rationale for applying different rates to different classes. In this chapter we consider some of the criteria for designing rate structures that have been advanced in the literature and offer some comments and explanation. We begin with a review of the basic components of a water rate structure. In section 2.3 we present in broad terms a set of criteria commonly used for evaluating the efficacy of a particular rate design. In sections 2.4 and 2.5 we prepare for the chapters to come by introducing the practical considerations that emerge when applying these criteria.

2.2 Components of a Water Rate Structure - Some Examples

Earlier this century, most urban water agencies were financed through fixed monthly charges. While these charges did not vary directly with consumption, they did sometimes vary according to customer characteristics, which, in the absence of metering, utilities used to identify probable consumption levels. In 1907, for example, Phoenix charged $1.50 per month for domestic water supply to dwelling houses; 25¢ per month each for water for horses and cattle; $1.00 for barber shops without fixtures, $1.50 for barber shops with one basin and one chair, and 50¢ for each additional chair; and rates of $1.00-$4.00 for stores and $2.50 and up for saloons. Since those days, water rates have changed, becoming simpler in some ways, more complex in others. In 1992, for example, Phoenix differentiated customers by meter size rather than by occupation. The rate structure involved a basic monthly service charge that varied by meter size and allowed usage up to 60 ccf per month in the winter (October - May) and 100 ccf per month in the summer (June - September), combined with a monthly volume charge for usage in excess of these amounts at a rate of 83¢ per ccf in the winter (December - March), $1.28

per ccf in the summer (June - September), and 99¢ per ccf in-between (April - May, October - November).

These examples are useful for illustrating the basic components common to most water rate structures. The first key distinction is between *flat charges* that are entirely independent of the quantity of water used (e.g., the monthly service charge of $1.50 per dwelling unit) versus charges that *vary* in some manner with the quantity used. Regarding the latter, one can further distinguish charges that vary *directly* with the quantity of water used (e.g., the volume charge of 83¢ per ccf) versus those that vary only *indirectly* or approximately with water use (e.g., a monthly service charge based on meter size or based on the number and type of fixtures). Thus, during the course of the century, Phoenix has switched from a system of fixed charges that were partly flat charges and partly based on factors that varied roughly, but imperfectly, with usage to a system with nothing that is a pure flat charge and a volume charge that varies directly with usage. Most other large urban water agencies have experienced a similar evolution in their approach to pricing water, and now have rate structures that combine various elements in a manner similar to Phoenix.

The Phoenix rate schedule illustrates two other increasingly common aspects of urban water rate structures. The first is that some elements vary seasonally, including the usage allowed as part of the fixed charge and the variable charge, which is higher in summer than winter. Indeed, the Phoenix rates contain two separate definitions of the seasons, one of which distinguishes the "shoulder" periods between the summer peak and winter off-peak periods. *Seasonally differentiated* charges like this have become more common in recent years -- Tucson was one of the first major urban water utilities to introduce them in 1975.

The second important distinction is that between variable charges that are *uniform-rate* versus *block-rate*. A uniform-rate variable charge is one where the amount paid per unit of consumption is the same over all units consumed; a block-rate is one where the unit charge varies, either decreasing with the amount consumed (*decreasing block-rate*) or increasing with the amount consumed (*increasing block-rate*). To the extent that the fixed monthly charge allows some usage for free, as in Phoenix, this implies a form of increasing-block rate, even if the remaining volume rate does not increase with additional usage. In the case of Phoenix during the summer, the cost per unit jumps from zero to the uniform-rate of $1.28 per ccf as usage rises above 100 ccf. Many utilities, however, are adopting a much stronger form of increasing-block rates that imposes an explicit premium for higher levels of usage. In 1993, for example, Los Angeles adopted a rate structure for single family residential accounts in which it charges $1.14 per ccf for usage up to 575 gallons per day in winter and $2.33 per ccf for usage above that amount. In the summer, the switch point changes to 725 gallons per day, and the rate for usage above this amount is $2.98 per ccf. The new rates complete a transition that had begun in 1977, when LADWP abandoned a declining-block rate structure in favor of a uniform-rate. Many other urban water agencies are in a similar transition: until 10-15 years ago, declining-block rates used to be common for large, non-residential accounts, if not for residential accounts, but are now giving way to uniform and increasing-block rate structures.

In addition, water agencies may levy other charges not yet mentioned. An example is a one-time charge levied when new customers are connected to the water system, known variously as a *connection charge*, *facilities charge*, or *capacity charge*. This is intended to recover capital expenditures for new facilities required to meet the projected demands of new customers. Furthermore, many water agencies have some special rates for particular classes of user. For example, there may be *life-line rates* which offer low-income customers some initial quantum of usage at a reduced price. It is also common practice for water agencies to offer service to some users outside their service area at special, higher rates; or to offer water for large irrigation users on an interruptible basis at specially reduced rates.

Having now seen the basic elements of a rate structure employed by urban water utilities, we return to the questions posed in the introduction to this chapter: How should these elements be combined in the design of a rate structure? What principles should be applied to judge a rate structure's merits? In the next section, we review some of the principles and criteria that have been advanced in the literature.

2.3 Criteria for Designing Water Rates

Generating income to meet a utility's revenue requirements is the primary role of a water rate structure. But whether intended or not, any system of water rates also performs two other functions: it allocates costs among users and it provides incentives to users. To evaluate the effectiveness of a given rate structure, we must consider how well it performs each of its functions.

(1) Collect Revenue - rates raise revenues that permit a utility to cover its costs on a self-sustaining basis;

(2) Allocate Costs - rates serve to allocate costs among different types of use and user;

(3) Provide Incentives - rates provide price signals to customers which may serve as incentives for them to modify their use of water, taking advantage of the favorable aspects of the rate structure and escaping the unfavorable ones.

An effective rate structure, therefore, should do a good job of collecting revenue, allocating costs, and providing incentives. What is meant in each of these cases by "doing a good job"? What are the criteria for success? Below, we identify and comment on the basic criteria that are generally accepted in the literature.

2.3.1 Criteria for Revenue Collection

With regard to revenue collection, a rate structure should be judged along at least three dimensions. Each represents a facet of the utility's ability to operate on a self-sustaining basis and meet its current and future financial obligations.

(1) revenue sufficiency – the rate structure should provide revenues sufficient to allow the utility to operate on a self-sustaining basis. The rate structure should collect revenue sufficient to cover operating costs, such as salaries, chemical supplies, gas and electricity, taxes, capital replacement costs, planning costs, and system expansion costs.

(2) revenue stability – the rate structure should provide a stable and predictable stream of revenue over time and as circumstances change (e.g., in a drought). Stable revenues allow for more accurate budgeting, better planning, and can lower long-term financing costs. Conversely, unstable revenues increase the risk of insufficient cash flow and can raise long-term financing costs. The extent to which unstable revenues also are uncertain (e.g., because of a drought) increases these risks.

(3) administration – administration and billing costs should be balanced against the potential benefits of a more complex rate structure. The rate structure should be designed to accommodate rate modifications and updates.

2.3.2 Criteria for Cost Allocation

With regard to cost allocation, the two principle considerations are equity and economic efficiency. Equity involves at least two considerations:

(4) Fair Allocation of Costs – The rate structure should apportion costs of service among the different uses and users in a manner that is fair and is not arbitrary.[7]

(5) Avoid Cross Subsidies – The rate structure should apportion costs of service in a manner that avoids the subsidy of one group of users at the expense of another.

In the context of cost allocation, economic efficiency requires the following:

(6) Fully Allocate Private and Social Costs– The cost allocation created by a rate structure should fully reflect the private and social costs of providing service for the different users and uses that the utility serves.

2.3.3 Criteria for Providing Incentives

With regard to the provision of incentives, the criteria reflect the goals of economic efficiency, system load management, and the promotion of conservation.

(7) Static Efficiency – The rate structure should encourage the efficient use of water in terms of quantity used and timing of use. To this end, at the margin, the water rate should reflect the full private and social marginal cost of supply.

(8) Dynamic Efficiency – The rate structure should encourage an efficient pattern of growth in water use and an efficient pattern of system development over time. In this regard, the marginal rate should reflect the long-run rather than the short-run marginal cost of water supply.

(9) Conservation – The rate structure should provide proper incentives for conservation, including investment by water users in cost-effective water saving appliances, fixtures, and landscaping.

(10) Rate Structure Transparency – The rate structure should be transparent and easy for water users to understand so that it provides a clear price signal to them. In order to be effective at influencing their investment decisions, it should offer a predictable price signal with minimum risk of unexpected adjustments.

---

[7] The most difficult problem in cost allocation is the dividing-up of common costs (also referred to as joint costs). Common costs are costs that are common to two or more services or functions of the utility. Simple rules of thumb to allocate common costs often result in an arbitrary assignment of costs to different customer classes. We discuss this issue in greater detail in Chapters 3 and 4.

2.4 Applying the Criteria in Practice

Applying these criteria to determine the best rate structure in some practical situation is not necessarily an easy task for at least three reasons. The first is that some of the criteria may directly conflict, forcing water agency managers to make trade-offs among them. For example, there can be a conflict between revenue stability and economic efficiency. Since a large part of a water utility's costs are fixed in the short-run, the safest way to ensure revenue stability is to raise revenues entirely through a fixed monthly service charge. This totally insulates revenues from fluctuations in the quantity of water delivered, but is counterproductive from the perspective of efficiency because it provides no incentive to use water sparingly. Conversely, some degree of revenue instability is inherent in any rate structure with volume charges. As the reader considers the above criteria, other potential conflicts will become apparent as well.

The second reason, which builds on the first, is that more than one party is involved in the rate setting process and, typically, different parties place different weights on the alternative criteria because of their own particular interests and point of view. Beecher et al. (1990) describe ratemaking as "a continual balancing act among the divergent and often competing perspectives of utilities, consumers, and society." They characterize these perspectives through a series of questions which we reproduce here, matching them up with the criteria presented in the preceding section.

<u>Utility's Perspective</u>

- Does the rate structure fully compensate the utility so that revenue requirements are met? (Criteria 1, 2)
- Does the rate structure allow the utility to earn a fair return on its investment? (Criteria 1, 2)
- Is the rate structure strategically sound for load management, competition, and long term planning? (Criteria 1, 2, 7, 8)

<u>Customer's Perspective</u>

- Are both the ratemaking process and the rate structure equitable? (Criteria 4, 5)
- Are utility rates perceived to be affordable? (Criteria 10)
- Are both the ratemaking process and the rate structure understandable? (Criteria 10)

Society's Perspective

- Does the rate structure promote economic efficiency? (Criteria 6, 7, 8)

- Does the rate structure promote the appropriate valuation and conservation of resources? (Criteria 6, 7, 8, 9)

- Does the ratemaking process take into account priority uses of water? (Criteria 9)

- Are both the ratemaking process and the rate structure just and reasonable? (Criteria 1, 4, 5)

Because of these differences in perspective, and because many parties have a stake in the rate setting process, one should recognize that rate design in practice is an inherently political process. In some cases, the law will proscribe what is or is not acceptable in a rate structure, while in others, it will be the public's voice. In all cases, however, a balance must be struck among the different criteria. In chapters 5 and 6 we will reexamine the rate structures to be introduced in the next two chapters in light of the legal and political aspects of ratemaking.

The third reason why applying the criteria can be difficult is the complicated nature of water utility costs. In many cases, the costs of service do not vary with the quantity of service in a simple or direct manner. The water industry is highly capital intensive, even compared to most other utility industries, and its capital equipment is lumpy; therefore, many of its expenses are fixed costs that hardly vary with the amount of water delivered. Also, the capital equipment is generally long-lived, so that the capital stock employed in any given water system tends to include assets of several different vintages, acquired at different times and at vastly different costs. Because capacity cannot readily be altered at short notice while demand may be highly variable -- both by day and by season -- a water utility always needs to build some amount of *excess* capacity into its system just to be sure of meeting demand with an acceptable level of reliability. Who is responsible for these costs and what price will a utility's customers be willing to pay for reliability? To complicate matters further, parts of a utility's system often serve several different functions at the same time, so that there is a problem of allocating *joint* costs among separate beneficiaries. For example, distribution mains are sized to meet average-day, peak-day, and peak-hour demand. Or consider a storage reservoir, which provides carry-over supplies from one season to another, benefiting peak users, while at the same time providing base supply to every user. How should these joint costs be allocated between average-day and peak-day services? How should those contributing to the peak demand be identified?

The result is that it can be both difficult and frustrating to answer what people might think is a simple question: "How much does it cost to supply me with the water that serves my needs?" On the one hand, it may require an enormous amount of data and extensive technical analysis. On the other hand, even when the analysis has been completed there still may be no single, correct way to develop a rate structure that satisfies everybody. The very nature of water utility costs tends to create conflicts between the criteria of raising adequate revenues, allocating costs fairly, and providing incentives for efficiency and conservation. Indeed, it is disagreement over how to resolve those cost-driven conflicts that accounts for most of the major controversies over utility ratemaking. This is particularly true of the dispute between proponents of the traditional approach to ratemaking, based on what is known as the principal of *embedded cost*, and proponents of the newer approaches that were first introduced into the electricity industry in the 1970's, then spread to other regulated industries such as natural gas and telecommunications, and are now being introduced into the water industry.[8] The traditional approach pools many different costs into a single average. It typically looks backwards to the costs that the utility has already incurred, and emphasizes the estimation and allocation of historical (i.e., embedded) average cost. In terms of the criteria presented above, it places the main emphasis on those that concern revenue sufficiency and stability, criteria 1 through 4. The newer approaches place primary emphasis on sending users the right price signal about the scarcity value of the commodity. They are generally forward looking; they focus more on marginal cost than on average cost, and replacement cost rather than historical, book cost, as the basis for pricing. In terms of the criteria, they place more weight on those that concern efficient use of a scarce resource, criteria 6 through 9.

In our experience, there is little appreciation among the public for the complexities of water utility costs and, therefore, little understanding of the judgments and compromises involved in the search for a good rate structure. In the next section, therefore, we offer a brief tutorial on the economic concepts of cost and the issues these raise for ratemaking.

2.5 Water Supply Cost Complexity and Ratemaking

We start off by imagining a fairy tale water utility that obtains its water supply from a magic spigot deep in a forest. The spigot is guarded by elves who also are responsible for installing and maintaining the distribution network that serves the utility's customers. For all these services, the elves charge the water company $0.05 per hundred cubic feet of water delivered to

---

8 Another common term for embedded cost pricing is fully distributed cost pricing.

the utility's customers. The staff of the utility consists of one retired elf, who deals with all the customers, sends them bills for water service, collects their payments, and handles all dealings with the elves who operate the magic spigot. Since he already has a pension, he performs these services for free.

In these circumstances, ratemaking is simple -- everybody would surely agree on $0.05/ccf as the appropriate retail charge for water. What makes the story a fairy tale is not the cheap price for water but rather the simple cost structure. The water utility only incurs costs as and when water is delivered. The costs vary directly with the quantity of water delivered. All units of water delivered cost the agency exactly the same amount of money.

The real world is different in every respect. As noted above, the water industry is highly capital intensive; most of its costs are capital costs, incurred when capital assets are installed, rather than operating costs, incurred as water is delivered. Once installed, the capital is not particularly malleable: the capacity of the system cannot be quickly expanded if you suddenly need more; nor can it be dismantled and disposed of profitably if you need less. Economists make a basic distinction between *fixed* versus *variable* costs -- i.e., costs that do not vary with the quantity of service provided versus those that do. Thus, most water utility costs are fixed costs.

There are, however, several complications to the distinction between fixed and variable costs. One arises from using different time frames for reference. We just indicated that, because capital is relatively unmalleable once installed, its cost is to be regarded as fixed. However, when viewed from a longer time perspective, all capital is variable -- existing capital will eventually wear out and, with sufficient time, new capital can always be added. This corresponds to economists' distinction between *long-run* and *short-run* -- in the short-run, the capital stock is fixed and its cost is a fixed cost; in the long-run, capital is variable and its cost constitutes a variable cost. How, then, should water be priced? Should its price reflect a short-run or long-run perspective? If prices reflect short-run costs, this may encourage patterns of use that are poorly suited to future circumstances when more expensive sources will be required. On the other hand, if prices reflect long-run costs, the utility may be in the position of charging for facilities before they exist.

Economists also make a basic distinction between *average cost* and *marginal cost*. Average cost is simply total cost divided by the quantity produced (e.g., if it cost $10 to produce 5 apples, then the average cost would be $2 per apple); marginal cost is the change in total cost of producing an additional unit (e.g., if it cost $8.50 to produce 4 apples and $10 to produce 5, then the marginal cost of producing the 5th apple is $1.50). The distinction between average and marginal cost applies in both the short-run and the long-run. Thus, there are short-run average and marginal costs, when capital is in place and its cost sunk, and long-run average and marginal costs, when capital is adjustable and all costs are variable. When all units of a commodity cost the same to produce, regardless of quantity, as in the fairy tale example, there is no difference between average and marginal cost -- the two costs are equal. Otherwise, however, there will be a difference between average and marginal cost, and it can matter greatly which type of cost is used as the basis for setting water rates.

Differences between average and marginal cost can arise for many reasons. A frequent example of divergent short-run average and marginal cost is that of a utility with a fixed capital stock and a constant unit operating cost. In the short-run the marginal cost is simply the unit operating cost -- the cost of producing an additional unit. The average cost, however, consists of the unit operating cost plus the cost of the fixed equipment averaged over the amount of production. The more capacity is used, the more product over which to spread the fixed cost and, therefore, the lower the average cost per unit. Hence, there is a declining short-run cost curve lying above a horizontal short-run marginal cost curve, as shown graphically in Figure 2-1. Historically, the notion illustrated here -- that with increased production one can spread fixed costs and therefore one has declining average costs -- underlies many of the arguments for declining block prices that were used prior to the 1970s. Often it was argued that large volume customers contributed disproportionately in thinning out these fixed costs, and therefore, should receive a lower rate.

A more difficult case for a utility is when cost is falling in the short-run but rising in the long-run, as illustrated in Figure 2-2. This phenomena has to do with the fact that new supplies may be considerably more expensive to develop than past supplies. While a utility may experience economies of scale for a single project -- similar to the situation shown in Figure 2-1 -- it may experience diseconomies of scale over a range of projects -- meaning that cost is increasing as new capacity is added to the system. Consider a utility developing a supply of water. It secures the most convenient and least costly source available and builds a reservoir. With the investment made and the reservoir built, the cost of the water depends on the level at which the reservoir is utilized. As use of capacity increases the average cost of supplying a unit of water falls because the fixed cost is being spread over more and more units. The short-run marginal cost of supplying a unit of water may also be falling because of scale economies associated with the project. When the capacity of the first reservoir is fully utilized, however, the utility will need to develop a new source of supply to meet new demand. It will likely have to develop a source not quite as conveniently located or of poorer quality than the first.

Figure 2-1
Short-run Cost of Supply for Single Project

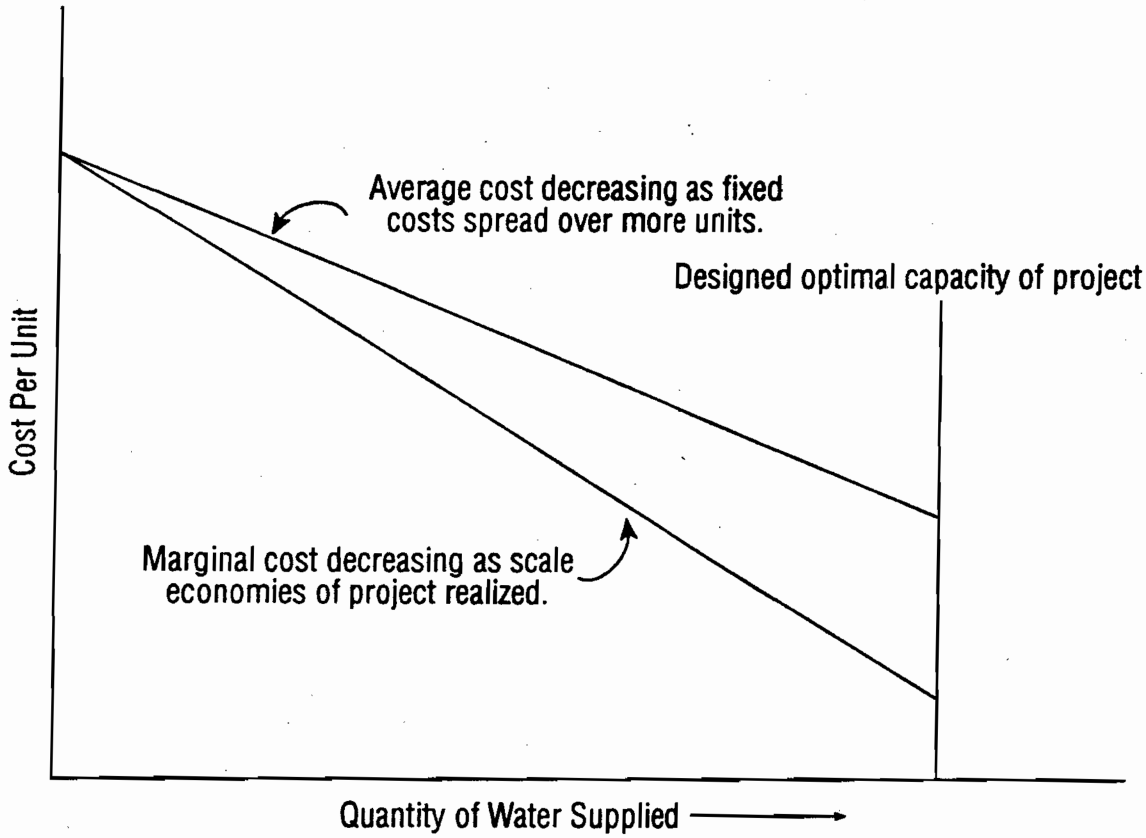
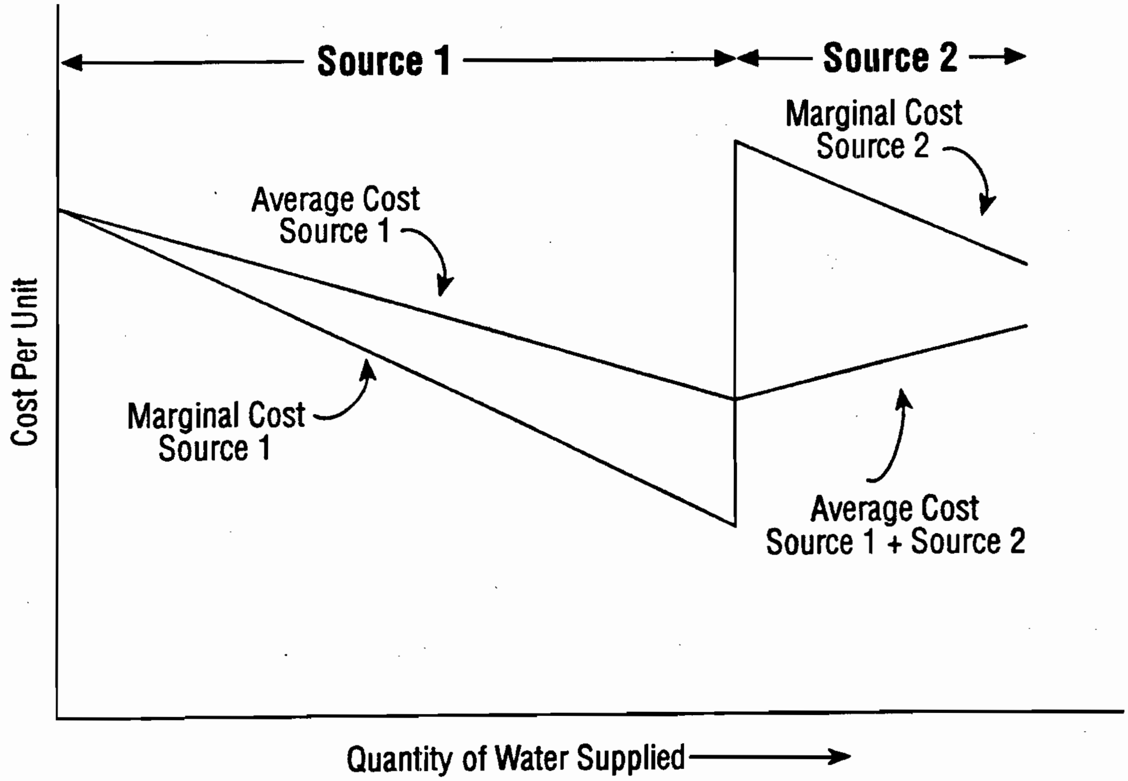


Figure 2-2
Change in Cost Structure
Following Addition of New Supply



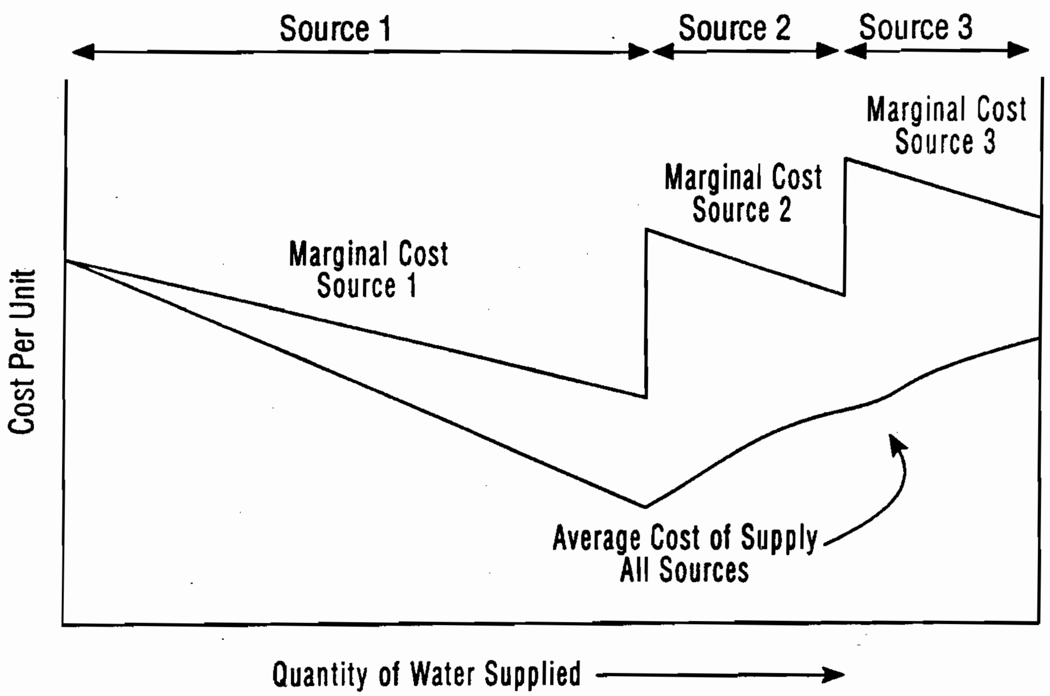
Suppose the cost of the second source is much higher than the first. How should the utility factor these two conflicting cost perspectives into its rate structure? Should rates reflect the lower unit cost of the first source, the higher unit cost of the second source, or some sort of average cost of both sources?

If we consider a utility with multiple sources of supply of varying vintages and costs, we get a situation like that shown in Figure 2-3. If the cost of water from these sources is arrayed from least to most expensive, as illustrated in the figure, we get an increasing long-run marginal cost of supply. In this case, the average cost would lie *below* the marginal cost; while rising, it is not rising as fast as marginal cost because it blends the high and the low costs together. Now we can better see the complexity involved in answering the simple question "How much does it cost to supply me with the water that serves my needs?" Can I lay claim to the cheap water corresponding to the first segment of the cost curve, and convince my fellow customers that *they* should be allocated the more expensive water corresponding to the other segments of the curve?

This is the fundamental dilemma of ratemaking when average and marginal costs diverge -- it is, in effect, an allocation exercise with many possible solutions. One approach is to average all costs over all users. Another is to charge all users the marginal cost for at least some of their usage. The first approach may be considered simpler, but it is not necessarily less arbitrary. Economic theory supports some version of marginal cost pricing on the principle that *all* users draw on the system at the margin and should be signaled the scarcity value of the resource.

Whenever average and marginal costs differ, if all units are priced at the average cost then total revenue automatically covers total cost. This is not true when all units are priced at the marginal cost -- total revenue falls short of marginal cost when marginal cost lies below average cost (Figure 2-1), but exceeds total cost when marginal cost lies above average cost (Figures 2-2 and 2-3). However, one should not leap to the conclusion that utilities *always* break even with average cost pricing or that they can *never* break even with marginal cost pricing. One reason is that we have focused so far only on variable charges while most rate structures in practice, as indicated at the beginning of this chapter, have several components, including fixed charges. The revenues from other charges can supplement the revenues from marginal cost pricing in the situation depicted in Figure 2-1. Conversely, when the situation is like Figures 2-2 or 2-3, where marginal cost pricing would more than break even, one can hold revenues down in other ways -- reducing some component of the fixed charge, for example, or charging less than marginal cost for part of a customer's usage (as with increasing block rates).

Figure 2-3
Cost Structure with
Multiple Sources of Supply



We have already mentioned the economic case for marginal cost pricing in terms of presenting an appropriate signal to users of the commodity. We should point out, however, that determining this cost is often very complicated. The complexity of a water system will require that the utility identify and measure not just one but several marginal costs. Some of a water utility's costs vary with the number or type of customers rather than the amount of water delivered. Other costs, such as for system capacity, depend on the *time* as well as the *amount* of water use.

Unfortunately, the complications do not end here. Many of the facilities for a utility provide several services simultaneously, and hence the cost for those facilities may be joint or common to those services. If a facility can be expanded to increase the level of output for one service without affecting the level of output of the other services connected to it, then its costs are said to be common. In this case, the share of the common cost for each service can be computed by calculating the incremental cost from increasing output for that service, holding the outputs of all other services constant. For instance, suppose a farmer stores three different crops in a barn. Then the common cost of the barn can be allocated to the three crops on the basis of how much of the barn's volume each crop requires.

If, on the other hand, the facility or production process can only provide services or outputs in fixed proportion to one another, then the costs of that facility or production process are said to be joint. For example, a distribution main jointly provides base capacity for off-peak service and extra capacity for peak service. To increase capacity by one unit for peak service necessarily increases capacity by one unit for off-peak service. In other words, producing one service implicitly results in the production of the other service, and their two costs are inextricably linked. As a result, the only economically definable cost for the distribution main -- average, marginal, or otherwise -- is for the composite peak/off-peak service it provides. This does not mean, however, that it is impossible to derive prices for the separate services that both are efficient and recover their joint costs of production.[9] They can be derived, but not from production costs alone. What makes the prices for the separate services determinate is the joint supply function, on the one hand, and the individual demands for each service, on the other (Kahn, 1970). It is the level of demand for each service that determines cost responsibility, and hence, the share of the joint cost applicable to each service. When one demand is disproportionately larger than the others, it will determine the level of production of the joint

[9] Competitive market forces do this routinely for most goods and services. The joint production of cotton seed and fiber, and their respective market prices, is an example.

products, and is causally responsible for the joint costs. If competitive forces were brought to bear, its price would reflect the full marginal cost of the joint facility while the prices of the other joint services would not reflect any of the joint cost (Kahn, 1970). Although a special case in general, this is more typical for water facilities that jointly provide peak and off-peak service, and is therefore important to ratemaking, as we will discuss further in the next two chapters.

As a matter of principle, one wants water rates to reflect costs. However, there usually is not one cost but many costs. Which cost, then, should the rates reflect, and how? While all costs are real enough, it is how we classify them that matters for ratemaking. Any classification is a matter of judgment and, to some degree, inevitably arbitrary. In the tale about the elves there was only one cost. Therefore, questions of classification did not arise and there could be total agreement on the correct rate structure. That is why it was a fairy tale. In the following chapters, we will move far away from this land of make believe in order to address in more detail rate structures for the real world.

CHAPTER 3 EMBEDDED COST RATE STRUCTURES

The formula 'Two and two makes five' is not without its attractions.

– Dostoevsky, *Notes from the Underground*

3.1 Introduction

Embedded-cost methods are commonly used by retail water agencies to develop rate schedules. The AWWA M1 Rates Manual, for examples, presents two embedded-cost methods for developing a rate schedule: the Demand/Commodity method and the Base Capacity/Extra Capacity method. Many state public utility commissions require that investor-owned utilities calculate rates using an embedded-cost framework.

Embedded-cost methods essentially provide ways to calculate the average unit cost of water service and then translate this cost into a rate schedule. The objective is to create a rate schedule that (1) accurately reflects the costs of service; (2) fairly allocates these costs to the utility's customers; and (3) does not collect too much or too little revenue, but just enough to cover the costs of running the utility.[10,11] As they are commonly applied, embedded-cost methods often do not meet these objectives. In this Chapter we show why, and suggest what can be done to improve their performance.

3.2 The Embedded-Cost Method

There are literally dozens of embedded-cost methods. While they differ in detail, they share a common framework. This framework includes four steps, as follows:

(1) Determine the utility's annual revenue requirement, which is the estimated annual expenditure that it will incur operating the water system. (In utility finance jargon, revenue requirement is synonymous with annual costs.) This would include operating and maintenance expenditures, such as for salaries, chemicals, energy, and equipment; depreciation expenditures for replacing worn or outmoded plant and equipment; capital financing expenditures for repaying outstanding debt and financing new projects; and, in the case of investor-owned utilities, a return on invested capital.

---

10 In the case of investor-owned utilities, this would include a fair rate of return on investment.

11 These objectives correspond to criteria 1 (revenue sufficiency) and 4 (fair allocation of costs).

(2) Allocate the revenue requirement to functional cost categories. The utility's costs estimated in step 1 need to be converted into costs that can be recovered either as a fixed charge or as a volume charge. Functional cost categories provide a way to do this.[12] For example, the Base Capacity/Extra Capacity method presented in the AWWA M1 Rates Manual defines four functional costs, as follows:

- Base Capacity Costs include the *variable* costs of water service plus a portion of the *fixed* costs of system capacity. This cost category attempts to measure the variable cost of water delivery plus the cost of capacity needed to meet average-day demand. These costs depend on the average amount of usage.

- Extra Capacity Costs consist of a portion of the fixed costs of system capacity. Extra capacity costs attempt to measure the cost of capacity needed to meet peak-hour and peak-day demands *in excess* of average-day demand. These costs depend on the amount and *timing* of water usage.

- Customer Costs are the costs of metering, billing, administration, and other customer related services. These costs depend on the number of customers rather than the amount of water usage.

- Fire Fighting Cost are the costs of installing and maintaining hydrants and other fire fighting fixtures.[13]

(3) Distribute functional costs to customer classes according to usage characteristics. These characteristics are used to assign each class a share of the functional costs. For example, suppose a utility using the Base Capacity/Extra Capacity method estimates in step 2 that its annual extra capacity cost is $1 million. This is the amount the utility estimates it costs each year to provide capacity to meet above average demand. The utility might then distribute this cost in proportion to each of its customer class's above average demand.

---

[12] Ideally, they will have some logical relationship to system costs, though this has not always proven to be the case.

[13] The Base Capacity/Extra Capacity method defines these cost narrowly. The cost of distribution line capacity required to meet peak fire fighting requirements are not included. Rather, they are subsumed in the base and extra capacity cost categories. Other embedded-cost methods might include capacity requirements for fire-flow as a separate category.

(4) Design the rate schedule to recover distributed functional costs. The goal of this step is to design a rate such that when multiplied by class usage the total revenue from the class approximates its share of functional costs, as calculated in step 3.

An embedded-cost approach to rate design has several advantages that makes it a popular choice for utilities. First, embedded-cost rates are relatively inexpensive to develop. They can usually be estimated with financial and customer usage data that the utility has on hand. Second, embedded-cost rates can provide a stable and sufficient revenue stream, so long as demand is stable. They are designed to recover annual costs assuming a given level of demand. Third, embedded-cost rates attempt to allocate costs fairly. They attempt to distribute costs to those that create them. Fourth, embedded-cost rates are meant to provide an efficient price signal to guide customer uses. They attempt to reflect the full cost of providing water service.

These advantages, however, are not always fully realized by utilities employing embedded-cost rate designs. As we show below, common accounting practices, unaccounted for costs, and rules of thumb to allocate costs can result in rate schedules that are less equitable and less efficient than they otherwise could be.

3.2.1 Accounting for Unaccounted Costs

While embedded-cost rates are meant to fully reflect the cost of service, they frequently do not. Moncur and Fok (1993) identify two causes of underpricing:[14]

(1) the use of actual historical costs rather than current replacement costs when computing capital and depreciation expenses;

(2) failure to include the economic cost of certain assets in the calculation of costs; and

It is typical for utilities to use book cost rather than market or replacement cost to value their assets for the purpose of calculating depreciation expenses and rates of return. Book cost refers to the value of the asset when it was purchased, i.e., its cost when it was entered in the utility's books. Most investor-owned utilities are required to use book cost rather than current market cost to value assets for rate calculations. For example, of the 50 state public utility commissions

[14] Other costs much more difficult to quantify also are excluded from embedded-cost rate calculations. These include the economic value of water itself, and environmental costs of developing delivering, and disposing of water. As a scarce resource, water itself has an economic value that should be included in rates if they are to be fully cost-based. The costs of environmental disruption associated with water development also should be reflected in rates.

in 1991, only Arizona's allowed its investor-owned utilities to cost their assets based on fair market value (National Regulatory Research Institute, 1992). Through time, a productive asset's book value steadily declines relative to its replacement value unless periodic adjustments for inflation and other factors are made. The longer-lived the asset, the greater the mismatch between historic cost and current market value is likely to be. As a result, utilities tend to underestimate depreciation and capital expenses, and rates do not reflect the true economic worth of the resources being consumed.

Utilities also do not include the cost of some assets and other expenses in their rate calculations. For example, state utility commissions frequently do not allow investor-owned utilities to recover through rates the cost of using retained earnings to finance a project. If a utility were to obtain outside financing for a project, the interest it would have to pay would be considered a legitimate expense that should be reflected in rates. If, however, it uses retained earnings, the interest it forgoes (i.e., the opportunity cost of not investing the retained earnings elsewhere in the economy) are not considered a legitimate expense that should be reflected in rates. Development charges that accumulate in "contribution-in-aid-of construction" (CIAC) or similar equity accounts also are not usually included as a cost of capital. In the case of investor-owned utilities, this is to prevent investors from earning a return on capital that they did not invest, which is reasonable. At the same time, however, by precluding this capital cost, rates do not fully reflect the economic cost of providing water service. Certain expenses also often are excluded. For example, in California, investor-owned utilities cannot recover through rates the cost of public education.

Rates that do not reflect the full cost of providing service will encourage inefficient consumption. If we accept that customers purchase water up to the point that the benefit of additional consumption equals the cost, then a below-cost rate will result in too much consumption; the benefit of the last units consumed is less than the cost of providing them.

Embedded-cost rates could be more cost-based (and more efficient) than they typically are if utilities included all productive assets in the rate calculation, and valued those assets according to their current economic worth. Moncur and Fok estimate that the use of market value rather than book value and the inclusion of all economic assets in embedded-cost rate calculations would increase rates for some cities by more than 50%. For the city of Madison, Wisconsin, they estimate that adjusting the book value of assets for inflation and adding the implicit cost of retained earnings would increase the 1991 commodity charge by 54%. Including CIAC assets in the calculations would increase rates an additional 17%. A similar analysis for

San Diego found that its rates would increase by 39%; rates for Denver would increase by 58%; and rates for Honolulu would increase 60%.

3.2.2 Allocating Joint Costs

Embedded-cost methods often rely on simple rules of thumb to categorize and distribute system capacity costs. These rules suffer from two weaknesses: (1) they do not properly allocate joint costs; and (2) they assign cost responsibility incorrectly. This can reduce the efficiency and equity of the rate schedule.

Joint (or common) costs occur when a single production process jointly produces two or more goods or services. For example, a cotton field jointly produces cotton fiber and cotton seed. Suppose one were asked to categorize each of the costs involved in growing cotton as either a cotton fiber cost or a cotton seed cost. While this can be done for some costs, it clearly cannot for others. For example, is land rental a cotton fiber or a cotton seed cost? What about the cost of tilling and planting, is it a fiber cost or a seed cost? These costs are joint to the production of both fiber and seed.

Many of a water utility's costs are joint costs. A neighborhood distribution network, for example, jointly provides fire protection as well as regular service. Peak summer demands are supplied by the same pumping and treatment plants that supply off-peak winter demands. If a transmission line is sized to meet peak summer demand but is also partially utilized during off-peak periods, how much of its cost should be paid by peak users and how much by off-peak users? Most embedded-cost methods rely on simple rules of thumb to allocate these costs. The advantage to this approach is that it is easy. The disadvantage is that it may result in an inequitable and inefficient rate schedule, as noted by Brown and Sibley in The *Theory of Public Utility Pricing*,

> "The distinguishing feature of [embedded-cost] pricing is that the allocation of common costs is done without much reference to what one would regard as economically meaningful criteria ... allocations have been done in literally dozens of ways in different regulatory proceedings."

This is not surprising, allocating joint capacity costs can be complicated and rules of thumb are often necessary to make the problem tractable. One the other hand, the risk that simple allocation rules will distort the efficiency or equity of a rate structure must also be kept in mind. For example, commonly used rules to assign peak and non-peak responsibility for capacity costs can detract from both the efficiency and equity of an embedded-cost rate schedule.

These rules typically allocate capacity costs according to each customer class's variation in demand. For example, the Base Capacity/Extra Capacity method in the AWWA M1 Rates manual assumes class's with high variations in demand are more responsible for the cost of system capacity than customers with low variations. The rational given for this is the following: a base amount of capacity is needed to meet average demand; extra capacity is then needed to meet above average demand; customers' average demands should pay for base capacity and their above average demands should pay for extra capacity; this can best be accomplished by allocating cost according to the variation in each class's demand.

While this logic is appealing, it will not necessarily result in a cost-based rate schedule because it does not correctly identify capacity cost causation. As discussed in Chapter 2, capacity costs are a function of the quantity *and* timing of customers' demand. A water utility scales its storage and delivery system in large part to meet peak period demand. As peak period demand increases, the extra capacity needed to satisfy it also increases. As Martin, et al. illustrate in *Saving Water in a Desert City*, the point to realize in allocating capacity costs is that during the peak period <u>any</u> water use contributes to the peak. Consequently, during the peak period <u>any</u> water use, not just water use in excess of a class's annual average demand, is partly responsible for capacity; during the off-peak, water use, even if its is above average for the class, is not driving the need for extra capacity (Martin, Ingram et al. 1984).

Commonly used embedded-cost rules of thumb to allocate capacity costs frequently neglect this link between cost and time of use. Two practices in particular can detract from the efficiency and equity of a rate schedule. The first is charging for extra-capacity year-around rather than only during the period that it is actually needed. Rates will be more efficient if they reflect the cost associated with both the quantity and timing of service. If the rate spreads the cost associated with time of use across all periods, rates will be too high during low cost periods and too low during high cost periods. Consumption will be distorted away from efficient levels and costs will be distributed inequitably between peak and off-peak demands.

Second, if capacity costs are distributed to customer class's using simple peak to average demand ratios, class demand peaks that are coincident with the system demand peak should be used. Customer class peaks in demand that are non-coincident to the system peak bear no relationship to the need for extra capacity and should not be used to assign cost responsibility.

3.2.3 Embedded-Cost Rates when Costs are Rising

Traditional embedded-cost rate designs perform best when demand and costs are stable and worst when they are unstable. In a traditional design, rates are set to recover the annual revenue requirement based on a given cost structure and pattern of demand. When costs increase, the rate then in force will begin to collect insufficient revenue. The utility will need to increase the rate to recover the higher revenue requirement. Traditionally, this was accomplished by increasing the rate in proportion to the increase in the revenue requirement. If costs increase 10%, then increase the rate 10%. This approach does not account for the link between price and demand, and will undercollect revenue when demand is downward sloping.[15] Unless the elasticity of demand is taken into account during an extended period of rate escalation, the utility will find itself perpetually trying to catch its revenues up with its costs. This can be avoided by incorporating the elasticity of demand into the rate calculation.

3.2.4 Short-run Perspective of an Embedded-Cost Rate Structure

In terms of efficiency, a principle disadvantage of an embedded-cost rate structure is that by averaging together the costs of the whole system, it does not signal the customer the full economic cost of developing new supply. They produce rates based on average cost rather than incremental cost. If costs are rising over the long-run, this may lead to uneconomic investments in facilities spurred on by excessive demand prompted by low rates (Hirshleifer 1958; Hirshleifer, Dehaven et al. 1960; Linaweaver and Geyer 1964; Howe and Linaweaver 1967; Goldstein 1986).

3.3 Embedded Cost Rates - Conclusion

Embedded-cost methods develop rates that price water according to the average cost of service. They are relatively straightforward to calculate, not data intensive compared to marginal cost approaches, and can provide a reasonable price signal to customers if based on the full cost of service. Certain accounting practices, state public utility commission

---

[15] For example, suppose a utility's total cost is \$100 [TC_0 = \$100]. Further suppose that it is currently selling 100 units [$Q_0 = 100$]. To recover its cost, it sets price to \$1 per unit [P_0Q_0 = \$1*100 = \$100]. The elasticity of demand is -0.4, meaning a 10% increase in price will decrease demand by 4%. Now, suppose total cost increases 15% [TC_1 = \$115]. Based on an expected demand of 100 units, the utility raises its price to \$1.15, expecting to collect \$115 in revenue [P_1*$E(Q_1) = E(R_1)$ = \$1.15*100 = \$115]. Actual demand falls short of expectations, however, because it is somewhat sensitive to the rise in price. Actual demand is 94 units rather than 100 [Q1 = Q0(1-0.4(P1-P0/P0)) = 100*0.94]; and actual revenues of \$108.1 are 6% below target.

requirements, and rules of thumb to allocate common costs, however, can result in rate schedules that are less efficient or equitable than they otherwise might be.

CHAPTER 4 OTHER APPROACHES TO SETTING WATER RATES

When the well runs dry, we know the worth of water.

– Benjamin Franklin

4.1 Introduction

With limited water supplies, rising costs, and heightened environmental awareness, a new emphasis on efficiency with respect to allocating existing supplies and developing new supplies (including non-traditional sources such as conservation) has emerged in the water industry. Hand-in-hand with this new emphasis is a growing interest in marginal cost pricing and other related pricing practices.

In this chapter we describe some of the alternative approaches to pricing utility services that first emerged in the US in the electricity industry during the 1970's and 1980's. What these approaches have in common is that they set aside the narrow focus on meeting the utility's revenue requirements, which underlies the embedded cost approach, in favor of a broader set of goals that also include economic efficiency and the promotion of conservation.

Pricing to encourage more efficient use of water rests on the assumption that prices can change consumers' behavior, even for a basic commodity like water. Whether this is so is an empirical matter that certainly can vary with circumstances. A major point we wish to emphasize throughout this chapter, however, is that *how* prices are used matters every bit as much as *whether* they are used. Prices can be effective or ineffective as tools for influencing behavior depending on how they are deployed.

The relative importance of price and other factors as influences on water use is frequently debated among water industry professionals. Our own view is that this is an unnecessary dichotomy -- price matters, but so do other activities such as public education, advertising, and information dissemination campaigns. Indeed, the two complement one another. Raising prices without providing guidance to customers on methods whereby they might reduce their water use blunts the effectiveness of the price signal. Conversely, exhortation and cajolery unaccompanied by financial incentives may have limited impact. All too often we have heard water managers complain that customers are resisting appeals for conservation because it is not in their financial interest. Low prices at the margin are an impediment to conservation – if there is little money to be saved, this undercuts the case for changing one's behavior.

Thus, we view pricing not as a substitute for a utility's existing or planned conservation programs but as something intended to work in tandem with them and enhance their impact.

The hallmark of newer approaches to ratemaking, as we view it, is that they attempt to target or tailor prices in such a way as to create effective incentives to use water more efficiently.

The rates discussed in this chapter owe their origins to two distinct sets of ideas. The first is the notion of marginal cost pricing as the requirement for an efficient functioning of the economy. This idea originated around the turn of the century with the great English economist Alfred Marshall in the context of his analysis of the economic problems created by economies of scale in industries such as the utility industry. It appeared again during the 1930's in the context of debates on socialist economics and the question of how prices should be set in public or nationalized industries. A formal, mathematical proof of the case for marginal cost pricing was provided by the American statistician and economist Harold Hotelling in a seminal paper on modern welfare economics published in 1938. It was put to practical use after World War II by French economists running the national power company Electricite de France, which adopted marginal cost as the basis for rate setting and investment planning in the mid-1950's. This was brought to the attention of Anglo-American economists in some papers published in American academic journals in the early 1960's and triggered an avalanche of interest. By the early 1970s, there was a strong tide of support for marginal cost pricing throughout the U.S. economics profession.

The other factor behind these rates was the environmental movement that was getting under way in the U.S. in the early 1970's. There was much concern at that time with the trend by U.S. utilities to build large new power plants, both nuclear and fossil-fuel, which were considered to have undesirable environmental impacts. There were also some economic arguments arising from concern about exaggerated projections of future demand and overbuilding of electricity generating capacity, all of which could be avoided by giving more emphasis to alternative sources of supply, such as cogeneration, conservation and what subsequently became known as demand side management.

Joining together, these two tides of opinion finally made their weight felt in the electricity industry following the 1974 energy crisis. First, utilities and their regulators began to explore and ultimately adopt pricing structures more consistent with marginal cost concepts. In 1975, for example, LADWP began a five year rate study to determine the effects of time of day and seasonal based rates, both offspring of marginal cost pricing theory. Second, the passage of the Public Utilities Regulatory Policy Act of 1978 (PURPA), intended to reduce or eliminate regulatory barriers to the development of markets for non-traditional sources of electricity, encouraged utilities to seek supply sources with the lowest social opportunity cost. Again, much of the motivation for this legislation came from economists and environmentalists invoking marginal cost arguments.

4.2. Economic Argument for Marginal Cost Pricing[16]

The economic argument for marginal cost-based rates comes from a branch of economics known as "welfare economics" which deals with prescriptions for efficiency in the use of scarce resources.[17] It offers rules for the efficient total production of commodities, the efficient allocation of this total among individual producers and consumers, and the efficient expansion of production capacity over time through new investment.

The central prescription of welfare economics is that commodities should be produced and allocated to the point where their marginal benefit equals their marginal cost. Whenever this is violated, there cannot be full efficiency in the economy with respect to either total level of production or the allocation of this total among individuals. Marginal benefit is defined as the extra benefit from producing and consuming one more unit of the commodity; as explained in Section 2.5, marginal cost is the extra cost from producing and consuming one more unit of the commodity.[18] They can equally well be defined as the benefit lost, and cost saved, by producing one *less* unit; seen this way, marginal cost is synonymous with *avoided cost* -- the cost that would be saved (avoided) by reducing output by a small amount. Seen the first way, it is synonymous with incremental cost -- the added cost of a small amount of additional output.

A related concept that is very important to an understanding of resource allocation is *opportunity cost*. With a given productive capacity, a decision to produce more of any *one* commodity, say water for urban use, implies a decision to produce less of all *other* commodities. Thus, the cost to society of producing anything consists, really, in the other things that must be

---

16 Our discussion in this section closely follows that in Kahn (1970, Chapter 3).

17 Welfare economics refers to the study of events, policies, and actions that affect the general welfare of an entire economy. It is not is not, as the name might imply to some, the study of specific welfare policies or programs to mitigate poverty.

18 Economists distinguish between *private* and *social* marginal benefit and cost. Social marginal benefit (cost) is the extra benefit (cost) accruing to society as a whole. private marginal benefit (cost) is the extra benefit (cost) accruing to the individuals directly involved. To the extent that there are what economists call *externalities*, there can be a divergence between private and social benefits and costs. The modern economic theory of pollution, for example, grows out of the divergence between private costs (which exclude the effects of pollution emissions on other people) and social costs (which include the effects of these emissions).

sacrificed to produce it. In the final analysis, all cost is opportunity cost -- i.e., the value of the alternatives foregone.

Given the above concepts, what is the economist's rationale for marginal cost pricing? At any given time, an economy has a fixed bundle of resources, a finite productive capacity.[19] In this zero sum game, a decision to produce more of any one good implies a decision to produce less of all other goods as a whole. Therefore, a basic economic problem for the society -- the efficiency problem -- is to use these resources to maximum advantage. For this to happen requires that the benefit gained from consuming one more unit of a good equal the cost to produce it. In other words, that the marginal benefit equal the marginal cost.

How do we then move from the fundamental requirement of marginal cost equal marginal benefit to marginal cost-based pricing? Prices provide signals that influence the behavior of producers and consumers. When consumers make their purchase decisions, they balance their benefit from a commodity against its price. If they are to make choices that reap the greatest possible benefit from society's limited resources, the prices that they pay must accurately reflect the opportunity costs associated with the commodities they are considering. If their judgments are correctly informed in this way, they will, by their independent decisions, guide scarce resources into those lines of production that maximize the net benefit to society *as a whole*.[20]

Could not prices equal, say, average rather than marginal costs (assuming the two are different) and do this just as well? If price had *no influence at all* on demand, it would not necessarily matter. But, in fact, the demand for all commodities is in some degree, at some point, responsive to price. As a practical matter, we know this to be the case for water, as we saw in Section 1.3. Then, if consumers are to decide wisely from society's point of view whether to take somewhat more or somewhat less of any particular item, the price they pay must reflect the cost of supplying somewhat more or somewhat less -- in short, the marginal opportunity cost. Suppose, instead, that buyers were charged more than the marginal cost for a particular commodity; they would then buy less than the socially optimum quantity -- welfare could be

---

[19] This finite productive capacity can grow or shrink with time; but for any given moment, it is, in a very real sense, fixed.

[20] It is important to emphasize the "as a whole" in this statement. Economic concepts of efficiency remain quiet with regards to distributional concerns. It is quite possible to have an efficient allocation that is grossly inequitable. The question here becomes should equity concerns be addressed independent of or simultaneously with efficiency concerns. Economists generally favor that the two problems be addressed independently, though this prescription by no means enjoys unanimous consent.

improved if they consumed more of the good in question and less of all other goods as a wnole. Some consumers who would have consumed more of the good in question, and less of other goods, will refrain from doing so because the price exaggerates the good's opportunity cost. Conversely, if price is set *below* marginal cost, they will buy more of the commodity (and less of all other commodities taken together) than is social optimal. Producers are diverting more resources to the production of this commodity than customers would have willingly authorized, had the price fully reflected the marginal opportunity cost.

4.2.1 The Choice of Time Frame for Marginal Cost Pricing

Although the basic economic argument for marginal cost pricing may seem compelling, we cannot end the discussion there and simply state that price should equal marginal cost. To do so is to ignore several important qualifications and practical limitations to this general rule. When it comes to moving from the theory of marginal cost pricing to its implementation, as with anything else, complications arise, and one is forced to deal with myriad details. In this and the next sections, we focus on two major issues: how one decides on the marginal cost concept to be used as a basis for setting prices, and how one ensures that the overall rate structure brings in sufficient revenue to meet the utility's requirements.

In Section 2.5 we presented a brief tutorial on the economic concepts of cost and pointed out the distinction between *short-run* and *long-run* marginal costs, as well as the related distinction between *fixed* and *variable* costs. Fixed costs, i.e., the costs associated with fixed inputs, are costs that do not vary with the quantity of service provided. Variable costs, i.e., the costs associated with variable inputs, are the costs that *do* vary with the quantity of service provided. By definition, the marginal cost associated with the use of fixed inputs is zero – the quantity of output changes, but there is *no* change in the quantity of these inputs, and so there is no change in this component of total cost. Only variable inputs generate non-zero marginal costs. In the short-run, the capital stock is fixed, and the marginal cost arises mainly from O&M costs. In the long-run, the capital stock can be replaced and expanded, and the marginal cost includes not only O&M costs but also capital costs.

By definition, then, short-run marginal cost is always less than long-run marginal cost. But, this gap is especially huge in the water industry because of its unusually high capital intensity.[21] Thus, it matters greatly in the water industry whether prices are set on the basis of short-run marginal cost, which is only a small fraction of total expenditure, or long-run marginal cost.

What is to be done? Different opinions have been expressed in the economics literature, and there clearly is no easy answer. All agree that price should *never* be set below short-run marginal cost. The question is whether it should be set any higher. The argument for short-run marginal cost pricing is best illustrated through the example of airline pricing. As Kahn has put it, no airplane should ever take off with empty seats as long as there exist some potential travelers who would be willing to pay the (almost negligible) short-run marginal cost associated with adding them to the flight roster. It is economically inefficient in this case to charge anything more than short-run marginal cost.[22]

This is not, however, a hard and fast rule. There are other considerations that may lead to a different conclusion. One consideration is price volatility. In some circumstances, short-run marginal cost could vary greatly over a short period of time as a result of fluctuations in either demand or supply (for example, in the case of a producer with access to many supply sources with very different costs). The variation in price that would result from short-run marginal cost pricing might be considered undesirable or counterproductive. Consequently, suppliers might prefer to abandon short-run marginal cost pricing in order to smooth out prices over time.

Another consideration is the impact setting price to short-run marginal cost may have on long-run investment. Whereas economic efficiency with respect to production from a *given* capital stock calls for setting prices equal to short-run marginal cost, economic efficiency with respect to investment and the determination of long-run capacity calls for setting prices equal to long-run marginal cost. If people are to reap the greatest possible benefit from society's limited

---

[21] For the water industry nationally, the asset requirement per dollar of revenue (i.e., the ratio of capital assets to annual revenue) is estimated at about $10-12; this is 3 to 4 times the capital intensity of the telephone and electric utility industries, and about 5 to 6 times that of the railroad industry.

[22] The airlines actually used this principle when they introduced cheap stand-by fares in 1966. But airline fares also reflect other economic considerations that are not necessarily socially efficient, such as price discrimination. The airlines reckon that the business demand for travel is far more inelastic than private individuals' demand for vacation travel. In order to discriminate between the two markets and maximize their profit from each, the airlines typically require staying over a Saturday night for their cheapest fares, since they figure this shuts out most business travelers. Then again, for stand-by-fares to be efficient requires the airlines be able to discriminate between different types of demand. Many airlines were compelled to discontinue stand-by-fares because they found that stand-by customers were making false reservations under assumed names to assure adequate seating.

resources and the capital stock is likely to be replaced or expanded in the foreseeable future, the prices of commodities must reflect the opportunity costs associated with not only the variable inputs but also the capital assets that are needed to produce them.

Suppose a producer will replace or expand his capital in the foreseeable future. Or, suppose that his customers will replace or expand *their* capital in the foreseeable future. In either case, charging short-run instead of long-run marginal cost provides an incorrect economic signal. If the wrong signal is given, this can lead both consumer and producer astray. If the price does not accurately reflect long-run marginal cost, the consumer may make investment decisions that are socially undesirable because they entail a long-run commitment to the use of a commodity for purposes that he would not consider justified if he had to pay the full long-run cost of producing it. Likewise, the producer may make investment decisions that are socially undesirable because they entail a long-run commitment to the supply of a commodity whose users would not find it worthwhile if they had to pay the full long-run marginal cost. This, in fact, is what happened on an unprecedented scale to the Washington Public Power Supply System -- whose acronym WPPSS is appropriately pronounced whoops -- resulting in the largest municipal bond default in U.S. history, which to this day continues to have serious financial repercussions for the Pacific Northwest.

Such questions about the correctness of investment decisions when prices are set below long-run marginal cost also have been raised throughout the history of the California water industry by critics of both urban and agricultural water policies. Gardner (1982) and others have long condemned the (wholesale) prices charged by the US Bureau of Reclamation for this very reason. From the beginning, the Central Valley Project has charged prices that were below long-run marginal cost -- indeed, for most of the last two decades it charged prices that were below short-run marginal cost. As a result, the critics argue, it continued to build new reservoirs and aqueducts that irrigators would not have considered worthwhile had they been required to bear the cost. These investment projects were politically viable only because their costs were effectively invisible -- they were averaged together with the costs of existing facilities and were spread over all beneficiaries of CVP water. But these phenomena -- pricing below long-run marginal cost and making investment decisions for new reservoirs that would not be justified if those who used the additional supply actually had to bear the cost -- are not unknown among urban water agencies, either.

As these considerations demonstrate, whether price should reflect short-run or long-run marginal cost will vary with the circumstances. As a general principle, however, we are inclined to follow Kahn's recommendation:

> The practically achievable benchmark for efficient pricing is more likely to be a type of average long-run incremental cost computed for a large, expected incremental block of sales, instead of short-run marginal cost estimated for a single additional sale. This long-run incremental cost (which we shall loosely refer to as long-run marginal cost) would be based on (1) the average incremental variable costs of those added sales and (2) estimated additional capital costs per unit for the additional capacity that will have to be constructed if sales at that price are expected to continue over time or to grow. Both of these components would be estimated as averages over some period of years into the future. (Kahn (1970) p. 85)[23]

Note that when Kahn speaks of average cost, he is referring the average incremental cost from adding a large block of new supply, <u>not</u> the average cost of supply from both old and new sources, which is the embedded-cost approach discussed in the previous chapter. When a utility prices its water service as Kahn recommends, customers are sent a signal that reflects the opportunity cost associated with their consumption level.

4.2.2 Meeting Revenue Requirements with Marginal Cost Pricing

In Section 2.5 we noted that, when prices are set equal to average cost, total revenue by definition covers total cost exactly. If prices are set any other way, such as marginal cost pricing, this does not hold. But, we also noted two qualifications. One is where average cost per unit stays constant as output changes; then average and marginal cost coincide, so that marginal cost pricing is the same as average cost pricing. The more important qualification is that there are *other charges*, and even *other ways of configuring marginal cost pricing,* through which utilities can ensure that their total revenue is adequate but not excessive to cover their total cost. Elaborating on these methods for meeting revenue targets is the focus of this section.

It has long been recognized that industries with decreasing average costs would not cover their total cost if they set prices equal to marginal cost. Two of the three main solutions have been known for over fifty years. One, involving what became known as the *two part tariff,* was suggested by Lewis (1941). This is where a utility combines a commodity charge based on

[23] This recommendation serves as a general principle rather than a hard and fast rule. It is easy to imagine situations where pricing at short-run marginal cost would be most appropriate. For example, if a wholesale agency had extra water available in its reservoirs that it either had to sell or release to make storage available for next year's run-off, the appropriate price would be the short-run cost of delivering the water to retail water agencies.

marginal cost with a fixed charge, e.g., a service charge or connection charge. Together, these make up the two parts of the tariff. The idea is that the fixed charge raises the additional revenue needed to cover total cost, but does not interfere with economic efficiency generated by having a commodity charge based on marginal cost. Since the customer pays the same fixed charge regardless of what quantity he consumes, only the commodity charge should influence his quantity decision, and this still provides the economically correct price signal. Hence, it is argued, the two part tariff can satisfy the goals of both economic efficiency and revenue adequacy.

Subsequently, the efficiency of the two part tariff has been challenged on two grounds. One is the argument that, even though consumers *ought* to disregard the fixed charge when they decide how much of the commodity to use, perhaps they do not. Perhaps they are impressed by the fact that, if they increase their consumption, they can spread this fixed charge over more of the commodity and so reduce the overall unit cost. If so, this would create an economically inefficient incentive to maximize consumption. When Mayor Bradley's Blue Ribbon Committee was meeting to consider LADWP water rates in 1992, it was persuaded by this argument and recommended eliminating fixed charges from the water rates. It is not clear, however, that the argument is valid.[24]

The second efficiency argument is given more weight by most economists. It arises from the fact that, while the fixed charge does not vary with the quantity consumed, it *does* vary with whether or not one consumes *anything at all.* If a consumer were willing to opt out of using the commodity entirely –e.g. disconnecting from the utility's distribution network – then he could avoid the fixed charge. A consumer in this position should be influenced by the fixed charge in making an economic decision. In this case, it can be shown that the fixed charge may be economically inefficient -- there may be situations where the social optimum requires that the consumer *not* exit the system, but his private incentive in the face of the fixed charge *does* lead him to exit the system. In those situations, social efficiency would prescribe a smaller fixed charge than that needed to fill the gap between total costs and the revenues raised through marginal cost-based commodity charges. If so, there is *no* solution that both generates adequate revenues and is perfectly efficient from an economic point of view; there are only what economists call *second-best efficient* solutions.

---

[24] Another reason for having some fixed charge is to cover costs that vary by customer rather than by quantity of commodity sold, such as the costs of administering customers' accounts. Economic efficiency suggests these be recovered not by a commodity charge but by a fixed service charge set equal to the marginal cost per customer served.

Before continuing, we should emphasize that all the foregoing discussion is predicated on there being decreasing long-run average costs. In the urban water industry today, the reality is more likely to be *increasing* long-run average costs, in which case long-run marginal cost pricing generates total revenues *in excess* of total costs. There is still an efficiency issue -- the utility would need to dispose of the excess revenues in a way that doesn't bias customers' decisions away from efficiency -- but it may be much easier to resolve.[25] For example, the utility could use some type of rebate that was independent of customer use. Furthermore, the efficiency problems that arise from having a fixed charge (or rebate) would vanish if one could be sure that the charge *did not* influence customers' exit/entry decisions. For water, more than for any other utility service, this is likely to be the case. The fixed payment that most urban water agencies need to charge are unlikely to drive consumers off the distribution system and, therefore, are unlikely to cause a two-part tariff to be economically inefficient.[26]

The second main solution to the revenue problem of marginal cost pricing in the presence of non-constant average costs is known as *Ramsey pricing*, named for an English economist who died at a tragically young age in 1928 after making two seminal contributions to mathematical economics that took almost 50 years to become fully appreciated. Ramsey (1927) modified the conventional economic efficiency analysis by adding an explicit constraint that commodity prices not only maximize social welfare but also break even.[27] He derived a complex formula for how one should adjust prices away from marginal cost *in inverse proportion to the elasticity of demand*. The intuition underlying Ramsey's formula is that, in order to preserve as much efficiency as possible, one wants to depart as little as possible from the pattern of consumption that would occur with unfettered marginal cost pricing, while still charging prices that secure the utility sufficient but not excessive revenue. One accomplishes this goal by imposing the *largest* price adjustments on the commodities whose quantity demanded is *least* sensitive to price, and the *smallest* adjustments on the commodities whose demand is *most* price-sensitive.

---

25 There also are legal issues, which we address in Chapter 5, associated with collecting too much revenue.

26 This is *not* the case with some other industries, such as cable TV for residential customers or long-distance telephone for large commercial and industrial customers. Those industries are more likely than water to have decreasing average costs, and they face a real prospect that their fixed charges could affect customer's entry/exit decisions. AT&T, in particular, has actively sponsored research since the late 1970s to develop a more efficient alternative to the two-part tariff known as *nonuniform* or *nonlinear pricing* [see Brown and Sibley (1986, Chaps 4, 5), Mitchell and Vogelsang (1991, Chap 5), and Wilson (1993)].

27 Because of the imposition of this constraint, his analysis was an exercise -- indeed, the *first* exercise -- in second-best efficiency.

The result is a form of cross-subsidization that yields a more efficient economy than if one had simply adjusted the price for all customers in the same way. There are two obvious problems with this approach, however. One is that the inverse elasticity pricing formula can be extremely complex and usually will require information on demand that simply is not available to most utilities. In order for Ramsey pricing to work, one must have detailed information not just for the demand of the commodity in question, but for all potential complementary and substitute goods as well. The other is that this particular solution to the revenue requirement problem relies on cross-subsidization, which, as we recall from Chapter 2, is in direct conflict with the principle of equity, and is in conflict with most regulatory codes governing the pricing of water.

There are other, *ad hoc* solutions to the revenue problem created by divergent marginal and average costs. For example, Brown and Sibley (1986) note that in Europe it is sometimes accepted that public utilities should price on the basis of marginal cost with the government making up revenue shortfalls out of tax revenues. In the U.S., with its preponderance of investor-owned utilities and its emphasis on PUC regulation, the tradition has generally been that utilities cover their own costs without taxpayer assistance in the case of decreasing long-run marginal cost, or that they do not earn excess revenues in the case of increasing long-run marginal cost. This has generally lead to average cost pricing rules in the former case and some form of proportionally scaled-down marginal cost-based pricing in the latter.[28] Either of these approaches distort the price signal to the consumer away from the true long-run marginal cost.

Another solution is to have increasing or decreasing block rates. These can be regarded as multi-part extensions of the two-part tariff. They will be discussed further in Section 4.4; we mention them here only to point out their implications for revenue adequacy. Suppose there are decreasing long-run average costs. If there is a form of declining block rates with the block where most consumption is located being priced at long-run marginal cost, while the earlier blocks (called the *inframarginal blocks*) are priced at some higher amount, this can provide many of the efficiency benefits of marginal cost pricing while still breaking even. Conversely, if there is increasing long-run marginal costs (which we consider more likely to be the case for the urban water industry) one wants a form of increasing block rates with the inframarginal blocks priced below long-run marginal cost in order that total revenues match total costs.

---

[28] The latter approach avoids the undesirable cross-subsidization implicit in Ramsey pricing by applying a uniform upward or downward scaling to marginal cost-based prices so that revenue equals total cost.

4.2.3 Marginal Cost Pricing -- Summary

Rate structures based on marginal cost pricing precepts are intended to provide price signals that will result in a more efficient allocation and use of a scarce supply of water. Efficient consumption requires that the benefit derived from consuming one more unit of a good equal the cost of supplying it. If the benefit is less than the cost, it is inefficient to produce additional units; the resources require to do so could be more productively employed elsewhere in the economy. By setting price equal to marginal cost, consumers are able to compare the benefit of additional consumption with its associated cost and make efficient choices. In this way, production may be guided towards more efficient levels.

Marginal cost pricing differs from average cost pricing in important ways. Whereas, marginal cost pricing reflects the cost of producing an additional unit, average cost pricing reflects the unit cost of producing all units. If costs are decreasing as output increases, average cost will be above marginal cost. If costs are increasing as output increases, average cost will be below marginal cost. In either case, prices based on average cost will not result in efficient consumption choices because consumers will not be matching marginal benefit with marginal cost.

Marginal cost prices are more difficult to calculate than average cost prices. To calculate average cost prices, a utility needs some understanding of its total costs and production. Dividing one by the other yields an average price. To calculate marginal cost, on the other hand, a utility needs fairly detailed information on costs for a variety of plant and equipment, some of which may not yet be built. This data may be expensive and difficult to obtain and of uncertain reliability.

Marginal cost pricing can result in over or under collection of revenue. Collection of revenue insufficient to cover cost obviously is not sustainable in the long-run. Collection of too much revenue may also be a problem if a utility is constrained to earn zero profit. Various strategies to satisfy the zero profit constraint while retaining the efficiency properties of marginal cost pricing have been proposed. Two-part tariffs and multiple block rates, for example, can be designed so that price at the margin reflects marginal cost, while prices for infra-marginal consumption are set so that the utility breaks even. Another strategy, Ramsey pricing, is to set different prices for different customers or customer groupings according to the magnitude of their elasticities of demand.

4.3 Seasonal Rates

The simple formulation of marginal cost pricing requires that all customers purchasing a good should be charged the same price, namely the good's marginal cost. But suppose -- as we suggested in Chapter 2 -- that it costs more to provide the good during certain periods than others. Suppose, for instance, that a water agency does not have one but several sources of supply, and that its most expensive sources are required only when demand is high, whereas its cheaper sources are sufficient when demand is low. How then should the utility price its water? Under marginal cost pricing, demand should be equated to cost at the margin. But since we have two different marginal costs, we get two different prices at.different times. This can be considered an example of peak-load pricing: the service is priced higher during the peak demand period, when cost is high, than during the off-peak period, when cost is low. For some utilities the major variation in demand occurs during different parts of the day (e.g., early evening, middle of the day, night) or during different days of the week (e.g. weekday versus weekend). For water, while there are daily and weekly cycles, the main variation is seasonal -- summer versus winter -- because outdoor water use varies with seasonal changes in temperature, precipitation, and plant evapotranspiration. For this reason, the terms seasonal pricing and seasonal rates are commonly used in the water industry when discussing peak-load pricing. In this section, we discuss both the theory and the application of peak-load pricing for water service in terms of economic efficiency, cost allocation, and incentives for conservation.

4.3.1 Principles of Peak-load Pricing

It is safe to say that the economics profession is fairly unanimous in its prescriptions for pricing capacity required to meet peaks in demand. In general, there are two. Both follow from marginal cost theory, and both assign capacity costs to those that are causally responsible for them. One is appropriate in instances where variation in demand is random, and therefore unpredictable; the other is appropriate when the variation in demand is systematic, and therefore predictable, at least to a reasonable degree of accuracy. As it turns out, both situations occur with urban water service. Fire service is perhaps the best example of an unpredictable demand variation that requires that excess capacity be on hand. Summer sprinkling is an example of a much more predictable demand variation that also requires that excess capacity be on hand. Before discussing in more detail each pricing prescription, however, additional insight into the nature of the service provided by a water utility may be useful.

A fundamental condition of urban water service is that the company must be ready to deliver at any time. In this sense, as Hirshleifer, et al., aptly state, the utility must stand "ready to enter into a contract for delivery at the option of the buyer." It is the customer rather than the utility that is in the proverbial driver's seat determining the quantity of service the utility must provide at each given moment. To oblige the demands of its customers with a given level of reliability, the utility requires an excess capacity above the *average* level of demand it can expect. All urban water systems are designed with this in mind.

Essentially, then, a customer holds an option to utilize this reserve capacity, and can exercise this option at any time. The cost of the option to the utility is the cost of providing additional capacity. Efficiency requires that the customer's willingness to pay for the option equal the cost of the option. In other words, the customer gaining access to additional capacity must be willing to pay the incremental cost of providing it. If the price for additional capacity is less than its marginal cost, too much capacity will be demanded. If customers were to face the true cost of the additional capacity, they would scale back their demand and apply the savings to more valued alternatives. Conversely, if the price is above its marginal cost, too little will be demanded. In this case, if customers were to face the true cost of additional capacity, they would scale back their demands for other goods and services and apply the savings to the purchase of more capacity.

With this in mind, we can state the economist's rule for pricing capacity in relatively simple terms, as follows:

> if the same type of capacity serves all users, capacity costs *as such* should be levied only on utilization at the peak. Every purchase at that time makes its proportionate contribution in the long-run to the incidence of those capacity costs and should therefore have that responsibility reflected in its price. No part of those costs should be levied on off-peak users. (Kahn 1988)

This is the rule of "Peak Responsibility" : those that create the peak should pay for the peak. However, as we alluded above, identifying those responsible depends on whether the capacity is required to satisfy random or systematic spikes in demand. If it is random and can occur at any time (e.g., fire service), the procedure is to allocate the cost of the necessary capacity according to the maximum instantaneous demand a customer or class of customers can be expected to place on the system (Hirshleifer, Dehaven et al. 1960). This can be accomplished with a "capacity" charge for the extra capacity that a utility holds in readiness for its customers unexpected demands. It is important to note that this should not be a volume charge but rather a fixed charge that reflects the cost of providing capacity sufficient to insure a reasonable degree of reliability in service.

If, on the other hand, the variation in demand is systematic, occurring during a certain period or periods (e.g., sprinkling demand), then the price for service during the peak period(s) should include the cost of capacity that makes consumption at the peak level possible.[29] This results in a very different pricing formulation then when the variation in demand is random. In this case, a volume charge is not only appropriate, it is necessary for efficiency to obtain. When peak demand occurs during a specific period, then any consumption during that period contributes to the peak and thus to the need for capacity. In other words, any consumption during the peak period is, in part, responsible for the capacity required to satisfy it. Whether a customer's consumption is small or large during that period is of no relevance, since it is the sum of *all* demands that creates the peak. Therefore, the price of a unit of water during the peak period for all customers should reflect the cost of providing this additional amount of water.

It is also true, however, that capacity available for service during the summer peak is also available during the winter off-peak. Indeed, if summer ceased to exist, part of the capacity would still be needed to serve winter. Should not, then, winter be responsible for the share of capacity it requires? Should, then, summer only be charged for the *extra* capacity above and beyond that needed to serve winter? The answer to both questions, in general, is no.[30] As Kahn states:

> Any attempt to shift capacity costs to the off-peak demands, by raising prices for that service above its own separate, incremental cost ... , will cause available production capacity at that time ... to be wasted, and would cut off purchasers willing to pay the additional cost of serving them. Any reduction of the peak ... price below the full joint cost ... would stimulate additional purchases at the peak, requiring additions to capacity that would not be made if buyers had to pay the full opportunity costs of the additional resources required to supply them. (Kahn 1988)

---

[29] Note that the timing of the peak is different for water than, say, electricity. With the latter, the peak tends to occur at certain hours of the day (e.g., late afternoon) and on weekdays instead of weekends. While daily and hourly variation in water use also occurs, the ability to store large amounts of water and the need for fire flow capacity make it of much less significance than for electricity. For water, the seasonal peak is the focal point of peak pricing issues.

[30] Actually, a better general answer to the two questions posed would be "it depends". However, answering the questions in this way would lead us too far astray. The reader interested in a very lucid and detailed discussion of capacity costing from the economist's point of view is referred to (Hirshleifer, Dehaven et al. 1960; Kahn 1988).

In the case of water service, it is the second instance that is of most importance. The underpricing of service during the peak period results in over consumption, and, in the long-run, will encourage over development of water resources. From a practical standpoint, would adjusting the peak period price to reflect the full cost of water service have much impact on consumption? The answer depends on the responsiveness of demand to price. As we recall from Chapter 1, residential demand for water service during the summer (the peak period for water service) is markedly more elastic than for the winter, and that water demand for outdoor uses is more elastic still. Since outdoor use tends to drive the peak, this would suggest that underpricing service during the peak period could indeed have a significant impact on consumption.

4.3.2 Implementing Seasonal Rates

Before a utility can implement seasonal rates, however, it must address several additional considerations. These include the stability of the peak period with respect to price and time, measurement, and administrative feasibility. We address each of these in turn.

A legitimate concern for a utility is that the institution of peak-load pricing may simply shift the peak to another period. Faced with a higher price in the peak period, customers shift, en mass, to the off-peak period where price is lower.[31] Whether or not this is a significant possibility depends on the cross-price elasticity between the two periods (i.e., the degree to which a change in price in one period affects the demand for the good in another period). In the case of seasonal rates for water, it is very unlikely that this would pose a problem. To a great extent, the difference in demand between the peak and off-peak periods is determined by factors other than price, such as climate. While an increase in the price of water during the peak period is likely to induce a decrease in consumption during that period, it is not likely to induce an appreciable increase in consumption during the off-peak period. Overall, the level of demand is likely to fall.

Another equally counterproductive possibility exists, however. Suppose that upon instituting a seasonal rate schedule the utility discovers that while the average level of demand during the peak period has declined, the actual peak day or hour demand, which determines

[31] There are many examples, such as telecommunications (where customers switch some calls and faxes to the evening), airlines (where people switch vacations to off season months), and electricity (where, for example, electric storage heaters are designed to use electricity during the evening hours). Given that the peak for water is seasonal rather than daily or hourly, peak shifting is much less likely to occur.

the capacity requirement, has remained unchanged. This phenomenon has been labeled "needle peaking" (Beecher, Mann et al. 1990). If we were to graph the pattern of consumption the peak would appear as a needle. Overall, the demand for water during the peak period has declined, which is good from the point of view of conservation, but the level of capacity required to meet demand has remained unchanged, which may be very bad from the point of view of utility finance. It means that a greater amount of capacity must stand idle for a longer period of time, thus deteriorating annual load factors and possibly eroding revenue.[32] Whether this is a significant risk is largely an empirical question that must be answered with the hard hand of experience.

A final consideration is that the peak may shift through time for reasons unrelated to price. Changes in technology, preference, or social policy each have the possibility of shifting the peak period. The relevance of this is that consumers may base long-term investments in water conservation, in part, on the cost of water during the peak period, which ideally would reflect the long-run marginal cost of additional supply. As the peak shifts, so too will the long-run marginal cost, and thus what was a wise investment before the shift may no longer be so wise after the shift. Again, this does not appear to be a considerable risk in the case of water service, where hydrologic conditions maintain a very stable peak.[33]

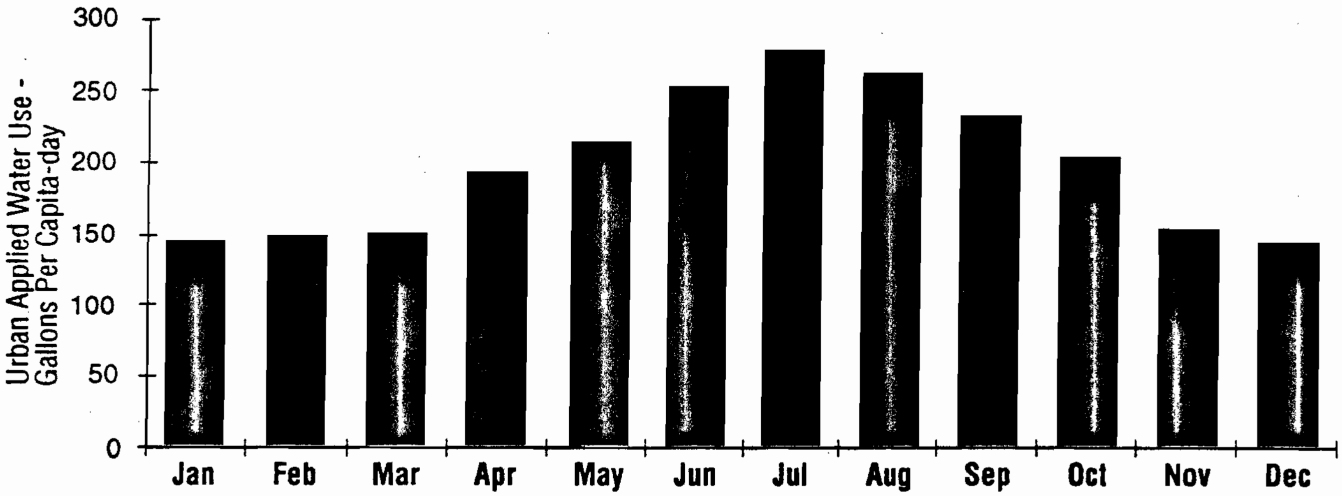
Before a utility can institute peak pricing it must have some way to differentiate between peak and non-peak consumption and to measure system costs that are associated with peak period consumption versus those that are associated with off-peak consumption. In the case of urban water service, identifying the peak season of demand is straightforward. Patterns of demand such as that shown in Figure 4-1 are fairly typical for urban water agencies in the west. As discussed in Chapter 1, urban water utilities also experience daily peaks and valleys in demand in a manner similar to those seen in the electricity supply industry. To measure these demands, however, requires meters that can mark both the amount of consumption and the timing of consumption , which currently is beyond the grasp of almost all water utilities.

---

[32] However, if the decline in demand is purely price motivated, and demand is price inelastic (as all empirical studies suggest), then a positive change in utility revenue would result from a price increase.

[33] Technology, too, has been shown to exert its influence on water consumption patterns, sometimes in rather amusing ways. For example, starting in the late 1940's and early 50's, water agencies started to detect small evening peaks in demand recurring at fifteen and thirty minute intervals. Eventually, it was determined that these peaks were the result of concentrated use of bathroom fixtures during TV commercial breaks. Perhaps with the adoption of ultra-low-flush toilets, evening demand will again resemble its pre-TV pattern.

Figure 4-1
Average Monthly Urban Per-Capita Water Use
Statewide - 1980-87



Therefore, the practical application of peak-load pricing for water service is restricted to seasonal rates.

Measuring system costs associated with peak consumption often is less straightforward, though not impossible. Marginal costs associated with a particular service or class of customer should be assigned accordingly. This corresponds to the pricing rule that a good must cover its variable operating cost. Many system costs, however, are joint costs, and it is sometimes difficult to determine which use or combination of uses is the driving factor. This is particularly true for assigning causal responsibility for capacity. Fire service and peak summer service are the two principle candidates. As we discussed above, the pricing rules for the two types of service are very different -- one being a fixed charge, the other being a volume charge. Thus, it is important that one not be assigned cost responsibility for the other. In general, fire-flow requirements will determine sizing for the distribution and storage network in the immediate vicinity of customers, while seasonal peak demands will determine sizing for most major storage facilities, transmission lines, or pumping and treatment plants.

Finally, before adopting a seasonal rate structure -- or any price structure for that matter --a utility should be reasonably confident that the gains from doing so will exceed the costs. Marginal cost principles, in addition to guiding the pricing of a good such as water, also can be applied to the feasibility and cost-effectiveness of a particular price structure itself. It is always possible that the cost of designing and administering a peak-load price structure will exceed its benefits. Whether this is a likely occurrence depends on numerous factors, including: the cost differential implicit in serving peak and off-peak uses; the price responsiveness of demand during the peak; the cost of measuring and assigning system costs to peak and off-peak uses; and, the cost of metering and billing peak and off-peak uses. For example, given the existing stock of metering equipment in place, it is very unlikely that potential water savings and load factor improvement derived from time-of-day rates would exceed the added cost of metering and billing necessary to implement them. On the other hand, it is very possible that seasonal pricing, which could utilize existing metering devices and would not require drastic changes in billing procedure, would be efficient; though this is not guaranteed to always be the case (Narayanan, Beladi et al. 1987).

4.3.3 Seasonal Rates - Summary

Seasonal rates are a potentially effective means for realizing more efficient use of scarce water resources when demands on a water utility's system vary systematically across seasons. Their primary advantage is that they provide to consumers an accurate signal of the cost of consumption, including the cost of capacity, in a given period. In this regard, seasonal rates have several advantages over more traditional approaches to pricing capacity, including:

- Consumption within periods responsible for capacity costs pay for those costs. Traditional approaches typically spread these costs over all periods. This can increase inefficiency and decrease equity. It increases inefficiency by underpricing water service in the peak period, thus encouraging too much consumption, and by overpricing water service in the off-peak period, thus encouraging too little consumption. It decreases equity because off-peak users, by paying a share of cost that they are not causally responsible for, implicitly subsidize the consumption of peak-users.

- All uses during the peak period are recognized as contributing to and are charged for the cost of meeting the peak. Traditional approaches often base the assignment of costs on non-coincident peak demands that may bear no relation to the actual system peak. If, for instance, a customer class's peak consumption occurs during the system's off-peak, it may actually be beneficial in terms of system load, and should not be discouraged.

- Ideally, seasonal rates will also reflect the full cost of capacity required to meet the peak rather than just that portion in excess of average demand. Traditional approaches, on the other hand, go to great lengths to identify whether capacity is meeting average demand or peak demand requirements. In fact, the capacity is jointly meeting both, and causal responsibility for costs depends on the relative magnitudes of peak and off-peak demands. In cases where the differential between the two demands is large, the peak period will bear responsibility for the costs.[34]

Exuberance for seasonal rates (if indeed such a thing is possible for a water rate schedule) must be tempered by the practical considerations of application. Seasonal rates may present a somewhat more complicated, and hence more expensive pricing formulation. Therefore, the costs of design and administration must be carefully weighed against the potential gains in efficiency. Still, the increasingly common occurrence of a seasonal rate differential would suggest that these costs are not prohibitive for many urban water utilities.

---

[34] The exception being for capacity required for fire flow.

4.4 Increasing Block Rates

Increasing block rates are rates where the volume charge increases for successive quantum of water. For example, a rate that charges $1 per unit for the first 100 units and $1.50 per unit for all subsequent units is an increasing block rate. Many utilities now employ increasing block rates of one variety or another.

An increasing block rate is often proposed to meet two objectives: (1) to promote customer conservation; and (2) to satisfy revenue sufficiency constraints. Its potential to promote customer conservation is well described by the following passage:

> The impact [on conservation] of an increasing block rate design is best illustrated by comparing it to the simplest alternative -- a uniform design where all water is sold at the same price. Because of revenue sufficiency and other constraints, it can be assumed that both rates are initially designed to recover the same total revenue. The increasing block design, then, must contain one or more prices which are higher than the uniform design, and one or more which are lower. Customers facing the higher prices at the margin will, in theory, use less water than they would under the uniform design; customers facing lower prices at the margin will use more. The increasing block design will conserve water if the sum of decreases in use exceeds the sum of increases. ((Metropolitan Water District of Southern California 1991)

The expectation is that demand in the high blocks will be more elastic than demand in the low blocks, resulting in a net decrease in water use. The empirical evidence on elasticity reviewed in Chapter 1 suggests that this is not an unreasonable expectation; but neither is it a guaranteed result. The schedule's design, customer attributes, and regional climate will be important factors to any outcome.

Increasing block rates also are advanced as a means to preserve revenue neutrality. For example, when marginal costs are increasing a marginal-cost based rate will collect more than the utility's revenue requirement. A commonly proposed solution to this is an increasing block rate that sets the marginal price for only the last block and sets price for prior blocks so that the utility breaks even. If all customers face the marginal cost for at least some of their consumption, the efficiency properties of a marginal cost rate design will be preserved. The problem, of course, is that it is exceedingly difficult to design a rate schedule that results in every customer facing the marginal rate for at least some of their consumption. More typical is a rate where some customers pay a high price and some pay a low price. Unless this is cost justified, questions of equity will arise.

4.4.1 Case Studies of Increasing Block Rates

In Section 2.2, we noted that increasing block rates have been used by a number of water agencies at various times. In 1990, for example, sixteen of the sixty plus water agencies served by the Metropolitan Water District of Southern California had increasing block rates, as shown in Table 4-1 (some other agencies within Metropolitan's service area have adopted increasing block rates since then, including Los Angeles whose experience is discussed in this section). Those rates were not necessarily adopted as conservation rates, but it may be instructive to analyze them from that perspective. In most cases, there are only two blocks, as opposed to having many, finely graduated blocks. There is considerable variety with regard to the magnitude of the price differentials between the blocks: in nine cases, the differentials are in the range of five to fifteen percent; in three cases, they are in the range of thirty-five to sixty-five percent; and in five cases, they exceed one hundred percent. What is most striking is the location of the switch points. In ten of the sixteen rates, the switch point is at 125 - 250 gallons per account per day. One thinks of a typical single-family residential account as having three or four people in the household with each having an indoor use of between 60 - 80 gallons per day. Thus the total indoor use for the typical account may range between 210 - 280 gallons per day. Added to this amount is outdoor use, which will vary considerably by account, but that could easily average an additional 100 - 200 gallons per day, bringing the total to between 310 - 480. What these numbers suggest is that the majority of the utilities -- ten of sixteen -- employing increasing-block rates in Metropolitan's service area are locating switch points at levels too low for most single-family households to attain without a very substantial change in the pattern of household water use. A more appropriate location to induce households to switch from above the switch point to below might be somewhere in the range of 300-500 gallons per day.

There are some cases where water districts have adopted increasing block rates with the promotion of conservation as an explicit consideration in the design of the rate structure. We consider three of the more informative examples. The first involves a small water district serving irrigators in the San Joaquin Valley rather than M&I users. While the agricultural context is very different from what we have been discussing, in this instance, the basic principles of rate design remain the same and are particularly well illustrated by this example. The second example is Tucson Water, which has employed increasing block rates, in part to promote conservation, for well over a decade. The third is the Los Angeles Department of Water and Power, which adopted increasing block rates in 1993.

Table 4-1
Retail Agencies in MWD Service Area with Increasing Block Rates in 1990

Agency	Pop. Served	Rate Structure (Marginal price in $/AF) (g/d = gallons/day/account)
LOS ANGELES COUNTY		
Beverly Hills	34,300	$0 up to 250 g/d $433 up to 374 g/d $469 up to 748 g/d $503 up to 1500 g/d $525 above 1500 g/d
Glendale	166,100	$159 up to 125 g/d $182 above 125 g/d
Inglewood	103,000	$601 up to 199 g/d $806 above 199 g/d
Las Virgenes	54,400	Zone 1: $283 up to 150 g/d $828 up to 300 g/d $558 above 300 g/d Zone 2: $923 up to 150 g/d $1032 up to 300 g/d $1198 above 300 g/d
Long Beach	419,800	$177 up to 125 g/d $362 above 125 g/d
Pasadena	132,200	$148 up to 374 g/d $317 above 374 g/d
Pomona	118,000	$136 up to 150 g/d $227 above 150 g/d
Santa Monica	96,500	$266 up to 250 g/d $301 above 250 g/d
ORANGE COUNTY		
Fountain Valley	55,600	$331 up to 250 g/d $392 above 250 g/d
San Clemente	35,100	$0 up to 125 g/d $540 up to 648 g/d $810 above 648 g/d
RIVERSIDE COUNTY		
Riverside	199,400	$65 up to 150 g/d $261 above 150 g/d
SAN DIEGO COUNTY		
El Cajon	83,200	$365 up to 250 g/d $421 above 250 g/d
La Mesa	52,200	$365 up to 250 g/d $421 above 250 g/d
San Diego	1,027,360	$409 up to 125 g/d $468 above 125 g/d
VENTURA COUNTY		
Oxnard	124,000	$0 up to 125 g/d $457 up to 250 g/d $482 above 250 g/d
Simi Valley	94,500	$283 up to 685 g/d $414 above 685 g/d

SOURCE: Metropolitan Water District of Southern California. Water Conservation Pricing Approaches of the Metropolitan Water District Staff Report, October 1990, Table A-1.

What these three utilities have in common is that they attempted to use the rate structure to change the distribution of water use by targeting consumption at the high end of the range of consumption. By way of explanation, when one sees data on distribution of water usage by individual residential accounts there is almost always a distinctive *long right tail* -- a small fraction of the users at the high end account for quite a substantial fraction of total use. Rather than the symmetric, bell-shaped distribution of the normal probability distribution, one finds something closer to the distribution in Figure 4-2. For instance, the top 24 percent of single-family residential accounts in the LADWP service area in 1988 accounted for 47 percent of the total use, and the top 10 percent accounted for almost 27 percent of the total use. Overall, usage per residential account varied from as little as 25 gallons per day to as high as 22,400 gallons per day (there were 5 accounts using this amount). Some of this variation is due to differences in household size or lot size -- but not all. Some of it is due to life-style, habit, and preference (e.g., whether or not people bother to fix leaky faucets, how much attention they pay when watering the yard, etc.).[35]

The existence of this distinctive pattern of household water use, with a long right tail, raises a fundamental question about the design of conservation-oriented water rates. If one wants to reduce overall water use by $x\%$, should one aim to *shift the entire distribution* of water use to the left by $x\%$, or should one seek to *change the shape of the distribution*, pulling in the right tail in such a way as to reduce overall use by $x\%$? The first strategy aims to change *every* customer's water use. The second *targets* customers with substantially above average consumption; it rests on the notion that, if one could just attract their attention and induce them to change their pattern of use, this could lower total use without having to impact the behavior of the entire customer base. Which is better is an empirical matter that will certainly vary with circumstances. If targeting a large incentive at the fraction of users in the right-tail is more effective than offering a smaller incentive for everybody, this strengthens the case for the strategy of changing the shape of the distribution. That was the objective in each of the three case studies, to which we now turn.

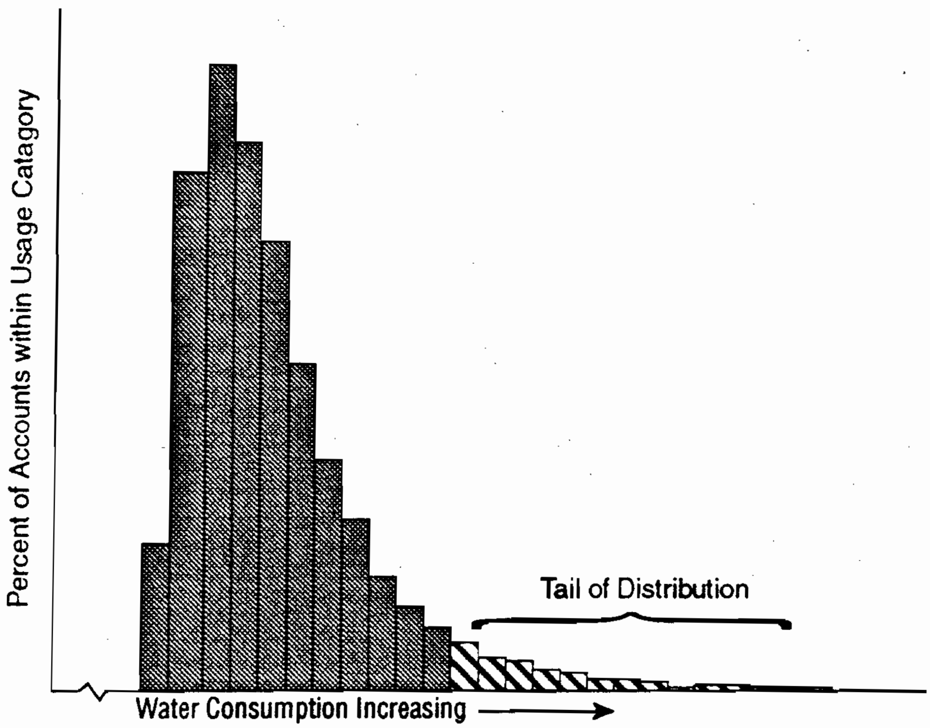
<u>Example 1: Broadview Water District</u>

Broadview water district is located near Firebaugh, California. It serves an area of about 10,000 acres of farmland, making it one of the smaller agricultural water districts in the state.

---

35 The variation in use does *not* reflect difference in price, since these households were all paying the same, flat-rate price in 1988.

Figure 4-2
Typical Distribution of Residential Account's by Consumption Level



Broadview is one of several irrigation districts on the west-side of the San Joaquin Valley implicated in the pollution of Kesterson Wildlife Refuge or the nearby San Joaquin River with selenium laden drainage water. Interim water quality standards for reaches of the San Joaquin River that serve as outlet for Broadview drainage were announced in 1987-88. To comply with the standards, Broadview would have to reduce its drainage into the river by about 15 percent. It was determined that this could be accomplished by reducing the amount of irrigation water applied to crops by approximately 10 percent, since over-irrigation was identified as a major cause of high drain flows.

In October 1988, the start of the 1989 crop year, the district announced a new rate structure. Previously, the district had charged a uniform commodity charge of $16 per acre-foot of delivered water, together with a fixed assessment of $42 per acre served by the district. Part of these charges covered the cost of the district-wide drainage collection and conveyance system. The annual cost to operate and maintain this system amounted to about $21 per acre-foot of collected drainage water or, when averaged over the roughly 25,000 acre-feet of water delivered annually by the district, about $3.08 per acre-foot. The district felt, however, that a surcharge of $3.08 per acre-foot would be too insubstantial to elicit a significant reduction in on-farm water use. Instead, it adopted a two tier rate structure; water in the first block continued to be priced at $16 per acre-foot; water in the second block would be priced at $40 per acre-foot. The idea was to use the revenues from the sales in the second block to cover the drainage system costs. Separate switch points were set for each crop grown in the district; these were set at 90 percent of the district-wide average water usage for the crops over the period 1986 through 1988, as shown in Table 4-2.[36] This was based on the goal of reducing applied water in the district by 10 percent.

The effects of the new rate structure on water use in 1989 and 1990 are shown in the last two columns of Table 4-2. In most cases, there was already some reduction in water use in the 1989 compared to 1986-88 levels. For cotton and wheat, however, efforts at conservation were offset by unusually hot temperatures during key parts of the 1989 growing season. In these cases, the real responses to the new rates came in 1990, when average water use fell 11 percent below the 1986-88 level for cotton and by 5 percent for wheat. For tomatoes, water use fell by 6 percent compared to 1987-88; for alfalfa seed, it fell by 9 percent; and for melons, it fell by 15 percent.

---

36 This corresponds to the notion that switch points, to be effective, must be set at levels that are within the reach of customers being asked to conserve. If the district had simply set a single switch point for all crops, irrespective of individual crop water requirements, the effectiveness of the rate would have been greatly reduced.

Table 4-2
Crop Specific Water Use in Broadview Water District, 1986-1990

Crop	Average Use 1986-88 (af/acre)	Switchpoint (af/acre)	Average Use 1989 (af/acre)	Average Use 1990 (af/acre)	% Change 1986-88 to 1990
Cotton	3.20	2.9	3.34	2.84	-11
Tomatoes	3.22	2.9	2.73	3.03	-6
Melons	2.11	1.9	1.93	1.79	-15
Wheat	2.30	1.9	3.02	2.18	-5
Sugarbeets	4.58	3.9	3.73	2.54	-45
Alfalfa seed	2.06	1.9	1.84	1.88	-9

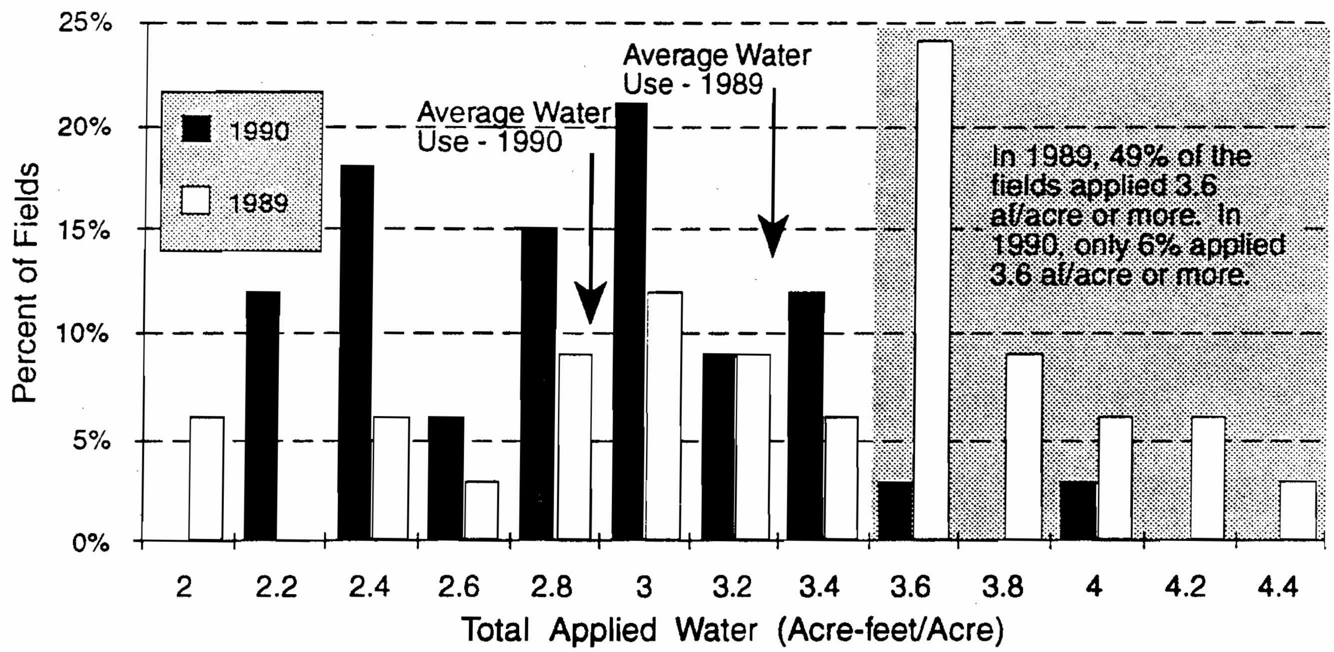
SOURCE:
Dennis Wichelns and David Cone, "Irrigation District Programs Motivate Farmers To Improve Water Management and Reduce Drain Volume." Presented at US Committee on Irrigation and Drainage, 12th Technical Conference on Irrigation, Drainage and Flood Control, San Francisco, CA, November 13-16, 1991.

One can get more insight into the effects of the new rate structure by looking at field-level data that show the change in the distribution of water use. Cotton is the single most important crop in Broadview, being grown on about half the district acreage. Field-level application rates for cotton are available for 1989 and 1990, as shown in Figure 4-3. Cotton was grown on 32 fields in 1989 and 33 fields in 1990. In 1989, the modal application rate (i.e., the rate that was applied on the greatest number of fields) was 3.6 acre-feet per acre. On eight fields, the application was higher than this amount. In 1990, the modal application rate fell 17 percent to 3 acre-feet per acre, close to the switch point of 2.9 acre-feet per acre. As shown in the figure, only one field had an application rate above the 1989 modal rate of 3.6 acre-feet per acre.

Since 1990, water use in Broadview has continued to decline. In 1991, for example, almost no water use on cotton fell in the second block (3.0 AF/a and over). However, the data from 1991 onward are confounded by the effects of the California Drought, because of which deliveries from the Central Valley Project to Broadview and other agricultural contractors were greatly reduced in 1991, 1992, and 1993. It is clear, nevertheless, that there has been a substantial reduction in water use in Broadview and this is likely to be permanent, because there have been numerous improvements in irrigation practice since 1989 which have now become fixtures in the district. There is no doubt that these changes were greatly encouraged by the new rate structure. It offered a large and highly visible price differential combined with a switch point located at a level that was likely to be attainable by most water users.

There are many features of the Broadview case that make it unique, but also highly illustrative of the factors that contribute to the effectiveness of an increasing block rate structure. First, the farmers in Broadview faced more than just a price incentive; there were strong pressures on the district because of the drainage situation and the threat of further action by the state if it failed to reduce drainage volume. Second, the Broadview farmers had detailed knowledge of their water use and monitored it fairly closely; they were fully cognizant of the link between how much water they used and how much money they paid, which is fundamental to customer responsiveness to price. Third, farmers in the district had direct access to information on possible conservation strategies, which was constantly being offered them by the District. Fourth, the district is tiny -- barely over a dozen growers who all know one another. Data on water use in each field within the district were posted in the district office every month so that everybody could see who was doing what. Thus, there was considerable peer pressure to cut back on water use. Fifth, the district was selective in what it chose to target. It did *not* target cropping patterns -- it made no attempt to shift these from more to less water intensive crops. Instead, it chose to target the promotion of irrigation efficiency. It designed a

Figure 4-2
Applied Water Use For Cotton - Broadview Irrigation District



system that left farmers free to grow whatever crops they wanted, as long as they did so in a normatively efficient manner. At the same time, this ensured that the same rate structure was being targeted on a extremely homogenous group of users (only those growing the particular crop).

Thus, the Broadview case study satisfied each of the factors discussed above that influence price responsiveness and make increasing block rates effective conservation measures: rates were tailored to target homogenous groups of users; the rate differentials presented users with significant incentive to curtail use; users themselves had a firm understanding of their use and how they could curb it; and normative expectations were clearly established and adhered to.

In an urban water district the situation is different in every respect. Urban utilities seldom face a homogeneous group of users for which to tailor rates. Typically, a utility's customers are differentiated only up to the point of primary class (e.g., residential, commercial, industrial, and institutional). Within any class, however, the degree of customer heterogeneity is still apt to be very great, which will make it much more difficult to develop a rate with price differentials and switch points relevant to a wide variety of customers. Also, as noted above, most urban users have very little awareness or understanding of their water use and the factors that affect it -- unlike farmers, they are not professional users of water, and their livelihood does not depend on it. Thus, a key factor contributing to the effectiveness of a rate to motivate behavior -- knowledge of use -- may be missing or muted in the urban setting. Finally, normative expectations become increasingly difficult to convey as the size of the group increases. In Broadview, the small number of growers made it relatively easy to impose a group standard. The same will seldom be the case for an urban utility.

Given these differences, it would not be reasonable to expect the same type of conservation potential from an increasing block rate in an urban setting. But, the goal would be to try to replicate the factors that accounted for its success in Broadview to the maximum extent possible. An indication that this may be feasible comes from the experience in Tucson, where an increasing block rate has measurably decreased high-end consumption and resulted in significant water savings.

Example 2: Tucson Water Department

Tucson Water first instituted an inclining block rate in the mid-1970's. In 1977, it combined a four block rate for summer months with a flat rate for winter months. By 1986, it had switched to a six block rate for both winter and summer, though rates in the summer were set higher than in the winter. Water rates were increased each year between 1977 and 1986, though in inflation adjusted dollars water bills either remained constant or declined until 1982. Starting in 1982, rates were adjusted in a way to make it more expensive in real as well as nominal dollars to consume above average amounts of water but left water bills for customers using average or below average amounts mostly unaffected. The notion was to motivate without coercing customers using considerably above average amounts of water to scale back their consumption. They were left free to continue on as they had, but it was going to cost more.

Between 1982 and 1986, water bills calculated with inflation adjusted dollars for single family residential customers with usage three times 1978-79 average usage increased by 26 percent. Over the same period, bills for customers with usage equal to 1978-79 average usage increased about 7 percent, while bills for customers with usage equal to one-half this amount increased about 9 percent. The impact of this rate adjustment on consumption, particularly with respect to usage in the upper blocks, was analyzed by Cutherbert and Nichols (1987). Their analysis indicates that the rates were relatively successful in shifting consumption out of the upper blocks. Normalizing for weather and other factors, they found that for the period 1982 to 1986, annual usage in the upper three blocks as a share of total residential usage declined from 8 percent to 6.6 percent. More importantly, however, weather normalized average monthly usage in the peak usage months of June-July was 11 percent lower in 1986 than in 1982, but was left unaffected during the non-peak winter months, clearly suggesting that the rate structure was effectively curtailing discretionary outdoor use during the peak season when the utility is most at risk for shortage. Annual water savings were estimated to be 550,000 ccf in 1983, 1,000,000 ccf in 1984, 1,500,000 ccf in 1985, and 1,300,000 ccf in 1986. By 1986, then, estimated annual savings from shifting the distribution of consumption away from the high-end range equaled an amount of water sufficient to supply more than 9,000 customers.

Example 3: Los Angeles Department of Water and Power

In the summer of 1991, in the face of a serious drought that had required rationing and emergency water rates to cope with a 15 percent shortfall in supply, Mayor Bradley appointed a Blue Ribbon Committee to consider LADWP's water rates for the future. The Committee made an extensive analysis of LADWP's costs and then proposed an increasing block rate structure designed specifically to target higher use customers.[37] The Committee was committed to the basic principle of marginal cost pricing, but wanted to ensure revenue sufficiency in the face of

[37] One of the present authors (MH) served as a technical adviser to the Committee.

a rising marginal cost of new supply sources. It saw the increasing block rate structure coupled with carefully designed automatic rate adjustment mechanisms as a means of providing revenue sufficiency while at the same time promoting conservation through a targeted incentive. The Committee also was concerned that the adoption of the new rate structure might be misconstrued by some simply as a revenue generating device rather than as a means to promote more efficient use of water. Therefore, the structure was designed to be revenue neutral in comparison to its predecessor for the first year following implementation.[38] Another advantage of the two-tier structure was that it protected more conserving users, whom the Committee felt should be rewarded for their efforts, not penalized. After some analysis, the Committee decided that reclaimed water should be taken as the marginal source of supply for the purpose of estimating off-peak marginal cost.[39] In addition, it proposed seasonally varying rates, based on the change in marginal cost between off-peak and peak periods, with the peak users covering certain of the capacity costs of treatment, transmission and distribution.[40] The committee felt that the incentive offered by two-tier rates could still be effective for consumers who were below the switch point, as long as it was sufficiently close that the higher price loomed in their consciousness and could influence their behavior. Finally, the Committee felt strongly that there should be the same rate structure throughout the city, as opposed to having different structures (e.g., different switch points or different prices) in different areas.

The rates that the Blue Ribbon Committee recommended are shown in Table 4-3 and Table 4-4. The rates are structured differently for single-family residences compared to other users. For single-family residences, the switch point is located at 525 gallons/account/day. The other customer classes -- multi-family residential, commercial and industrial -- are considerably more heterogeneous than the single families and it was felt that, for them, the switch point should be based not on some absolute level of use that would be the same for all users within the class but

---

[38] After the first year, revenue would be allowed to vary from that generated under the old rate schedule. The rate ordinance guarantees LADWP a base revenue of $277 million and includes a Water Revenue Adjustment Factor to adjust rates if revenues fall short of targets.

[39] It was understood that this could change over time as new developments occurred, such as increased supplies becoming available from demand side management or water markets sales by agricultural users.

[40] Having adopted flat-rates in place of declining block rates in the previous drought (1977), LADWP introduced seasonally differentiated flat-rates in 1985. The summer rate was initially set at about 15% higher than the winter rate; by 1992, the differential had risen to 25%. The Committee felt, however, that this differential was too small to reflect the real differences in correctly calculated peak and off-peak marginal costs, as well as too small to attract much notice from water users.

Table 4-3
Water Rates Proposed by LADWP Blue Ribbon Committee

Normal Year Rates

	Price in Low Block ($/ccf)	*Switch Point*	*Price in High Block ($/ccf)*	
			Winter	Summer
Residential				
Single-Family	$1.71	525 gallons/day	$2.27	$2.92
Multi-Family	$1.71	125% of winter use	NA	$2.92
Non-Residential	$1.78	125% of winter use	NA	$2.92

Drought Year Rates

	Price in Low Block ($/ccf)	*Switch Point*	*Price in High Block ($/ccf)*
Residential			
10% Shortage	$1.71	475 gallons/day	$3.70
15% Shortage	$1.71	450 gallons/day	$4.44
20% Shortage	$1.71	425 gallons/day	$5.18
25% Shortage	$1.71	400 gallons/day	$6.05
Multi-Family			
10% Shortage	$1.71	115% of adjusted winter use	$3.70
15% Shortage	$1.71	115% of adjusted winter use	$4.44
20% Shortage	$1.71	110% of adjusted winter use	$5.18
25% Shortage	$1.71	110% of adjusted winter use	$6.05
Non-Residential			
10% Shortage	$1.78	115% of adjusted winter use	$3.70
15% Shortage	$1.78	115% of adjusted winter use	$4.44
20% Shortage	$1.78	110% of adjusted winter use	$5.18
25% Shortage	$1.78	110% of adjusted winter use	$6.05

Table 4-4
Normal Year Rates Adopted by LA City Council

	Normal Year Rates			
	Price in Low Block ($/ccf)	*Switch Point*	*Price in High Block ($/ccf)*	
			Winter	Summer
Residential				
Single-Family	$1.14	575 gallons/day	$2.33	
		725 gallons/day		$2.98
Multi-Family	$1.14	125% of winter use	NA	$2.92
Non-Residential	$1.21	125% of winter use	NA	$2.98

rather on a *relative* level of use, namely usage in the winter. Thus, in the winter there is a single block rate; in the summer, the first block applies to consumption up to 125 percent of winter consumption, and the second block rate applies to consumption beyond this level.[41] The second block rate is the same for *all* users and is based on the estimate of LADWP's marginal cost of supply; it differs between summer and winter because of the peak-load pricing design. The rate for the first block varies among customer classes and is set so as to meet the revenue targets for that class.

In addition, there is a *separate* set of rates for *drought* years. The idea is to set down ahead of time the principles that will be followed when it comes to adjusting water rates in the course of a drought. The same type of block rate structure still is applied in a drought year, but it is modified in two ways to adjust to the shortage. First, the switch point at which the second block commences is reduced, roughly in proportion to the severity of the shortfall. This means that the higher price will be triggered sooner during a drought than during normal supply years. Second, the rate charged in this second block is raised to equal what the Committee estimated to be the rationing price that would equilibrate demand to supply, given the shortfall.[42]

In January 1993, the LA City Council adopted a rate ordinance which closely followed the Committee's recommendations. The main change was to raise the switch point for single-family residential accounts and differentiate it by season, placing it at 575 gallons/account/day in winter and 725 gallons/account/day in summer. This was done mainly to accommodate the interests of residents living in the San Fernando Valley who face a warmer climate and tend to have larger lots than residents in the downtown and coastal areas.

LADWP staff had estimated that 71 percent of the single-family residences would have a lower water bill in a normal year with the new rate structure than with the flat rate structure it replaced. They could have a higher bill in some of the summer months when their usage spilled over into the higher priced block, but not for the year as a whole. Residents of the San Fernando Valley would be somewhat more adversely affected – only 61.4 percent would have a

---

41 It should be noted that sewer charges in the LADWP service area are based on the volume of water used during the winter period. This provides a countervailing incentive against artificially boosting wintertime use for the sake of lowering summer water bills.

42 These equilibrium prices were based on an analysis of LADWP's experience in the summer of 1991 when inverted block rates with punitive upper rates were introduced temporarily to cope with the drought.

lower bill over the course of the year. By contrast, in the downtown and on the westside, respectively, 84 percent and 91 percent of the single-family accounts would have a lower bill with the new rate structure.

In fact, these predictions have borne out well during the first twelve months with the rate experiment. By the end of the fall of 1993, however, it had run into strong political opposition from the San Fernando Valley. Several factors were responsible. In addition to the change in the water bill, there had been an increase in sewer charges, which are included on the water bill along with various other city charges. Customers attributed all of the change to the water rates. This coincided with the first hot months, September and October, when many households experienced a higher bill for the first time, after several months of lower bills. Moreover, a new Mayor of Los Angeles had been elected with considerable support from the San Fernando Valley. As a result, the Mayor reconvened the Blue Ribbon Committee November 1993 and directed it to reconsider the rate structure to make it more equitable for residents of the San Fernando Valley. At this writing, the Committee has still not made a decision; it is inclined to preserve the two-block rate structure, but it is now willing to allow the structure to vary more by user characteristics. It is examining the possibility (and administrative feasibility) of allowing the switch point to vary among consumer classes -- for example by climate zone, or by lot size.[43]

4.4.2 Increasing Block Rates - Summary

An increasing block rate is often proposed to promote customer conservation through higher rates without violating revenue sufficiency constraints. For example, increasing block rates have been used by several utilities to reshape the distribution of consumption by discouraging high-end uses. They can be formulated based on average or marginal cost data. If designed so that every customer pays the marginal cost of service for at least some of their consumption, they can mimic the efficiency properties of marginal cost pricing. The heterogeneity of customer demands, however, makes this exceedingly difficult to accomplish. More typical is a block rate structure that results in some people to paying a higher price for service than others. Unless there are cost based reasons to do so, concerns about equity will emerge.

---

43 Another idea -- having the switchpoint vary by household size -- has been rejected because of the administrative infeasibility of monitoring it.

4.5 Priority Pricing

In the natural gas and electricity utility sectors, separate pricing of firm and interruptible service is widely used as an alternative to peak and off-peak pricing. The notion is to offer customers the choice of reliability of service, where the less reliable service is priced at a lower rate and the more reliable service at a higher rate. Customers opting for less reliable (interruptible) service agree to curtail or cease consumption whenever the demand on the utility exceeds some pre-specified level; in return, they pay a lower price. Customers opting for more reliable (firm) service pay a higher price, but are guaranteed a higher level of reliability of service regardless of demands on the system. Thus, priority is priced. Both the utility and its customers can benefit from this type of pricing scheme. The utility benefits because it can alter the system load factor in a manner that is beneficial (e.g., smooth out the peaks and valleys of demand) and can sometimes forestall expansion of system capacity. Customers can benefit because they are given greater freedom of choice with respect to what type of service (how much reliability) they receive. Within the limits set by the menu of priority rates, they determine how reliable a supply. They are allowed to self-select as to whether they prefer to curtail consumption or pay a higher price during periods of shortage or limited capacity. While fairly common to other utility sectors, priority pricing (as we shall refer to it in this discussion) is not common in the urban water industry. In the remainder of this section, we explore the concept of priority pricing in more detail, and consider its applicability to the urban water sector.

4.5.1 Rationale for Priority Pricing

The goals of priority pricing are two-fold. First, it helps utilities match the cost of supply to the value of service, and thus better allocate a scarce resource in an environment of increasing supply and cost uncertainty. Second, it allows customers with flexible demands to lower their cost of service without significantly reducing the quality of their service. Proponents of priority pricing see it as offering practical solutions to some of the everyday problems of utility management, such as:

- How can the unit cost of service be fairly and economically lowered to customers with flexible demands willing to curtail usage during shortages?
- How can the development of expensive new supply be avoided or forestalled?
- How can the need for new transmission and distribution capacity be avoided or forestalled?

- How can the cost to customers of shortages be reduced?

Priority pricing typically is used as an alternative or supplement to non-price rationing to manage emergency shortages. Rather than impose rationing on all customers, priority pricing allows customers that can curtail usage at lowest cost to self-select themselves. Customers with low outage costs are likely to select a less reliable level of service if it is offered at a lower cost than regular service, whereas customers with high outage costs are likely to select a more reliable level of service at a higher cost. In this way, the overall cost of the shortage to the customer base is reduced. The concept is very similar to that of emission credit trading for the control of pollution, whereby firms with the lowest pollution control costs willingly assume a disproportionate share of emission control in return for financial consideration. Priority pricing is also viewed as a way to achieve longer term management objectives, such as better load management and deferred system expansion. Both objectives can reduce the financial risk for a utility; the first by improving utilization of existing capacity and the second by limiting exposure to financially risky investments. Thus, by recognizing the diversity in customers' needs in terms of time of use and service quality, utilities can provide incentives that promote complementarity in usage patterns (Chao, Oren et al. 1988).

4.5.2 Can Priority Pricing Work in the Water Industry?

The adoption of priority pricing in the electric, gas, and telecommunication sectors has made significant strides in recent years. In the electricity sector, for example, priority pricing for industrial customers is common and several experimental programs for commercial and residential customers are underway (Electric Power Research Institute ; Chao, Oren et al. 1988). Electric utility regulators are encouraging and in some instances mandating that utilities unbundle the services offered to customers. For instance, in the mid-1980s, Massachusetts ordered Boston Edison to offer interruptible rates to all customers and indicated that all electric utilities in Massachusetts will eventually be told to do the same (Chao, Oren et al. 1988). In a 1987 speech, the Federal Energy Regulatory Commission Chairwoman stated that utilities should work "to allocate electricity supply to its most valuable uses when the amount demanded exceeds available supply."[44] She went on to state that "[c]ustomers should be apprised of the cost consequences of different reliability levels and then be allowed to choose the level that suits their needs and pocketbooks."

---

44 As quoted in Chao, Oren et al. (1988)

Could urban water utilities and their customers benefit by adopting a similar approach to managing demand? Certainly the concepts are applicable. During shortages, water agencies frequently request or mandate that customers curtail usage. It is not uncommon for this request to be followed by a request for additional revenue through higher rates, leaving those conserving with a sense that they are being punished for their efforts. With priority pricing, customers understand up front the reliability and cost of the service they are being provided, and, in principle, the utility should be able to devise a menu of rates that will provide both the necessary curtailment in usage and sufficient revenue during shortages. Additionally, utilities that are approaching present system capacity also may be seeking ways to defer costly expansion. Priority pricing is one way to do this. Customers most able to reduce their demands on the system during peak periods are provided incentive to do so.

While the concepts of priority pricing are applicable to the urban water sector, there are three factors that may impose practical constraints on its adoption by water agencies. The first is metering and control technology. To be able to institute priority pricing, utilities must be able to unbundle the services they provide, which requires relatively sophisticated metering and control technology (Chao, Oren et al. 1988). It requires that utilities be able to monitor customers' usage to ensure that they are adhering to the terms of the contract. Water utilities can do this to a certain extent with existing meters. However, monitoring would be limited to ex post verification of consumption, which would increase the risk to the utility that contract terms were being violated. This probably could be counteracted with a system of fines that penalized customers for violating their contracts. Since priority pricing would most likely be employed to curtail usage during the peak season, utilities would need to synchronize the billing cycle with the season cycle.

The second factor is the utility's practical ability to guarantee contractual levels of service reliability. For electricity there are spot markets and flexible (but expensive) types of generating capacity such as gas turbines which can be used as a last resort. While water markets are evolving, these are mainly not spot markets, and there is nothing yet analogous to gas turbine sources of supply that can be activated at short notice. Given the nature of supply, it may be harder in the water utility industry to maintain a given degree of reliability at short notice. The water utility, therefore, may need to include escape clauses in the contracts or decrease the level of reliability offered. Whether a sufficient number of customers will still find the price/reliability tradeoff worthwhile will depend on the overall terms of the contract. Ultimately, the precision with which a utility could offer a level of service will depend on its reliability planning process. As water utilities become more experienced with reliability planning, they will be able to offer more differentiated levels of service.

The third factor is consumer attitudes to coping in a drought and the emphasis that people place on solidarity and everybody pitching in together. In some focus groups on water reliability in various parts of California, when the question of priority pricing was raised the reaction was somewhat negative. Participants said that they thought there would be intense peer pressure from friends and neighbors to comply with requests for voluntary conservation in the face of a drought. This would undercut the benefit from having contracted for a higher than normal level of reliability -- if you had the only green lawn in the neighborhood, it just wouldn't be worth it. Such considerations may be less of a factor for commercial or industrial users; it may be that those sectors offer the best prospects for priority pricing in the water industry at this time.

As with any rate structure, the utility will need to weigh the cost of achieving its demand management objectives with priority pricing versus other approaches. The costs to its customers as well as the costs to itself need to be considered. For instance, it is possible that the administrative costs would be less for non-price rationing than for priority pricing, but that the outage costs to the customers would be greater. The sum of the two would then determine which is the most desirable from a social perspective.

4.5.3 Priority Pricing - Summary

Priority pricing represents a marked deviation from traditional pricing of utility service. It offers to the customer price/service combinations on several levels rather than just one. By doing so, it offers customers the ability to better match their needs with available supply, and therefore can improve efficiency. Before a utility could implement priority pricing, however, it would require sufficiently accurate metering and control technology as well as the ability to ensure a system reliability congruent with the levels of service it is offering. Given hydrologic variability, this will be more difficult for water suppliers than for other utilities with more stable supply sources. Still, as the water sector's experience with reliability planning deepens, it is likely that this type of pricing will become increasingly attractive.

4.6 Rate Structures and Utility Finance

During California's 1987-92 drought, urban water supply agencies faced considerable financial hardship as a result of declining sales. It is, therefore, of little surprise that these same utilities view with some concern rate structures touting to further discourage the use of water.

Utilities are caught in a classic "Catch-22". While they need their customers to conserve water, they also need their customers to continue to buy water.

Revenue derived from rates provides a significant share of a utility's operating capital and is integral to its overall financial structure and well being. A utility considering a new rate structure will want to know how it may affect its financial structure, cost of capital, and ability to meet current obligations.

4.6.1 Capital Structure of Urban Water Supply Agencies

The urban water supply industry is capital intensive, requiring the construction, maintenance, and replacement of large-scale, long-lived plant and equipment. To finance these investments, the industry relies on a variety of capital instruments. As shown in Table 4-5, these can be grouped into four categories: (1) internal revenue sources, (2) external financial market sources, (3) customer sources, and (4) government sources. As shown in the table, some instruments available to publicly-owned utilities, such as general obligation or revenue bonds, are not available to investor-owned utilities, and vice-versa. Table 4-6 shows the general differences in debt structures for publicly-owned and investor-owned utilities. Whereas, publicly owned utilities finance most of their capital expenditures with municipal bonds, privately owned utilities mostly rely on retained earnings.

A utility's rate structure enters into this picture in two interrelated ways. First, revenue from rates is itself a source of capital financing for the utility. Utilities use retained earnings or operating surpluses, depending on whether they are privately or publicly held, to finance new projects and service debt. These are perhaps the lowest risk, lowest cost funds available to a utility. Second, revenue stability can be an important factor influencing a utility's credit rating, which determines accessibility to and cost of external financing options.

4.6.2 Risk, Credit, and Rate Structure

To determine credit risk, credit rating agencies assess a utility's ability to meet current and future financial obligations. There are several ways to assess risk. A variety of factors must be taken into account, including regional business environment, utility capital structure, and regulatory environment. Financial ratios can give some indication of the riskiness of a particular aspect of a utility's financial position. Table 4-7 lists financial ratios commonly used to assess the risk position of water utilities. Profitability, profitability trend, liquidity, leverage, growth and efficiency, and efficiency and profit are the principle indicators for investor-owned utilities. Publicly owned utilities operate without profit, so profitability measures are not

Table 4-5
Capital Financing Mechanisms for Water Utilities

Sources of Financing	Privately Owned Utilities	Publicly Owned Utilities
Internal Revenue Sources of Capital Financing		
1. Expensing of Capital Improvements	Yes	Yes
2. Retained Earnings	Yes	No
3. Depreciation	Yes	Yes
4. Tax Credits	Yes	No
5. Operating Surpluses	No	Yes
External Sources of Capital Financing		
1. Common-stock sales	Yes	No
2. Capital payments from affiliates		
a. Debt advances	Yes	No
b. Equity contributions	Yes	No
3. Debt sales		
a. Mortgage bonds	Yes	No
b. Debentures	Yes	No
c. Commercial paper	Yes	Yes
d. Demand notes	Yes	Yes
e. Demand bonds	Yes	Yes
f. Short-term notes	Yes	No
g. General-obligation bonds	No	Yes
h. Revenue bonds	No	Yes
i. Revenue-anticipation notes	No	Yes
j. Bond-anticipation notes	No	Yes
4. Bank and institutional loans		
a. Lines of Credit		
(1) Construction	Yes	Yes
(2) Working capital	Yes	Yes
b. Collateral and property loans	Yes	Yes
Customer Sources of Capital Financing		
1. Construction advances	Yes	Yes
2. Construction contributions	Yes	Yes
3. Connection charges	Yes	Yes
4. Property assessments	Yes	Yes
5. Special rate structure for new connections	Yes	Yes
Government Sources of Capital Financing		
1. Federal loan programs	Yes	Yes
2. State infrastructure bank loans and grants	Yes	Yes
3. Economic development programs	Yes	Yes
4. Joint-venture authorities for construction	Yes	Yes
5. Conversion of municipal departments to authorities	No	Yes

Source: Reproduced from Beecher, et al. (1993) Table 4-2.

Table 4-6
Capital Structure of Publicly and Privately Owned Water Utilities

Source	Percent of Total
Publicly Owned Utilities	
Tax-exempt municipal bond market	60%
Operating surpluses	20-30
Intergovernmental aid	5-10
Other sources	5-10
Bank loans	
Special tax assessments	
Developer contributions, etc.	
Privately Owned Utilities	
Retained earnings	40-50%
Stocks and taxable bonds	20-30
Industrial revenue bonds	10-20
Other sources	20-30
Bank loans	
Developer contributions, etc.	

Source: Beecher, et al. Table 4-1.

Table 4-7
Financial Indicators Used to Assess Credit Risk

Indicator	Ratio
Profitability	Cash flow/sales
Profitability trend	Retained earnings/common equity
Coverage	Revenue available for debt service/total debt service requirement
Liquidity 1	Current assests/liabilities
Liquidity 2	Net fixed assets/total assets
Leverage 1	Book common equity/total assets
Leverage 2	Total debt/total assets
Leverage 3	Current liabilities/total debt
Growth and efficiency	Sales/total assets
Efficiency and profit	Operating revenues/operating expenses

Source: Adapted from Beecher, et al. (1993) Table 4-9 and Table 4-10.

applicable. An alternative is the *coverage ratio*, which measures net revenue available for debt service as a fraction of the utility's average annual debt service requirement. The higher the coverage ratio, the more cash reserves available for debt obligations and the lower the risk of default.

Sales revenue factors into each of the ratios listed in Table 4-7 directly or indirectly. A rate structure that decreases a utility's revenues or introduces additional volatility can affect the ratios negatively and may result in a lower credit rating. Utilities may initially resist alternative rate designs because they are unfamiliar with the consequences they hold for sales revenue. The rate analyst will need to demonstrate that an alternative design can promote more efficient use of water without introducing an unacceptable level of revenue volatility.

4.6.3 Rates and Revenue Volatility

Swings in demand produce volatile revenues. The degree of volatility is determined in part by the rate design. For example, a 10% reduction in demand would not produce any change in revenue if the rate were a simple fixed charge; would decrease revenue by 10% if the rate were a uniform volume charge; and would decrease revenue by less than 10% if the rate were some combination of the two. The desire for more efficient rate structures must be balanced against the utility's need for stable revenues.

In general, utilities have progressed from rates that provide very stable revenues but undesirable efficiency properties to rates that provide less stable revenues but have more desirable efficiency properties. At the turn of the century, the dominant rate structure consisted of a fixed charge. A fixed charge makes revenues impervious to swings in demand but does not provide any incentive to use water efficiently. By the 1960's, the dominant rate structure was the declining block rate. Revenue in the higher priced initial blocks provided stability; but the declining blocks could result in uneconomic consumption. By the 1980's, the uniform rate schedule became widely used. Uniform rates eliminate the inefficiency of declining block rates, but also make revenue more susceptible to swings in demand. Now increasing block rates are gaining popularity. While providing incentive to conserve, they recover a significant percentage of the revenue requirement in the more demand variable upper blocks.

We return to our elfin utility with the magic spigot in the woods introduced in Chapter 2 to further illustrate the different revenue properties embodied in declining block, uniform, and increasing block rate designs.[45] Suppose the total annual cost for our elfin utility is $100. Because the elves are heavily regulated, they are not allowed to earn a profit but can only recover the cost of providing service. Annual demand is usually about 2,000 ccf but can vary from this level depending on weather and economic activity. The elves have been charging a uniform rate of $0.05/ccf but now are thinking of switching to a block rate. The retired elf that runs the utility lost his pension during a reckless gambling spree in Las Vegas and now is working for salary. Having learned his lesson in Vegas, he does not want to gamble again with new rates. To reassure himself, he does a few calculations involving the proposed rates.

Two alternatives to the uniform rate have been proposed to the Board of Director elves: the first is a declining two block rate, with the price of the second block half that of the first and a switch point set so that 80% of consumption falls in the first block and 20% in the second; the second is a inclining two block rate, with the price of the second block twice that of the first and a switch point also set so that 80% of demand falls in the first block and 20% in the second. All three rates are designed to recover $100 if demand is 2,000 ccf.

We reproduce the retired elf's calculations in Table 4-8, which shows the prices, demands and revenues for the three rate schedules. The elf has noted from his long experience managing the utility that it is not uncommon for sales to decline by as much as 10%, either because of an economic downturn, or because scheduled maintenance at the spigot has temporarily reduced supply which required rationing. Therefore, he calculates the impact a 10% decline in sales would have on revenue under each of the three rate designs. To analyze the block rates, he considers two different scenarios: in scenario 1, only sales in the second block decline while sales in the first remain unaffected; in scenario 2, sales in both blocks decline by an equal percentage.

The results of his calculations also are shown in Table 4-8. Several observations can be made. First, under both scenarios, the uniform rate caused sales revenue to decline by 10%. This can be generalized to the following statement:

- With a uniform rate, an x% variation in sales will result in an x% variation in sales revenue.[46]

---

[45] There is no need to include a rate that is purely a fixed charge since its revenue properties are clear.

[46] The following generalizations refer only to the change in revenue from sales. With revenue from fixed charges or other sources invariant to sales, the change in total revenue would be less than depicted here.

Table 4-8
Elf Utility Rate Proposal

	Price ($/ccf)		
Block	Declining	Uniform	Inclining
First	0.056	0.05	0.042
Second	0.028	NA	0.084

	Demand (ccf/yr)		
Block	Declining	Uniform	Inclining
First	1600	2000	1600
Second	400	NA	400

	Sales Revenue ($/yr)		
Block	Declining	Uniform	Inclining
First	89.6	100	67.2
Second	11.2	NA	33.6
Total	100.8	100	100.8

Scenario 1:
Demand declines by 10%; for block rates, only demand in second block is affected.

	% Decline in Sales Revenue		
Block	Declining	Uniform	Inclining
First	0	10%	0
Second	6%	NA	17%
Total	6%	10%	17%

Scenario 2:
Demand declines by 10%; for block rates, demand declines by 10% in each block.

	% Decline in Sales Revenue		
Block	Declining	Uniform	Inclining
First	9%	10%	7%
Second	1%	NA	3%
Total	10%	10%	10%

For example, if R = F + PQ, where R is total revenue, F is fixed revenue, and PQ is sales revenue, then the percentage change in revenue given a percentage change in sales revenue $PQ/(F+PQ) \leq 1$.

Second, under scenario 1, the declining block rate caused sales revenue to decrease by less than 10% while under scenario 2 it caused revenue to decline by 10%. This result may also be generalized to the following:

- With a declining block rate, an x% variation in sales will result in a less than an x% variation in sales revenue if the sales variation is concentrated in the upper blocks; but will result in an x% variation in revenue if it is evenly distributed across all blocks.

Third, under scenario 1, the inclining block rate caused sales revenue to fall by more than 10%, while under scenario 2, it caused revenue to decline by 10%. This can be generalized to the following statement:

- With an inclining block rate, an x% variation in sales will result in a more than x% variation in sales revenue if the sales variation is concentrated in the upper blocks; but will result in an x% variation if it is evenly distributed across all blocks.

From his analysis, the retired elf concluded that a declining block rate has the capacity to dampen the variation in revenue inherent in a uniform rate while an inclining block rate has the capacity to amplify it. He also noted that the degree of dampening (in the case of a declining block rate) or amplification (in the case of an inclining block rate) has two principle determinants: (1) whether the response in demand is concentrated in a subset of blocks or more evenly distributed across all blocks; and (2) the height of the blocks. As the old elf saw, if the response in demand is evenly distributed across blocks, an inclining block rate will not amplify nor will a declining block rate dampen revenue volatility relative to the uniform rate. The degree of dampening or amplification of revenue volatility increases linearly with the concentration of the response in the upper blocks.[47] The price differential between blocks also affects revenue volatility. The larger the differential, the greater will be the dampening or amplification relative to a uniform rate.[48]

---

[47] The relationship between the change in revenue and the distribution of the sales decline is described by the following formula: $\Delta R = P_2 \Delta Q + \alpha \Delta Q(P_1 - P_2)$; where ΔR is the change in revenue, ΔQ is the change in demand, P_1 and P_2 are the first and second block prices, and α is the fraction of the sales decline that occurs in the first block.

[48] The relationship between revenue volatility and price differential can be described as follows: let $P2 = \beta P1$ with $\beta > 0$; then with a two block rate, $R = \alpha P_1 Q + (1-\alpha)\beta P_1 Q$; and $dR/dQ = P_1(\alpha + (1-\alpha)\beta)$, which says that the change in revenue given a unit change in sales increases as β increases.

A utility will want to know how much additional revenue volatility to expect if it switches to a rate structure that has potential to produce more volatile revenues. This is a difficult question to answer, since, as we saw with the elves' utility, it depends on the particulars of the rate design and the behavior of demand, both of which are likely to be much more complicated in the real world than in a fairy tale. Currently, research is underway to address this question, but results are not yet available. One might surmise from the prevalence of utilities employing increasing block rate structures that revenue volatility associated with that design is of a manageable level. One the other hand, many of these same utilities have found it difficult to manage swings in demand caused by extended drought. The prudent course, then, is for a utility to carefully evaluate the possible revenue outcomes for the rate structures it is considering, and choose one that best matches its efficiency goals while still meeting its revenue stability requirements.

In addition, other methods might be used to insulate utilities against too much revenue volatility. Los Angeles Department of Water and Power (LADWP), for example, has recently adopted two rate schedules, each tailored to a different set of cost and demand conditions. The first pertains to the costs and demands associated with normal or above normal water years and the second to the costs and demands for below normal water years.

A second approach is to use a rate adjustment clause to better balance the risk of revenue shortfalls between the utility and its customers. These type of clauses originated in the energy utility sectors during the 1970's when fuel prices were unstable. For example, the Energy Cost Adjustment Clause (ECAC) used in California establishes a balancing account for the utility to keep track of over or under collections, which are then refunded or recovered through rate adjustments in subsequent periods. This approach, however, can create a conflict between maintaining the financial stability of a utility by reducing the risk associated with varying revenues and maintaining incentives for the utility to minimize its costs. To preserve the incentive to minimize expenditures, adjustment clauses generally allow only partial rather than full recovery of undercollections.

Publicly-owned water utilities frequently employ reserve accounts to even out revenue fluctuations. Rates recover annual expenses plus a little extra to be held in reserve for periods of restricted cash-flow. This approach is consistent with cost-based rate making principles, since managing cash flow is an essential aspect of running a utility.

4.6.4 Rates and Utility Finance Over the Long-Run

Alternative rate designs discussed in this chapter have the capacity to strengthen a utility financially over the long-run. Peak load pricing, increasing block rates, and conservation surcharges that improve long-run resource allocation and operational efficiency can forestall costly system expansions(Beecher, Mann et al. 1993). Conservation surcharges, for example, can be calculated to recover the avoided or deferred cost of additional capacity that would be required if growth in demand was not reduced. These approaches are consistent with integrated resources planning where they would be viewed as alternative sources of supply as well as sources of revenue.

4.6.4 Ratemaking Reforms to Improve Utility Finance

Regulatory or administrative lag associated with ratemaking can destabilize revenue, discourage investment, and increase financial risk for a utility. Investor-owned utilities, in particular, face this risk since the flexibility with which they can design or adjust rates is largely determined by a regulating authority. Expedited rate proceedings and preapproval of expenditures are possible ways to lower the financial risk to utilities associated with the regulatory process.

Investor-owned utilities will be reluctant to incur costs for conservation and other demand-side-management programs if it is uncertain that these costs will be recoverable. Expenditure preapproval would reduce this uncertainty and lower the risk associated with these investments. Like rate adjustment clauses, however, expenditure preapproval can reduce the incentive to minimize costs since risk has been shifted from the utility to the ratepayer. A preferred approach would preserve the incentive to control costs by having the utility share the risk of its expenditure decisions while at the same time keeping this risk at a manageable level. A set of preapproved criteria for expenditures could accomplish this. The criteria would establish reasonable and prudent cost standards for different investments and expenses. A utility could then expect to recover costs through rates so long as they were in accordance with the cost criteria.

4.6.5 Rate Structures and Utility Finance -- Summary

Revenue from rates are an integral part of a utility's financial structure. Swings in demand increase risk of insufficient revenue, which may exclude a utility from taking advantage of some finance options and raise the cost of others. The degree of revenue volatility is determined in part by the rate design. For example, it was shown that a declining block rate has the capacity to dampen the variation in revenue inherent in a uniform rate while an inclining block rate has the capacity to amplify it. The degree of dampening or amplification depends to a large extent on the particular features of a block rate schedule and the behavior of demand.

Over the long-run, the rate designs discussed in this chapter have the capacity to strengthen a utility financially. Peak load pricing, increasing block rates, conservation surcharges, and priority pricing that improve long-run resource allocation and operational efficiency can forestall costly system expansions and result in a more even utilization of capacity, both of which can lower unit costs. Finally, reforms to the ratemaking process itself -- such as expedited proceedings or formalized criteria to recover certain costs -- could further reduce revenue uncertainty and improve system management and planning.

CHAPTER 5 LEGAL ASPECTS OF RATEMAKING IN CALIFORNIA

You should punish your appetites rather than allow yourself to be punished by them.

-- Epictetus, *Fragments*

5.1 Introduction

As discussed in Chapter 2, because of the different perspective to rates held by utility, customer, and society, and because many parties have a stake in the rate setting process, rate design in practice is an inherently political process. In many cases, the discretion that an urban water supply agency has in setting water rates to promote conservation depends in part on the legal and regulatory codes that govern it. In general, we have found that publicly-owned water utilities enjoy broad legal discretion in setting rates to manage demand, recover cost, and stabilize revenue. Investor-owned utilities, on the other hand, do not have a similar flexibility. These utilities are subject to traditional rate of return regulation, and through its authority to permit or disallow cost recovery, the California Public Utilities Commission (Commission) exerts great power over the pricing behavior of investor-owned utilities. This chapter reviews the legal and regulatory constraints in setting water rates as they pertain to urban water supply agencies. We divide the discussion into two sections. In the first we review the California Water Code and case law as it pertains to pricing by publicly-owned water utilities and in the second we review the California Public Utilities Code as it pertains to pricing by investor-owned water utilities.

5.2 Ratemaking Powers of Publicly-Owned Water Agencies

An examination of the Water Code and relevant case law reveals two important facets of the ratemaking powers of publicly-owned water agencies that directly pertain to water conservation:

(1) The California Water Code under section 375 authorizes any public entity which supplies water at the retail or wholesale level to "encourage water conservation through rate structure design;" and

(2) Municipalities and special districts have generally been accorded discretion by the courts in establishing rates that collect revenue in excess of annual expenditure and using excess revenue to fund the replacement, expansion, and *management* of the system.

5.2.1 Water Code Section 375

On August 25, 1993, the Governor signed into law A. B. 1712, amending California Water Code section 375 to codify the right of public water agencies to adopt and enforce conservation programs within their service areas, and to adopt water rate structures that promote water conservation. Specifically, the amended section provides that "notwithstanding any other provision of the law ... any public entity which supplies water at retail or wholesale ... may ... adopt and enforce a water conservation program to reduce the quantity of water used ..." and "with regard to water delivered for other than agricultural uses ... may also encourage water conservation through rate structure design." Water Code section 375. Section 375 defines a "public entity" to include cities, counties, special districts, or any other political subdivision of the state. Water Code section 375. Section 375 does not define what constitutes a conservation promoting rate structure, which seems to imply that the law gives agencies discretion in the design of conservation-oriented rates. In past court decisions, for instance, the court has generally presumed that an agency's adopted rate structure is both reasonable and necessary for the management of its supply system.

5.2.2 Case Law Affecting Rates for Public Water Agencies

A general principle that has guided the setting of water rates is that rates should, to the degree possible, be proportional to cost of service, and that water should be provided at the lowest possible cost. This principle is reflected in numerous Government and Water Code sections defining the ratemaking powers of public water supply agencies. For instance, Government Code section 54514 requires local agencies to supply water "at the lowest possible cost," and Water Code section 31007 directs County water districts to only "collect charges sufficient to ... provide for repairs and depreciation of works owned or operated by the district and to provide a fund to pay principle on bonded debt as it becomes due," implying that water rates should not exceed cost of service. However, it is clear from case law that the courts interpret cost of service rather broadly, leaving to the supply agencies themselves the determination of what constitutes reasonable costs of service.

For instance, as to the discretion accorded water agencies in setting rates, in General Engineering and Drydock Company v. East Bay Municipal Utilities District (1932) 126 Cal. 349, the court interpreted the language in Water Code Section 47180 stating that "tolls and charges shall be proportioned as nearly as possible to the services rendered" to mean that it is up to the district to determine what is "possible" in setting rates. In Arcade County Water District v. Arcade Fire Department (1970) 6 Cal.3d 232, the court concluded that "it must be presumed that the district adopted reasonable rules and regulations as to the supplying of water... thus, the presumption of reasonableness applies to the rates so fixed." As for what constitutes cost of service, the California Supreme Court in Hansen v. City of San Buenaventura (1987) 42 Cal.3d 1172, 1187 interpreted Government Code 54514 (cited above) to mean that cities are only barred from earning unreasonable profits, but are allowed to earn a "reasonable return on investment," and should be compensated for replacement funds, debt payments, *risks incurred in operating a water system*, and the opportunity costs of owning a water system. Hansen, 42 Cal.3d at 1182.[49]

5.2.3 Legal Challenges to Conservation Oriented Rate Structures

The pending case Brydon v. East Bay Municipal Utility District, Civ. No. A060031 appears to be the only example of a legal challenge to a water agency's authority to adopt conservation-oriented rates during the most recent drought. In this instance, the Petitioners made two arguments challenging the legality of the adopted rate structure, both of which the district court rejected. First, the Petitioners argued that the increasing-block rate structure adopted by East Bay MUD's Board of Directors was unreasonable because it did not reflect cost of service and it unreasonably discriminated against water users in hot-climate areas. Second, they argued that it produced excess charges which constituted a "special tax" that required approval by two-thirds of the electors in the district.

- **East Bay MUD's Drought Rates**

In reaction to the recent drought, EBMUD established a 15 percent mandatory conservation goal based on 1986 figures and adopted a Drought Management Program, which included an inclining block rate structure. Under the inclining rate structure, each single-family residential customer was allotted 250 gallons of water per day. Water use in excess of the allotment was billed at a higher rate, with this rate increasing with consumption.

- **The Reasonableness Argument**

The Petitioners argued that East Bay MUD had failed to show that the inclining block rate structure was rationally related to the District's conservation goals since water users had voluntarily reduced water use (the previous year) by a greater amount than the mandatory amount established by East Bay MUD. Appellant's Brief at 9. Moreover, it was argued that East

---

[49] There also exists the possibility that Hansen will be interpreted to apply to all public water agencies, not just municipalities. This interpretation, however, has not yet been rendered by the courts.

Bay MUD had failed to show that the cost of providing water at an amount above the 250 gallons per day was greater than the cost of providing water below that amount. Id at 16. Appellants argued that East Bay MUD should have applied an "incremental cost" approach--that is water rates should have been reasonably allocated among customer groups reflecting the costs of water services. Id at 21. Moreover, appellants believed that the inclining rate structure unreasonably discriminated against customers in hot-climate areas because they were forced to use more than the allotted amount of water and therefore had to pay the higher block rates. Id at 15.

East Bay MUD successfully argued that the adopted rate structure was reasonable. First, it argued that the 15 percent mandatory conservation goal was imposed because its records indicated that voluntary conservation (several months before the said adoption) was below that level. Appellee's Brief at 4. Second, it argued that the adopted rates reflected cost of service in the sense that the total revenue collected did not exceed the overall cost of providing water services to customers in the district and that industry accepted methods were used to calculate rates based on an "embedded cost approach," where "[t]he rates were calculated to generate the revenue required to meet the District's annual maintenance and operation expenses and the rate-funded capital costs and debt service." Id at 30-31. Finally, East Bay MUD argued that the inclining rate structure did not unlawfully discriminate against water users because the rates uniformly affected all single-family residential customers and that it uses a reasonable basis to classify its customers into categories, such as single family residential and multi-family residential. Id at 35-37.

- **The Special Tax Argument**

The Petitioners further argued that because the inclining rate structure was not based on the reasonable and actual cost of providing water services and therefore resulted in excess charges, such a rate structure constituted a "special tax" which must be approved by a two-thirds majority of the District's electors. Appellants' Brief at 24. In arguing that the inclining rate structure was a "special tax," the Petitioners had to show that the rate bore no relationship to cost of provided service and that East Bay MUD is a "special district." Id.

East Bay MUD successfully argued that the adopted rate structure was not a special tax because charges were tied to the receipt of water services and that the rate structure was designed to help customers lower consumption and avoid paying penalties. Appellee's Brief at 16-20. Finally, East Bay MUD argued that it is not a "special district." Id at 27. The district court ruled in favor of East Bay MUD. The case is currently being appealed.

5.2.4 Powers of Specific Types of Publicly-Owned Water Agencies

While there are dozens of legal forms of water supply agencies in the state of California, we restrict the following review to the three most common publicly-owned water suppliers in urban areas. These are County Water Districts, Municipal Utilities Districts, and Municipalities. This review examines the ratemaking authority granted to these governmental entities under statutory law prior to the amendment of Water Code section 375. These governmental entities appear to have enjoyed rather broad ratemaking powers even prior to the amendment of section 375. The amendment codifies their powers with respect to setting rates to manage water shortages and promote conservation within their service territories.

- **County Water Districts**

As previously mentioned, Water Code section 31007 authorizes County Water Districts to "collect charges sufficient to ... provide for repairs and depreciation of works owned or operated by the district and to provide a fund to pay principle on bonded debt as it becomes due." Thus, section 31007 authorizes County Water Districts to establish reserve funds to cover past debt and to provide for the replacement of existing works. Within this context, these districts are given discretion to design rates that charge in excess of annual expenditures, with said excess to be applied to debt payment and facility replacement and expansion. Moreover, Arcade County Water District, 6 Cal.3d.232, suggests that actions taken by districts with respect to water rates will be presumed reasonable by the courts.

In addition, Water Code section 31024 authorizes County Water Districts to establish rules and regulations for the sale, distribution, and use of water. Under this section, it appears to be reasonable for a district to establish a conservation program, including conservation promoting rates, if it was deemed by the district's governing body to be necessary for regulating the sale, distribution and use of water.

- **Municipal Utilities Districts**

Water Code section 71616 authorizes Municipal Utilities Districts to set rates to "raise revenues sufficient to ... provide for repairs and depreciation of works ..." and to "provide a reasonable surplus for improvements, extensions, and enlargements." Thus, under section 71616, Municipal Utilities Districts are clearly authorized to collect revenues in excess of expenditures and to establish surplus funds for system improvements, which could be interpreted to include conservation and other water saving or reclaiming projects. In support of this, Water Code Section 71610.5 authorizes a district to "undertake a water conservation program to reduce water use and may require as a condition of new service that reasonable water saving devices and water reclamation devices be installed to reduce water use." What constitutes a conservation program is not specified, but it appears that Municipal Utilities Districts could justify the use of conservation promoting rates and the application of surplus revenues to water saving activities under this section of the Water Code.

- **Municipalities**

Both general law and chartered cities are given broad discretion in setting water rates. The operation of water works by general law cities is governed under Government Code sections 38730 and 38742. These sections authorize a general law city to "acquire, construct, repair, and manage ... all physical works needed to supply its inhabitants with water and also to sell the water from such a system." Under Article IX, Section 7, of the Constitution, a general law city is empowered to enforce within its limits local laws not in conflict with general laws and thus is accorded discretion in setting water use and rate ordinances. Chartered cities are governed by Constitutional provisions only and thus the design of water use and rate ordinances also is left to the city's governing body. Moreover, Hansen v. City of San Buenaventura, 42 Cal.3d.1172, reaffirmed that water supply municipalities are entitled to a fair rate of return on their investment in water supply infrastructure, and are entitled to recoup from their customers the cost of maintaining and running the water supply system. In essence, a municipality is authorized to operate the city's water system as its governing body deems reasonable and beneficial to the city. The design of rates and collection of excess revenues appear to lie within these broad powers. For this reason, it has been noted that of the three legal forms discussed here, municipalities have the greatest flexibility in the design of rates and recovery of costs.

- **Publicly-Owned Water Supply Agencies - Summary**

Statutory and case law does not appear to preclude publicly-owned urban water agencies from adopting a wide array of conservation promoting rate structures. Particularly in light of the recent amendment of Water Code section 375, the use of any of the conservation promoting rate structures to be discussed in Chapter 4, including increasing block rates, seasonal rates, excess use surcharges, and value-of-service based rates, appears to be legally defensible under existing law. In addition, the law accords publicly-owned urban water agencies discretion to establish surplus funds and collect revenue in excess of annual expenditure for purposes of system management and improvement. Thus, the use of revenue stabilizing funds, also to be discussed in Chapter 4, to even out the collection of revenue, particularly during drought, appears to be justified under existing law.

5.3 Ratemaking Powers of Investor-Owned Water Agencies

In general, we find that under existing Public Utility Code the investor-owned water utilities are accorded very little discretion in setting rates and recovering costs, and that this could greatly impair their ability to adopt conservation promoting water rates and their willingness to invest in conservation programs.

5.3.1 Mutual Water Companies

The one exception to this general finding is with respect to mutual water companies which are not subject to rate making regulation. A mutual water company, as defined in P.U. Code section 2725, is a private water company that sells water only to its shareholders. As long as a mutual water company does not furnish water for public use, i.e., deliver water to non-shareholders, it is not subject to the regulations of the Public Utilities Commission (P.U. Code section 2705). With respect to urban water agencies, however, mutual water companies represent a minute share of deliveries, mostly to small, remote communities.

5.3.2 Investor-Owned Water Agencies Supplying Water For Public Use

The more typical case is the investor-owned water utility that supplies water to non-shareholders. As a private utility providing water for public consumption, it is subject to traditional rate of return regulation and comes under the jurisdiction of the Commission. To initiate a change in rates, Class A investor-owned utilities (those with 10,000 or more customers) must file for a rate review with the Commission. In a general rate case, the utility's application for rate relief will be investigated and possibly contested by Commission staff, and then resolved through the Commission's hearing and decision processes. Class A utilities may file for rate relief every three years.

Investor-owned utilities with fewer than 10,000 customers (e.g., Class B/C/D) may go through an abbreviated, less formal process for rate relief. These utilities can petition the Commission for rate relief by filing an "advice letter" supported by a set of standardized supporting documents. Formal hearings are generally avoided (Assembly Utilities and Commerce Committe 1992).

As to the level and structure of rates, P.U. Code section 451 charges that rates must be "just and reasonable." The Commission has interpreted this to mean that rates should be based on cost-of-service and designed to recoup operating expenses, taxes, depreciation, and a return on net investment. With its authority to permit or disallow the recovery of certain costs, and its ability to dictate how costs can be recovered through rates, the Commission greatly limits the discretion with which investor-owned utilities can set rates.

The Commission has recently issued an Order Instituting Investigation ("Risk OII") to determine whether regulatory changes are needed to address financial and operational risks faced by investor-owned water utilities. The Commission has issued several decisions addressing issues affecting smaller water utilities (Class B/C/D) and will subsequently address issues affecting Class A companies.[50] Below we discuss specific regulatory issues that directly affect the ability of investor-owned utilities to adopt some of the conservation promoting rate structures presented in Chapter 4.

5.3.3 Adjusting Rates To Limit Demand

As already discussed, investor-owned water utilities under the jurisdiction of the Commission must gain Commission approval prior to adjusting their rates. Moreover, they are limited to applying for rate relief once every three years. These regulatory provisions would seem to prevent, or at least make difficult, the adoption of a rate schedule with automatic price adjustments to limit demand when certain supply constraints are reached (such as LADWP recently adopted), or the use of other price mechanisms that allow for a quick response to a water supply emergency. A recent survey of Commission ratemaking practices for water utilities found that in California the average length of time to complete a rate case was six to seven months for small systems and seven to nine months for large systems (National

---

[50] Decisions issued by the Commission for Class B/C/D utilities as a result of the Risk OII include the following:

1. Class C/D utilities may petition for annual cost-of-living rate adjustments. D.92-03-093 at 33-35.

2. Class C/D utilities are authorized to establish memorandum accounts to record and collect "unanticipated repair expenses." Id at 39-41.

3. The range of return for Class C and D utilities was increased from 11.6% to 12.1% and from 13.9% to 14.4%, respectively. Id at 45. The range of return for Class B utilities will continue to be dealt with on a case by case basis. Id at 46-47.

4. Class D utilities may recover up to 100% of their fixed costs with a service charge while class C utilities may recover up to 65% of their fixed costs. Class A/B utilities are only permitted to recover up to 50% of their fixed costs. Id at 47.

5. An operating ratio method can be used simultaneously with the return on net investment method in the calculation of rates for Class C/D utilities. The operating ratio method calculates "a margin over operating and maintenance expenses." Whichever method results in a higher return may be used. Id at 57-58.

Regulatory Research Institute 1992). During the most recent drought, the Commission granted several utilities the right to ration water use and assess a penalty for excessive use. However, in order to gain this relief, a utility must be able to show the existence of an imminent water emergency (D. 92-05-032), by which time it may be well into its water shortage.

5.3.4 Revenue Stabilizing Funds

Existing Public Utility Commission regulations also seem to prevent investor-owned water utilities from instituting revenue stabilization funds. A revenue stabilization fund allows a utility to ration demand without being pressured to recover the attendant revenue loss through rate increases during the rationing period. For the fund to work, the utility must be able to collect revenues in excess of annual expenditures in some years (e.g. when water supplies are normal or high) so that it can collect revenue less than annual expenditures during years when rationing is necessary. The fund is used to limit the risk of "rate shock" to customers asked to pay more for less water in order to meet the utility's revenue requirement. The cost-of-service criteria under P.U. Code section 451 would seem to prevent the use of this type of fund in most circumstances. During the most recent drought, utilities were allowed to establish memorandum accounts to track fluctuations in expenses and revenues as a result of reduced water sales and higher costs associated with conservation measures. D.90-07-067. In theory, they could then use these accounts to file for rate relief to recover drought costs. However, there remains considerable uncertainty regarding what costs the Commission will or will not allow to be recovered. Moreover, to the extent that the memorandum accounts were part of procedures adopted in reaction to the recent drought, the Commission has disclaimed that such procedures will be automatically applied to future droughts. D.92-09-084. Thus, for investor-owned utilities, considerable uncertainty surrounds the use of rates to recover drought costs.

5.3.5 Cost Recovery Risks and Conservation Investments

Because the Commission has the power to disallow cost recovery, utilities are reluctant to incur costs unless they are reasonably confident that the Commission will allow them to be recovered. This caution applies to conservation investments in particular, where precedents for recovery may not exist because of the newness of these types of investments. Thus, while the Commission encourages utilities to develop water management programs to address ways to manage water resources in the long-run (D. 90-08-055), the utilities themselves are understandably cautious in incurring potentially unrecoverable costs. Currently, the determination of cost allowance proceeds on a case-by-case basis. One way to reduce the uncertainty and encourage conservation investment would be for the Commission to establish a list of conservation practices for which the associated costs are recoverable. Such a list would allow investor-owned utilities to proceed more confidently with their conservation plans.

5.3.6 Recovery of Fixed Costs

Existing Commission regulations allow Class A utilities to recover up to 50 percent of their fixed costs through fixed charges and the remainder through volume charges. Smaller utilities can recover up to 100 percent of their fixed costs through fixed charges. Currently, Class A utilities are recommending that the Commission also allow them to recover up to 100 percent of their fixed costs through fixed charges. From the viewpoint of conservation, shifting cost recovery from the volume charge to a fixed charge sends the wrong message to the customer. The per unit cost of water becomes cheaper at a time when the cost of providing service is increasing. Class A utilities, however, are in a difficult situation. Because their ability to adjust rates, use surplus funds, or use balancing accounts to stabilize revenue during rationing periods is very restricted, and because at least 50 percent of fixed costs must be recovered through volume charges, they frequently are unable to cover their fixed costs when demand drops off during a shortage. Again, from the point of view of promoting water conservation, allowing investor-owned utilities to adopt more flexible rate structures and other revenue stabilizing mechanisms would be preferable to achieving revenue sufficiency by loading most of the costs into a fixed charge that is invariant with water use.

5.3.7 Investor-Owned Water Utilities - Summary

The Public Utility Code appears to limit the ability of investor-owned utilities to adopt flexible rate structures that promote conservation. Moreover, strict cost-of-service criteria appear to prevent the use of surplus funds or balancing accounts to stabilize revenues in all but emergency situations. Certainly, under existing regulations, investor-owned water utilities are not currently able to factor these more flexible approaches to rate design and revenue stability into their water management plans. There is a possibility, however, that forthcoming decisions from the Commission's ongoing Risk OII will allow for more flexible approaches in the future.

5.4 Conclusion

At this time there exists an odd juxtaposition between the ratemaking powers of publicly-owned and investor-owned water agencies in California. While the two types of utility often serve side-by-side communities, they have very different abilities to manage demand, recover cost, and stabilize revenue. The California Water Code and case law offer publicly operated utilities broad discretion in choice of rate structure and use of revenue stabilizing funds. This is not the case for investor-owned utilities. They are subject to rate-of-return regulation, which strongly influences rate structure, cost recovery, and conservation investment decisions. As publicly-owned utilities adopt more progressive, conservation-oriented rate structures, investor-owned utilities will not be in a position to follow suit.

CHAPTER 6 PUBLIC INVOLVEMENT IN RATEMAKING

After all, even the most meticulously designed water rate structure is worthless unless it is politically feasible.

-- Betsy Reifsnider, *Public Participation in Rate Setting*

6.1 Introduction

By 1976 it was clear to Tucson Water Department that the city faced a growing dilemma: demand for water service was rapidly outpacing supply, particularly during peak summer months. Working with its rate consultants and the City Council (which had authority over water rates), the Water Department developed an ambitious capital expansion program to increase the city water system's capacity.

During the winter and spring of 1976, the Water Department developed a new rate structure to help finance this expansion. The new rates would sharply increase monthly water bills for most customers. The City Council held one public hearing addressing the need for new water rates and informing the public as to likely impacts. The transcripts of that meeting, however, reveal that the those in attendance did not understand why a rate increase was required, did not believe that Tucson had a water supply problem, and did not want to pay more for water. The Council and Water Department continued their water rate deliberations with little further public involvement. In the early part of June, the Water Department presented the new rate schedule to the City Council for approval. For the first time, the city would adopt cost-of-service based rates. This was going to affect the various classes of customers quite differently than in the past. The Council was told that bills would decrease -- possibly by as much as 40 percent -- for low volume customers, while bills would increase -- possibly by as much as 85 percent -- for high volume customers. What was not conveyed to the Council or what it failed to grasp was that any customer with monthly usage in excess of 6 ccf was going to pay more for water, perhaps a lot more. The average summer monthly usage for a residential account at that time was about 20 ccf. That essential point, however, eluded the Council, which adopted the rate schedule. The first bills calculated with the new rates were received during the summer, which, as Murphy's law would have it, was abnormally hot and dry.

Public outcry was sharp and immediate. Bills for many customers were dramatically higher than predicted by the Water Department -- some bills increased by as much as 400 percent. An unsuspecting public was caught by surprise. Moreover, the rates were so complex, involving multiple blocks and no less than eight different lift zones, that most customers were unable to recompute their bills to verify the charges. This added to their frustration and anger. Within two weeks of the first billing the public's wrath took shape in a recall petition for Council members that voted to adopt the new rates. In September, the recall petitions were filed with more than twice the necessary signatures.

In November, the city manager, in an attempt to quell the unrest, appointed a Citizens' Water Advisory Committee (CWAC) -- consisting of a cross-section of the community affected by the rate increase -- to review Tucson's water supply and rate issues. In January, the Council members that had voted for the new rates were soundly defeated by candidates that promised to "rollback" the rates to their previous levels. The new Council majority, however, in deference to prudence decided to wait until the CWAC completed its report on the city's water situation before acting. Within a month the CWAC released a report that endorsed, ironically, the previous Council's cost of service-based rate schedule and recommended an additional 10 percent rate increase.

At the same time, through the media and public meetings, the city's current water supply options and costs, its future needs, and the need to begin demand management programs to curtail peak summer use were brought to the attention of the public. Rather than "rollback" the rates, as they had promised to do, the new Council adopted the rate structure recommended by the CWAC. The revised rate ordinance essentially was identical to previous Council's except that rates were even higher. This time, however, the public understood and accepted the change.[51] Thus ended one of the most bizarre and ironic exercises in ratemaking history.

Its lesson is obvious. If the public does not perceive rates to be affordable or equitable, it will make its discontent be felt, and possibly derail the entire ratemaking process. Conversely, if the public is well informed, made to understand the issues at hand, and allowed access to the ratemaking process, they can provide a positive force for change to a more effective rate design. In this chapter, we consider public involvement in ratemaking.

6.2 Customer Perceptions and Rate Acceptability

Research from earlier droughts indicates that customer perception of the fairness or effectiveness of particular conservation policies -- including price rationing -- is a key factor in their acceptance (Bruvold 1979). Customers that believe they are being treated unfairly or that the conservation program is ineffective will be less likely to accept the program. On the other hand, those that perceive the program as fair or effective will be more likely to accept the

---

[51] The source for this story is Martin, et al. (1984).

program. The same relationship holds true for rate structure policies, as Tucson's experience in the mid-1970's demonstrated. As we discussed in Chapter 2, the principle concerns about a water rate for most customers are (1) is it affordable and (2) is it equitable. Since, for most customers water is generally affordable, the focal point often becomes is the rate equitable. There are many possible ways to define what is equitable, but in the case of utility ratemaking, the issue of equity is usually connected to the issues of cross-subsidy -- i.e., how much of my water bill is recovering the cost of somebody else's consumption -- and of cost recovery -- i.e., is the utility collecting excessive revenue. If customers perceive that a rate structure will result in either of these outcomes, they will not accept it. The case of Brydon v. East Bay Municipal Utility District, discussed in Chapter 5, is an instance where a segment of customers felt that the utility's rate structure resulted in both cross-subsidy and unreasonable revenue collection. Los Angeles Department of Water and Power is facing similar challenges with respect to its new rates.

6.3 The Vocal Minority

Given the diversity of uses within and across customer classes for most urban water utilities, it may always be the case that a shift in rates will create both winners and losers. Water bills will increase for some and decrease for others. For those that receive an increase there is likely to be dissent, particularly if the increase is large or if they sense that they are being singled out. For the purposes of this discussion, we refer to this group as the vocal minority. The vocal minority plays an important role in the ratemaking process in two ways. First, it may become a blocking coalition, and attempt to preserve the status quo by preventing the adoption of or significantly modifying a proposed rate structure. Second, it can be for the utility what the canary was for the coal miner; alerting the utility to problems with the rate structure that were not initially considered or recognized.

The vocal minority can be particularly effective as a blocking coalition because its size and membership allow it to make its needs and desires more keenly felt by decision makers than can the "silent" majority. Its relatively small size lowers its organizational costs. More importantly, the risk of losing income discourages free-ridership and encourages group cohesiveness. In contrast, group size and potential for free-ridership make it very difficult for the silent majority to organize itself as a countervailing force to the vocal minority. Thus, the vocal minority may effectively prevent a change to a new rate structure even if the utility's customers as a whole would be made better off by the change. Conservation-oriented rates designed to discourage some type of discretionary use may be particularly susceptible to actions by a vocal minority. Those significantly vested in the use in question are likely to resist the change.

The vocal minority can be a particularly effective canary as well. Given the complexities of rate design and the heterogeneity of customers, it is always possible that a new rate structure will affect some customers in unintended and undesirable ways. Customers will certainly alert the utility if this occurs thus providing it with important information necessary to modify the rate structure. In this way, the vocal minority can help the utility to refine and improve its rates.

6.4 Providing Information

As the experience of Tucson illustrated, customers' are much more likely to accept a rate schedule if they understand its justification and believe that it accurately reflects the cost of providing service. In other words, customers need information that explains the rationale and cost basis for a rate schedule. Such information does not need to be extremely technical, but it should address basic questions that are certain to be raised by customers. For example, a utility instituting a seasonal rate for the first time might consider providing information to address questions of the following sort:

- Why are rates higher in the summer than the winter?
- If rates are cost based, why does it cost more to provide service in the summer than in the winter?
- We use the same amount of water in the summer as in the winter, why should we have to pay a higher rate in the summer?
- The hot, dry climate requires that we use more water during the summer to keep the landscape healthy. Is it fair to charge us extra for this?

Providing satisfactory answers to these type of questions is useful in at least two ways. First it reduces the likelihood that a vocal minority will form a blocking coalition. As previously mentioned, research indicates that customers generally do not oppose higher rates or restricted water use, per se, but rather resist if they perceive the new regime to be unjustified. It is in the interest of the utility, therefore, to provide to its customers accurate information explaining and justifying the rate change. Second, anticipating and addressing the concerns of customers acts like a system of checks and balances. If the utility finds itself unable to satisfactorily answer valid questions about the fairness or efficiency of a rate schedule, that would be a strong indication for the need to reexamine the rate design.

6.5 Gathering Input from the Public

Experience suggests that it is best to involve the public early-on in the ratemaking process, particularly if the new rate structure is to depart significantly from the old one. Without this involvement, the risk that the public will balk at the new rates is greatly increased. Public involvement can take many forms, including public hearings, surveys or polls, and advisory committees.

Public hearings are the most common method, and offer the public an opportunity to address to those with authority over rates their issues, concerns, and questions about a rate proposal. Public hearings, however, only offer to the public an opportunity to react. For rate changes that promise a significant departure from the past, it often can be advantageous to have representatives of the public play a more proactive role in the ratemaking process. It is much less likely that the public will react negatively, at least en mass, to a rate design when they themselves participated in its development and selection. Of course, public participation can only be extended so far. It is likely to always be the case that some customers will object to a new rate design. However, citizens committees have proven an effective way to build broad support for a new rate structure.

The appointment of a Citizens' Advisory Committee representing the diversity of interests and viewpoints of the utility's customers is an increasingly common way to involve the public in rate setting. Such committees can help foster understanding, cooperation, compromise, and consensus among the various and often conflicting interests seated around the ratemaking table. The benefits of the Los Angeles Mayor's Blue Ribbon Committee for Water Rates were described by one of its members as follows:

> The results of this long, inclusive process were profound. Members of the coalition developed a strong working relationship with one another, balancing the sometimes contradictory goals they had set for themselves. Moreover, this coalition of citizens and water agency professionals produced a rate structure which each member had helped to design; each member understood the rate proposal, could explain it to his/her constituency, and could lobby for its passage. (Reifsnider 1993)

6.6 Conclusions

"No rate schedule can be successful without positive public perception and understanding. The importance of this particular aspect of rate structure establishment cannot be overstated." So concluded the Southern California Water Committee's *Conservation Through Rate Structure Task Force Report*. Public involvement is a key element to any ratemaking process. Keeping customers advised of impending rate changes, sending out information explaining rates and bill formats, and involving the public in rate proposals through hearings and public advisory committees will foster understanding and consensus that lead to a successful ratemaking process.

REFERENCES

Agthe, D., R. Billings, et al. (1986). "A Simultaneous Equation Demand Model for Block Rates." Water Resources Research 22(1):

Alvarez, F. (1993). Oxnard May Reduce Rates of Big Water Users to Lure Businesses. Los Angeles Times. Los Angeles, B6.

American Water Works Association (1979). Managing Water Rates and Finances. American Water Works Association.

American Water Works Association (1983). Water Rates. American Water Works Association.

American Water Works Association (1992). Manual of Water Supply Practices: Alternative Rates. American Water Works Association.

Arthur Young (1988). 1988 National Water and Wastewater Rate Survey.

Assembly Utilities and Commerce Committe (1992). Investor-Owned Water Utilities: Rates, Reliability, and Regulation. Assembly Utilities and Commerce Committee.

Baumann, D. D. (1986). Issues in Water Pricing. Water Pricing and Water Demand . Arizona Corporation Commission.

Beecher, J., J. Landers, et al. (1991). Integrated Resource Planning for Water Utilities. National Regulatory Research Institute.

Beecher, J. and N. Zearfoss (1992). 1992 NRRI Survey on Commission Ratemaking Practices for Water Utilities. National Regulatory Research Institute.

Beecher, J. A. (1990). "Value, Cost, and Price: Essay on Emerging Water Utility Issues." NRRI Quarterly Bulletin 11(2): 177-181.

Beecher, J. A. and A. P. Laubach (1989). Compendium on Water Supply, Drought and Conservation. The National Regulatory Research Institute.

Beecher, J. A., P. C. Mann, et al. (1990). Cost Allocation and Rate Design for Water Utilities. The National Regulatory Research Institute.

Beecher, J. A., P. C. Mann, et al. (1993). Meeting Water Utility Revenue Requirements: Financing and Ratemaking Alternatives. The National Regulatory Research Institute.

Bollman, F. H. and M. A. Meritt (1981). Community Response and Change in Residential Water Use to Conservation and Rationing Measures: A Case Study -- Marin Municipal Water District. Selected Works in Water Supply, Water Conservation and Water Quality Planning Eds. J. E. Crews and J. Tang. Fort Belvoir, VA, Institute for Water Resources, U.S. Army Corps of Engineers. 393.

Bonbright, J. C., A. L. Danielsen, et al. (1988). Principles of Public Utility Rates. Arlington, VA, Public Utilities Reports.

Brown, S. and D. Sibley (1986). The Theory of Public Utility Pricing. Cambridge, Cambridge University Press.

Bruvold, W. H. (1979). "Residential Response to Urban Drought in Central California." Water Resources Research **15**(6): 1297-1301.

Bureau of Domestic Commerce (1976). An Analysis of Price/Cost Sensitivity of Water Use in Selected Manufacturing Industries. U.S. Department of Commerce.

Carson, J. (1979). The Price Elasticity of Demand for Water. UCLA.

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

Chao, H.-p., S. S. Chen, et al. (1988). "Priority Service: Market Structure and Competition." The Energy Journal **9**(Special Issue on Electricity Reliability): 77-103.

Cuthbert, R. W. (1989). "Effectiveness of Conservation-Oriented Water Rates in Tuscon." American Water Works Association Journal **81**(33): 67-69.

Dasgupta, P. S. and G. M. Heal (1979). Economic Theory and Exhaustible Resources. Cambridge, Cambridge University Press.

Dawdy, D. R., L. D. James, et al. (1983). Demand Oriented Measures. Coping with Droughts Eds. V. Yevjevich, L. d. Cunha and E. Vlachos. Littleton, CO, Water Resources Publications.

Day, M. (1986). A Discussion of Empirical Evidence of the Conservation Impact of Water Rates. Water Pricing and Water Demand . Arizona Corporation Commissioner. 38.

Dziegielewski, B., E. Opitz, et al. (1991). Municipal and Industrial Water Use in the Metropolitan Water District Service Area: Interim Report No. 4. Prepared by Planning and Management Consultants, Ltd. for Metropolitan Water District of Southern California.

Electric Power Research Institute Priority Service Methods: A Review of Concepts and Status of Research.

Ernst & Young (1990). Ernst & Young's 1990 National Water and Waste Water Rate Survey. National Environmental Consulting Group, Ernst & Young.

Farnkopf, J. and S. S. Seeley (1993). Allocating the Cost of Reclaimed Water to Water and Waste Water Rate Payers.

Feldman, S., J. Breese, et al. (1981). "The Search for Equity and Efficiency in the Pricing of A Public Service: Urban Water." Economic Geography **57**(January): 78-92.

Feldman, S. L., R. Obeiter, et al. (1980). An Operational Approach to Estimating Marginal Costs of Urban Water Supply with Illustrative Applications. Wisconsin Public Service Commission.

Gewe, G. (1992). Addressing Engineers' Concerns for Marginal Cost Pricing in Urban Markets. Los Angeles Department of Water and Power.

Goldstein, J. (1986). "Full-Cost Water Pricing." American Water Works Association Journal **78**(2): 52-61.

Hall, D. (1993). Water Rate Design Based on Marginal Cost. Mayor's Blue Ribbon Committee.

Hanemann, W. M. (1992). Marginal Cost Pricing & The New LADWP Water Rates. Mayor's Blue Ribbon Committee on Water Rates.

Hirshleifer, J. (1958). "Peak Loads and Efficient Pricing: Comment." Quarterly Journal of Economics **72**(August): 451-62.

Hirshleifer, J., J. C. Dehaven, et al. (1960). Water Supply: Economics, Technology, and Policy. Chicago, IL, University of Chicago Press.

Hogarty, T. F. and R. J. Mackey (1975). "The Impact of Large Temporary Rate Changes on Residential Water Use." Water Resources Research **11**(6): 791-94.

Howe, C. and F. P. Linaweaver (1967). "The Impact of Price on Residential Water Demand and its Relationship to System Design and Price Structure." Water Resources Research **First Quarter**: 13-32.

Immerman, F. W. (1987). Financial Descriptive Summary: 1986 Survey of Community Water Systems. Office of Drinking Water, U.S. Environmental Protection Agency.

Kahn, A. (1988). The Economics of Regulation: Principles and Institutions. Cambridge, MA, The MIT Press.

Linaweaver, F. P. and J. C. Geyer (1964). "Use of Peak Demands in Determination of Residential Rates." American Water Works Association Journal **56**(April):

Mann, P. (1993). Water-Utility Regulation: Rates and Cost Recovery. Reason Foundation.

Mann, P., R. Saunders, et al. (1980). "A Note on Capital Indivisibility and the Definition of Marginal Cost." Water Resources Research **16**(3): 602-4.

Mann, P. and D. Schlenger (1982). "Marginal Cost and Seasonal Pricing of Water Service." American Water Works Association Journal **74**(1): 6-11.

Mann, P. C. (1987). "Reform in Costing and Pricing Water." American Water Works Association Journal **79**(3): 43-45.

Martin, W., H. Ingram, et al. (1984). Saving Water in a Desert City. Washington, D.C. : [Baltimore], Resources for the Future ; Distributed by the Johns Hopkins University Press.

Martin, W. E. (1988). Water Price as a Policy Variable in Managing Urban Water Use: Tucson, Arizona. University of Arizona.

Mercer, L. and W. D. Morgan (1986). "The Efficiency of Water Pricing: A Rate of Return Analysis for Municipal Water Departments." Water Resources Bulletin **22**(2): 289-295.

Metropolitan Water District of Southern California (1991). Water Conservation Pricing Approaches of the Metropolitan Water District.

Moncur, J. and Y.-S. Fok (1993). Water Pricing and Cost Data: Getting the Right Numbers. University of Hawaii Water Resources Research Center.

Moncur, J. and R. Pollock (1988). "Scarcity Rents for Water: A Valuation and Pricing Model." Land Economics **64**(1): 62-72.

Narayanan, R., H. Beladi, et al. (1987). "Feasibility of Seasonal Water Pricing Considering Metering Costs." Water Resources Bulletin **23**(6): 1091-1099.

National Association of Regulatory Utility Commissioners (1989). NARUC Annual Report on Utility and Carrier Regulation 1988. Washington, DC, National Association of Regulatory Utility Commissioners.

National Regulatory Research Institute (1992). 1992 NRRI Survey on Commission Ratemaking Practices for Water Utilities. National Regulatory Research Institute.

Nieh, B. (1979). A Probabilistic Optimization Model for Water Resources Capacity Expansion with Price Sensitive Seasonal Demand. UCLA.

Raftelis, G. A. (1993). Comprehensive Guide to Water and Wastewater Finance and Pricing. Ann Arbor, Lewis Publishers.

Reifsnider, B. (1993). Coalition Building for Marginal Cost Pricing. Mono Lake Committee.

Renshaw, E. F. (1982). "Conserving Water Through Pricing." American Water Works Association Journal **74**(1): 5.

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

Rice, I. M. and L. G. Shaw (1980). Water Conservation -- A Practical Approach. Water Conservation Strategies . Denver, CO, American Water Works Association. 73.

Rohde, R. (1976). Potential Legal Constraints on Price Conservation Programs. Department of Water Resources.

Russell, J. D. (1983). Seasonal and Time of Day Pricing. Water Rates: An Equitability Challenge . Denver, CO, American Water Works Association. 96.

Saunders, R., J. Warford, et al. (1976). The Definition and Role of Marginal Cost in Public Utility Pricing: Problems of Application in the Water Supply Sector. International Bank for Reconstruction and Development.

Saunders, R., J. Warford, et al. (1977). Alternative Concepts of Marginal Cost for Public Utility Pricing: Problems of Application in the Water Supply Sector. World Bank.

Sax, J., R. Abrams, et al. (1991). Legal Control of Water Resources: Case Studies and Materials. St. Paul, Minn., West Publishing Co.

Southern California Water Committee (1993). Conservation Through Rate Structure.

Steiner, P. O. (1957). "Peak Loads and Efficient Pricing." Quarterly Journal of Economics **71**(November): 585-610.

Stevens, T. H., J. Miller, et al. (1992). "Effects of Price Structure on Residential Water Demand." Water Resources Bulletin **28**(4): 681-685.

Turvey, R. (1976). "Analyzing the Marginal Cost of Water Supply." Land Economics **52**(May): 158-68.

U. S. Department of Commerce (1976). An Analysis of Price/Cost Sensitivity of Water Use in Selected Manufacturing Industries. Bureau of Domestic Commerce.

Vickrey, W. (1948). "Some Objections to Marginal Cost Pricing." Journal of Political Economy **56**(June): 218-238.

Wallace, L. P. (1987). Water and Revenue Losses: Unaccounted for Water. American Water Works Research Foundation.

Water Resources Research Institute of The University of North Carolina (1978). North Carolina Workshop on Water and Wastewater Pricing. North Carolina Workshop on Water and Wastewater Pricing, Water Resources Research Institute of The University of North Carolina.

Woodard, G. C. (1986). A Summary of Research on Municipal Water Demand and Conservation Methodologies. Water Pricing and Water Demand . Phoenix, AZ, Arizona Corporation Commission. 43-47.