Ex. AA-D-30

<u>CHARGING FOR DISTRIBUTION UTILITY SERVICES: ISSUES IN RATE DESIGN</u>

December 2000

Frederick Weston

with assistance from:
Cheryl Harrington
David Moskovitz
Wayne Shirley
Richard Cowart

The Regulatory Assistance Project
16 State Street, Montpelier, VT 05602
Phone: 802-223-8199
Fax: 802-223-8172
rapvermont@aol.com

ACKNOWLEDGMENTS

This paper could not have been written without the guidance, insights, and thoughtful comments of many people. The review given it by members and staff of NARUC's Committee on Energy Resources and the Environment was invaluable. Many thanks go to Bob Anderson, Denis Bergeron, Annunciata Marino, Bill Nugent, Ann Thompson, and Jim Gallagher.

Many others in the electric utility, regulatory, consumer advocacy, distributed resource, and environmental protection communities offered their time and expertise to the effort. I am very grateful to J. Riley Allen, Scott Anderson, Shimon Awerbuch, Bruce Biewald, Jim Cater, Bill Deehan, Ralph Cavanagh, Paul Chernick, Ennis Gidney, Ron Lehr, Bill Marcus, Tom Starrs, Margaret Suozzo, and Roger Sutton. Thanks also to my colleagues at RAP and to Eric Heitz and Katie McCormack of the Energy Foundation. Especial thanks go to Jim Lazar and John Stutz, who gave the report more of their time and energy than I had any right to expect: it is all the better for it.

Lastly, this report was prepared by the Regulatory Assistance Project for the National Association of Regulatory Utility Commissioners, under a grant from the Energy Foundation. The views and opinions expressed herein are strictly those of the authors and may not necessarily agree with, state, or reflect the positions of NARUC, the Energy Foundation, or those who commented on the paper during its drafting.

Thanks to all,

Rick Weston

TABLE OF CONTENTS

EXECUTIVE SUMMARY

The world of electricity is changing. Competition in generation and retail services is being introduced in states across the nation, and in countries around the globe. The responses have been swift and far-reaching: a proliferation of new industry players, innovations in technologies and product offerings, and institutional reforms to support the evolving markets and preserve the benefits of the old system. Even those components of the sector that appear to retain their monopolistic character have been affected.

Such is distribution. Still seen as an essential network and natural monopoly, it has quickly become a focus of new scrutiny. In states with competitively restructured markets, the wires network remains the sole regulated asset, but it too is more and more facing competitive pressure. Shifting patterns of demand and alternative technologies, especially distributed resources (end-use efficiency and small-scale, dispersed generation), have begun to change the uses of the lower voltage network. What it will look like, how it will function, in twenty years is anyone's guess. "We are not," so to speak, "in Kansas anymore."

This transformation has many ramifications and it poses unique challenges for policymakers. One in particular – how in the midst of structural change should prices for distribution services be set so as to promote economic efficiency, vibrant competition, fairness, and environmental protection? – is the subject of this report. The distribution network is no longer the seemingly static monopoly that it once was. The policies that regulators adopt should be devised with an eye to competitive service provision, to encourage innovative and environmentally sustainable energy use. They should not shortsightedly protect a *status quo* that, over the coming decades, will not be well-suited to the economy it serves. Pricing policy will have direct impacts on how that future is met.

In this report, we evaluate how rate structures are and should be set for electric distribution services, today and as the industry becomes increasingly competitive. We examine the various methods for determining distribution costs – embedded and marginal – and consider whether there are distinct features of the system that call for particular approaches to rate design. We then test the several approaches against long-standing principles of rate design and the characteristics of prices in competitive markets. This is not, however, merely an academic exercise; also involved are important practical considerations. Our intention is to offer regulators useful and uncomplicated guidance when designing rates for distribution services.

Our findings and conclusions can be summarized as follows:

- In a vertically integrated, monopoly regulated industry, customers have historically paid for generation, transmission, and distribution in bundled usage-based (per unit) prices. For lower-volume consumers – residences and small businesses, these prices have typically been volumetric, that is, energy (per kilowatt-hour) prices. For higher-volume purchasers, rates have often been broken into two parts – energy and demand (per kilowatt). Except for modest monthly customer charges intended primarily to cover billing and metering costs, utility rate structures have been such that changes in customer consumption were accompanied by corresponding changes in customer bills: which is to say, the more (or less) a customer consumed the more (or less) he or she paid.

- Efforts to restructure the electric industry, to create competitive markets for generation and retail services, have, in a sense, "uncovered" the distribution system and have encouraged utilities, consumer advocates, and regulators to re-examine pricing policies for what appears to remain a naturally monopolistic component of the industry. (The same can be said of transmission; however, our focus is on that part of the system that falls primarily under state jurisdiction.) The distribution network, which typically had accounted for anywhere from ten to forty percent of a vertically-integrated utility's total investment, has thus become the object of central concern to firms that no longer own generation assets. Recent utility proposals to restructure distribution rates can be seen as business strategies to mitigate risk and increase revenues, understandable in themselves, but not necessarily consistent with the overall public interest.

- It is not enough to assert a principle of economics to justify a particular rate design. Economic efficiency is an important consideration when structuring rates, but it is by no means the only one, or even the foremost. Fairness, rate stability, revenue stability, administrability, non-discrimination, and environmental protection are equally significant, and regulators often have to find ways to reconcile these sometimes competing goals.

- The potential effects of a rate design must also be weighed. Will it induce economically efficient behavior by both the utility and its customers? Will it promote societally least-cost production and consumption? How will it affect customers' costs for energy services? How does it shift revenue burdens among customer classes? What impacts will it have on company revenues? How does it affect the allocation of risk between customers and the utility? Always be aware of the revenue effects of a particular rate structure. Who benefits, who loses? Here the admonition to be practical cannot be stressed enough. Seemingly small changes in a rate design can have very significant consequences for different customers.

- The usefulness of cost analyses of the distribution system in designing rate structures and setting rate levels depends in large measure upon the manner in which the studies are undertaken. Cost studies (both marginal and embedded) are intended, among other things, to determine the nature and causes of costs, so that they can then be reformulated into rates that cost-causers can pay. Such studies must of necessity rely on a host of simplifying assumptions in order to produce workable results; this is especially true of embedded cost studies. Moreover, it is often the case that many of the costs (*e.g.*, administrative and general) that distribution rates recover are not caused by provision of distribution service, but are assigned to it arbitrarily. Too great a dependence on cost studies is to be captured by their underlying assumptions and methodological flaws. Utilities and commissions should be cautious before adopting a particular method on the basis of what may be a superficial appeal. More important, however, is the concern that a costing method, once adopted, becomes the predominant — and unchallenged — determinant of rate design.

- Distribution costs are driven primarily by demand, numbers of customers, and energy needs. This is true in both the short and long runs. Utilities are continually investing in distribution plant – new facilities, upgrades, and replacements – in response to changes in load.

- Since one object of regulation is to serve as a proxy for competition, to impose upon a single provider the disciplines of competitive markets, it is reasonable to consider the structure of prices in competition when pricing monopoly services. Two relevant facts emerge. The first is that goods and services in competition are invariably available and priced on a unit basis. And the second is that the extent to which more restrictive pricing schemes exist is a measure of the lack of competition in that particular market. In competition, a consumer who does not consume a product or service does not nevertheless pay for the mere *ability* to consume it. Thus, as a general matter, prices should be structured so that, if a consumer chooses not to purchase a good or service, he or she has no residual obligation to pay for some portion of the costs to provide that good or service. In this sense, from the consumer's perspective, costs should be "avoidable."

- Volumetric, energy-based unit prices for distribution services (equal to at least long-run marginal cost) are the preferred approach. This is particularly true for lower-volume consumers. Such rates promote long-run economic efficiency and are fair: they enable consumers to make cost-effective trade-offs between electricity consumption and alternative investments in distributed resources (and other energy sources) and they require that consumers pay only for the services they use. Time-of-use differentiations (peak and off-peak), reflecting the changing costs of delivery over time, would in many cases be a sensible variation on this approach.

- For larger volume customers, a multi-part price structure that differentiates between demand-related and energy-related costs will work, to the extent that, as with energy-only pricing, customers pay only for what they use and that, as their consumption changes, so do their bills. Fixed charges should not be "dressed up" as demand charges.

- Two often-cited arguments in favor of fixed rates are, one, that distribution costs are fixed and, two, that usage-based rates in excess of marginal cost will encourage customers to over-invest (*i.e.*, invest non-cost-effectively) in alternatives to service: end-use efficiency, distributed generation, alternative fuels, etc. Neither is persuasive. First, as already stated, distribution costs are not fixed: investment in distribution is constant and growing, and avoidable therefore. Second, there is very little evidence that consumers make uneconomic investments in efficiency, distributed generation, and alternative fuels even in those places where bundled retail rates for generation and delivery are very high (two or three times marginal cost). Other barriers to these alternative resources – lack of information, lack of capital, onerous interconnection requirements, etc. – all remain powerful inhibitors to consumer investment. Usage-based distribution rates will provide one incentive to investment in distributed resources (signaling that there are avoidable distribution costs), but by themselves they will not lead to *over*-investment.

- Fixed, recurring, unavoidable charges also violate certain principles of rate design. They do not necessarily promote economic efficiency, since they tell a consumer little about the costs that his or her consumption imposes on the system. This can lead to uneconomic consumption and degraded system reliability. Nor are fixed charges particularly fair, since consumers contribute equally to the utility's revenues, regardless of the level of their usage. Consequently, lower-volume and, in many cases, off-peak consumers would pay a disproportionate share of the network's costs. Those who make greater use of the network should bear a proportionately greater share of its costs. In addition, shifts from usage-based to fixed charges could have undesirable revenue impacts upon a company, either excessive losses or earnings, that would require perhaps politically unpalatable remedial actions. Revenue stability may be jeopardized, and public faith in the regulatory institutions threatened.

- Usage-based rate designs promote economic efficiency, fairness, environmental protection, and the deployment of distributed resources. Fixed charges, because they are unavoidable (except where a customer goes off-grid entirely), discourage cost-effective consumer demand responses and innovation, to which firms would likewise respond. The constant pressure for dynamic efficiency would be lost.

- Usage-based rate designs reward firms for increased sales. To the extent that such sales are economically inefficient and environmentally damaging, actions must be taken by regulators to remove the firm's incentive, to break the link between sales and profitability. A performance-based, per-customer revenue cap mechanism is a promising approach for doing so. It rewards a firm for increases in efficiency, while making it at the very least indifferent to the volume of throughput over its wires. To the utility, a per-customer revenue cap would produce revenues in just the way that fixed recurring charges would; however, the revenue cap enables prices for end-users to be set on a usage basis, thereby enabling them to make consumption decisions and alternative energy investments that are, in the longer term, most efficient.

In sum, we urge regulators to adopt pricing and rate-setting policies that will serve the longer-term public interests: fairness, economic efficiency, competitive provision and innovation, and environmental protection. In the distribution system, this calls for usage-based pricing – primarily volumetric (energy-based) but also, where appropriate, demand- and energy-based. Additionally, we recommend that policymakers implement revenue-cap performance-based regulatory schemes for distribution companies. In so doing, the firms would obtain much of the benefit that they seek through fixed pricing, while consumers and the economy overall would still retain the benefits that usage-based, competitive pricing provide. We note, however, that a revenue cap or similar mechanism should not be seen as a necessary prerequisite to the usage-based pricing structure.

I. INTRODUCTION

All regulators understand the importance of rate design. So do customers. Even those who have never given it a moment's consideration implicitly understand that there is a direct relationship between the price of a good and their willingness to purchase it. Price – both its level and its form – is a powerful determinant of consumer behavior. Accordingly, the setting and design of rates is one of the regulator's most effective means by which to achieve desired policy objectives.[1]

Rate design for distribution utilities is emerging as a critical issue in electric industry restructuring. The primary revenue-producing activity of companies that no longer bear the financial and business risks of generation is the delivery of electricity to end-users; accordingly, the pricing and recovery of their transmission and distribution costs has fast become a locus of debate among utility officials, consumer advocates, and regulators. Distribution investment can make up anywhere between ten and forty percent of a vertically-integrated utility's costs, depending on the demographic, geographic, and other cost (in particular, generation) characteristics of the company.[2] But today, with distribution more and more divested of generation, distribution-only companies now recognize the sensitivity of their profits to changes in through-put – in other words, the provision of distribution services now constitutes most, if not all, of their business risk – and they have begun to develop rate designs aimed at mitigating those risks. However, those rate designs will also have other effects (intended and not) of which regulators should be aware. In the coming months and years, the decisions about how those costs will be reflected and recovered in rates will have significant impacts on the overall economic efficiency and environmental sustainability of the electric industry.

All this is complicated by the fact that, today, in all the states that have introduced some form of retail choice, customers also pay wires charges for the "stranded" costs of their utilities' historical generation investments and for public benefits — *e.g.*, the energy efficiency programs, renewable energy resources, and research and development activities — that the previously integrated system once provided. Stranded cost charges can amount to up to twenty percent of a customer's total bill, while public benefits charges typically total no more than four percent.

Two broad views are coming into focus. The first holds that the costs of the electric grid, across wide ranges of demand, are not especially usage-sensitive, and therefore it is sensible to employ rate structures heavily weighted to high fixed recurring charges and low usage-based rates. In the minds of some advocates, there are precedents for this kind of pricing scheme in other regulated

1. By "rate design," we mean the *structure* of prices, that is, the form and periodicity of prices for the various services offered by a regulated company. The two broad categories of pricing are usage charges and fixed, recurring (non-usage-sensitive) charges.

2. The ranges vary more widely by customer class. In low-cost systems in a number of places (Washington, Oregon, Idaho, Montana, Nebraska, and parts of Canada) distribution typically makes up fifty percent of low-volume customers' costs.

industries — for instance, telecommunications, and cable — as well as in some competitive ones, such as Internet services.

The second camp maintains that distribution (and transmission) costs do in fact vary with usage, and they point to, among other things, the effects of congestion on prices at peak times. They go on to argue that there are sound economic reasons – going to efficiency and fairness – for why regulators have traditionally rejected the notion of high, fixed recurring charge structures in favor of consumption-based prices. They contend that customers should understand what the incremental costs of their consumption decisions are so that resources will be allocated to their most highly valued uses. Also, they assert that customers should pay the costs that they impose upon the system, so that their consumption neither subsidizes nor is subsidized by the consumption of others. Usage-based rate designs, these advocates say, serve, and have served, these long-held objectives very well.

Also, there is the argument that, regardless of the degree of a product's short-run cost sensitivity to usage, as a general matter, competitive markets price goods on a usage basis and it is appropriate for regulation to mimic that approach. Moreover, advocates of this position point out that such pricing schemes are more consistent with other long-held principles of rate design: simplicity, administrability, and public acceptability.

This debate turns, in part, on an understanding of the economic principles of rate design. There is no dispute that, as a matter of economic efficiency, rates should be set to equal the marginal costs of production, allowing customers to incur (or avoid) marginal costs as their demand increases (or decreases) and thereby promoting the overall efficiency with which energy is used in our society.

There is dispute, however, about how to measure marginal costs (in fact, about whether they are even meaningful, in either the short or long run) and how to best reflect them in prices. Advocates of high fixed recurring charges cite the advantages they provide — surety of distribution company revenues, known costs for consumers, simplicity, to name a few — and the "unavoidability" of significant infrastructure costs. Those favoring usage-based pricing regimes assert that marginal costs are correlated to usage — that is, to demand — and it is therefore appropriate to recover seemingly fixed capital costs in volumetric charges. They point out that such pricing sends more rational signals to consumers about the value of their electricity use, and of alternatives to it, such as energy efficiency investments and alternative fuels.

The debate raises many questions. Will non-usage-based pricing result in the imposition of high, uneconomic "stand-by" (or back-up) charges for distributed resources? Will cost-effective energy efficiency investments be foregone, since, even though they will reduce costs, they will be less likely to reduce the bills that retail customers pay? Is it possible that the overall reliability of the electric system will be degraded, because usage charges set below their long-run marginal costs (particularly on peak) might encourage uneconomic over-consumption, which could in turn tax the transmission and distribution system beyond its capacity to maintain deliveries? Will fixed charges in fact produce a stable and reasonable level of distribution revenues? Is it a misapplication of marginal cost pricing principles to assert that, where incremental consumption also carries with it an incremental addition to unpriced environmental costs, the price signal should nevertheless be set to cover only the provider's marginal cost of production?

The economic characteristics of the regulated natural monopoly raise other important issues, notable among them the need to assure that rates set to reflect long-run marginal costs are "adequate" to recover the cost of service, but not more. This reconciliation imperative offers opportunities for the application of sound pricing policies in ways that will encourage the most efficient use of society's limited energy and environmental resources. Is it possible, for example, to reflect the incremental environmental costs of increased electricity generation in volumetric charges for distribution services, and to do so in a way that does not provide a financial windfall to the distribution company?

This report addresses these questions and suggests some simple policy responses to them. Although competition in electricity has raised new challenges, it has not altered the fundamental role that regulators must play in overseeing an industry greatly "affected with the public interest." The long-standing objectives for regulation still hold, and they provide useful guidance as we move into the new century. Decisions that regulators make today, with respect to what seem to be the most mundane of issues, will have important consequences in the years to come.

II. THE NEW DEBATE: PRICING UNBUNDLED DISTRIBUTION SERVICES

A. History

When in 1882 Thomas Edison first offered direct-current electricity for sale in downtown New York, to power the light bulbs that he was selling, he charged his customers not according to the quantity of electricity they consumed, but rather according to the amount of lighting service, or lamp-hours, they used. Although his lamp-hour charges were based in part on the costs of generating and delivering electricity (and in part on the costs of the alternative – gas lighting), he preferred his method of pricing, reasoning, first, that the public understood lighting but not kilowatt-hours ("webers" in those days) and, second, that the selling of end-use service, rather than the intermediate commodity, would give firms stronger profit incentives for innovation in supply. However, the proliferation of end-uses and electricity suppliers made it increasingly difficult for Edison to maintain his pricing scheme, and in 1898 he disappointedly adopted the "kilowatt-hour" as the basis of rates.[3]

In the century since then, the commodity has remained the fundamental unit of electricity sales. Users have typically paid for their services through energy rates, or a combination of energy and demand rates, and modest customer charges.[4] For the most part, these rates have been "bundled," which is to say that they have covered all the costs of the various components of electric service — generation, transmission, and distribution. As a general matter, the notion that customers should be charged according to their usage was not seriously challenged, and innovations in rate design were driven as much by a utility's revenue needs as by the prescriptions of economics. And, so long as costs continued to decline as output expanded, as utility profitability remained secure while prices declined, there was little reason to upset this state of affairs.

This traditional approach to pricing, which largely approximated how competitive markets set prices, enabled customer costs (bills) to vary with their usage (energy) and capacity needs (demand). Since customers, both as individuals and as members of classes sharing similar usage profiles, imposed particular costs upon the electric system, a structure of prices that had customers bearing, to the greatest extent possible, the costs they caused in relation to their usage was generally seen to best serve the twin goals of efficiency and fairness.

This is still true. Although the electric industry is now in the midst of a tremendous restructuring, driven in large part by significant technological changes that have just as profoundly altered the economics of production, the fundamental goals of regulation and, in particular, of rate design remain unchanged. Generation and retail services are becoming more

3. Neil, Charles E., "Entering the Seventh Decade of Electric Power," Edison Electric Institute Bulletin, September 1942.

4. There have, of course, been variations on these approaches, for example, customer charges that include some amount of "free" usage (say, 100 kWh/month). Bonbright, p. 347.

and more competitive, but there are still the delivery services — transmission and distribution — to be dealt with. How should these services, particularly distribution, be priced so that equity, efficiency, and environmental protection can be promoted in the electric industry?

As of July 2000, twenty-three states have restructured, or have set deadlines for restructuring, their electric industries. Generation and other retail services will be offered by competitive providers. Electricity will be delivered to end-users over monopoly-owned wires, the prices for which are set by the Federal Energy Regulatory Commission (FERC), in the case of transmission, and state public utility commissions, in the case of delivery to final retail consumers.[5] In some places, companies will fully divest themselves of either their competitive or their monopoly operations; in others, structural separations will be put in place to prevent companies from using their control of bottleneck facilities to favor their competitive activities. In any event, competitive products will be priced not by regulators but by market forces (assuming the efficient functioning of the market), and monopoly services will continue to be regulated.

The manner of regulation of monopoly distribution services will surely vary from state to state — traditional cost-of-service regulation, price caps, revenue caps — but there will remain as always the need to design rate structures for those services. The unbundling of generation from transmission and distribution has exposed these services to a new scrutiny. Whereas previously distribution and related services might have made up anywhere from ten to forty percent of a vertically integrated company's costs, they make up virtually all of a distribution-only company's costs. Under these circumstances, it is only natural for firms to consider new ways of reducing the risks attending those revenue streams; however, not all actions taken to enhance the financial position of the utility are necessarily consistent with the greater, longer-term interests of the public and the economy.

B. What is Distribution?

For most readers, this question need not even be asked. Regulators, company officials, advocates, and other policymakers who daily inhabit the world of electricity readily understand the tripartite distinctions in the configuration of the electric system: generation, transmission, and distribution. Generation, of course, refers to the facilities that actually produce electricity.

5. The line between FERC and state jurisdiction is the subject of much debate; demarcating it clearly has been one of the many complex issues that proposals for federal legislation have tried to tackle. This paper assumes that a separation between state and federal jurisdiction along the lines historically drawn will remain largely intact.

Whether transmission and distribution in fact remain natural monopolies is a subject of some debate. From a narrow perspective, they do indeed appear to be natural: single providers that can most efficiently provide the delivery service. But a broader view suggests that alternatives to them are available: other energy sources, local generation, and efficiency, for example. But even those who argue that the monopoly status of transmission in particular will erode over the next few decades do not conclude that it does not possess significant market power and should not be regulated (in some fashion) therefore. See, for example, Awerbuch, Shimon, Leonard S. Hyman, and Andrew Vesey, *Unlocking the Benefits of Restructuring: A Blueprint for Transmission*, Public Utilities Reports Inc., Vienna, VA, Nov. 1999, pp. 8, 15, 25-28, 143-169, 203-240.

Transmission consists of the network commonly called the "grid": the high-voltage power lines and hardware used to transport electricity in bulk from generators to load centers where, through substations (which are really just large transformers), its voltage is reduced and it is fed into the distribution system.

The distribution system makes up the remainder of the electric system: the lower-voltage wires, transformers, and related functions (*e.g.*, hook-ups, metering, and billing) necessary to deliver power to homes and businesses. It can be further broken down into sub-components: primary and secondary lines and transformers. Primary lines, or "feeders," typically run from transmission sub-stations to transformers located closer to loads, where the power is further "stepped down" to be transported over secondary lines. Distribution, though similar to transmission in certain respects (it requires frequency synchronization and voltage support), differs in others. It is built with less flexibility and redundancy than is the bulk power grid, which is generally designed to withstand the failure of one or more lines. Typically, distribution systems "are one-way, one-path networks that branch into successively smaller limbs as they progress toward the customer."[6] However, some networking (interconnecting at several points) of these "radials" can be accomplished, thereby improving reliability.

The boundaries among generation, transmission, and distribution are not always easily drawn, insofar as the system as a whole requires close, integrated operation to prevent power failures. Moreover, investments of one type are often made to avoid higher-cost investments of another: for instance, when distributed resources such as end-use efficiency and small-scale generation are deployed in areas of concentrated load expressly to circumvent the construction of more expensive transmission and distribution plant or when wires are sized not merely to satisfy peak demand but also to reduce energy losses, thereby reducing generation and other delivery costs.

C. Pricing Distribution Services

1. Treatment of Distribution in the Vertically-Integrated Utility

Customers served by vertically integrated electric companies typically pay for their electricity through bundled energy (per-kWh) or through a combination of energy and demand (per-kW) charges. Included in those charges are all the costs incurred by the utility to generate, transmit, and distribute the commodity (a significant portion of which are the joint and common costs necessary to those services but directly attributable to none). Lower volume consumers (*e.g.*, residential and small commercial) commonly pay rates under two-part tariffs — energy and customer charges — differentiated in some cases by season or time of day, or both. Higher usage customers generally take service under three-part tariffs, with the costs of capacity removed from the energy charge and captured in a separate demand rate. In special cases, very large-volume

6. Fox-Penner, Peter, *Electric Utility Restructuring: A Guide to the Competitive Era* (Public Utilities Reports, Inc., Vienna, VA, 1997), p. 262.

users may interconnect at a higher level in the system, at primary- or transmission-level voltages, which requires dedicated facilities for those customers, for example their own sub-stations. The costs for those facilities are borne by the customer they serve, either directly or through charges negotiated with the utility. Commodity, capacity, and remaining delivery costs are covered in rates, bundled or otherwise. In all cases, the rates are subject to regulatory review and approval.

The recovery of capacity and energy costs through separate rate elements has historically been justified on economic efficiency grounds, but the focus of utility managers and regulators has invariably been on pricing and recovering the costs of generation (a firm's riskiest investments) and, to a lesser extent, on transmission. Distribution costs, a usually smaller share of total costs, have invariably been translated into per-kWh charges and recovered in energy rates generally or, in the cases of larger-volume customers, recovered through demand and energy rates that also recover generation and transmission costs.

The customer charge deserves special mention here. It is, after all, a fixed, recurring charge that customers pay whether they flip a light switch or not. Proposals for fixed monthly distribution charges are, in effect, simply requests to radically increase the historical customer charge that utilities have used to collect certain minimum costs, usually billing and metering, but also occasionally customer service drops and transformers. Customer charges gained some acceptability on the grounds that metering and billing costs are not related to usage at all, and were adopted by commissions in part on that basis and in part because they rendered the utility a certain amount of revenue stability. But regulators have long resisted disproportionate growth in such charges, primarily because of their adverse impacts on low-volume users. Whether the traditional customer charge will continue to make sense as the industry becomes more competitive is another aspect of the overall distribution pricing question.

2. Distribution in a Restructured Industry

Restructuring has been driven in large measure by the changing technology and economics of generation. By far the largest cost component of the electric industry, generation has shown itself amenable to competitive provision, as the economies of scale associated with increasing capacity of central stations have been largely exhausted. This has led, in those states that have restructured, to the separation of competitive from monopoly services (whether through divestiture or structural separations) and to the "unbundling" of services and rates: generation (usually including transmission) on the one hand and distribution (and other customer services) on the other.

3. Recent Proposals for Pricing Distribution Services

Restructuring, by formally separating distribution from generation and transmission, has precipitated new proposals for the pricing of distribution services. In 1999, Nevada Power Company and Sierra Power filed cost of service and rate design proposals that called for significant increases in the monthly recurring customer charges, large enough to fully cover their distribution costs. In the spring of 2000, the Nevada Public Service Commission approved the companies' filings for the most part.

Early in January 2000, Southern California Edison filed a proposal for a $17.00/month fixed customer charge for residential consumers.[7] Like the Nevada companies, SCE argued that distribution system costs are for the most part fixed and do not vary with load or throughput and therefore should be covered in fixed, recurring charges.

In Maine, in setting rates for distribution services prior to the full scale introduction of competitive generation services, the Public Utilities Commission explicitly rejected the proposals to impose fixed, recurring monthly charges. In the case of Central Maine Power, rates have been unbundled: $0.041/kWh for energy, $0.074/kWh for transmission and distribution (including stranded cost recovery), and a $7.40/month customer charge, designed to cover the costs of metering, billing, and customer service only. (The customer charge is, in fact, a minimum distribution charge: the $7.40 gives a customer up to 100 kWh of "free" delivery service in a month.)[8]

The current debate about pricing for distribution services is in essence a debate over whether customer charges (long an element of utility rate design, but for the most part quite small) should be significantly increased. This report analyzes the rationales for, and effects of, fixed and usage-based rates – in other words, compares and contrasts increased customer charges with volumetric rates. The narrower question of the traditional role of customer charges is taken up in Section V.A.5., below.

D. Pricing in Competitive Markets

It has often been said that regulation is meant, among other things, to serve as a proxy for competition, to impose upon a single provider the disciplines of competitive markets.[9] Therefore, when designing rates, it is appropriate for regulators to consider how competitive markets price their goods and recover their costs of production.

One important aspect of competitive markets is their inability to differentiate among consumers for the purposes of pricing. All consumers, regardless of the relative values they place upon the good in question, pay the same price. Ramsey pricing (discriminating among consumers according to their willingness to pay) is impossible, since attempts by some purchasers to resell

7. Application of Southern Californian Edison Company for Post-Transition Rates, A.00-01-009, filed January 7, 2000.

8. MPUC Docket 97-580, Order of 3/19/99. The Maine PUC intends to further unbundle rates by separating transmission from distribution.

9. Kahn, Vol. I, p. 17; Bonbright, p. 372; Pierce and Gellhorn, pp. 2, 47-48, 94-95.

the product to others who are willing to pay higher prices (*i.e.*, to earn profits by arbitraging the differences in elasticities) will quickly drive down prices to the intersection of marginal cost and demand — to the market clearing price.[10] However, it often happens that the cost to supply one market or region actually differ from those to supply another, typically having to do with increased delivery and other transactions costs that flow from unique features of the particular sub-market (*e.g.*, geography); in such cases, the price of the good will often vary from market to market.

How are products actually priced in competitive markets? Groceries, automobiles, fuels, agricultural products, appliances, communications services, entertainment, even electricity — the list is endless — are all priced in ways that reveal something about how competitive markets operate and about consumer preferences.

Commodity markets come closest to meeting the requirements of economic theory's "perfect competition." Sorghum, crude oil, pork bellies, to name just a few, are all traded in markets where both suppliers and buyers generally lack the power to unilaterally affect price, the product is homogeneous among all suppliers, and quality and price information is instantly available. The commodity is sold on a unit basis, and prospective buyers are not required to make minimum payments, even if they choose to purchase nothing.[11] As for virtually all goods, the production of these commodities invariably involves agricultural, extraction, or other processes that require suppliers to make investments in fixed assets (land, processing equipment, etc.) and to incur on-going operational expenses (labor, fuel, transportation, etc.).

Most retail sales take place under similar conditions. Grocery stores, department stores, boutiques all place a variety of offerings on their shelves. Consumers are free to pick and choose amongst them, according to their needs and wants, and purchase as little or as much as they wish. Included in the costs of their products are a host of costs — the fixed and variable costs of production, delivery, and marketing — which all must be recovered through the sale of the goods. The retailer's own costs too must be covered, but it is rare indeed that it charges potential customers simply to enter the premises.[12] But even if some retailers do impose such fees, they

10. "One of the virtues of pure competition is that it eliminates the possibility of price discrimination." Kahn, I, at 123. See also *id.*, pp. 64 and 133 (inc. fns. 18 and 19); Bonbright, p. 372-374; Bonbright *et al.*, p. 531.

11. Brokers who have seats on an exchange do, however, have to pay for that seat, or "access." Those payments support the operations of the exchange itself. And, though the costs of a seat may be quite high, they are, as a function of the total value of the transactions performed by the broker, quite small. Typically, they are recovered from clients through transaction-based service fees.

12. There are, of course, examples of retail operations (Price Club, Costco, Sam's Club) that require buyers to pay a fixed periodic (*e.g.*, annual) charge for access to their stores, where they then can enjoy discounts associated with large-volume purchases. The pricing scheme is designed to attract precisely those in-bulk purchasers, both households and businesses. It is important to note that the membership fee comprises a small portion of the total amount that a customer typically spends in the store and that the revenues from such fees do not begin to cover all the fixed costs of the enterprise. Often, for marketing purposes, the charge is waived.

At retail, there is often less homogeneity of products than there is in wholesale commodity markets.

(continued...)

are nevertheless avoidable: customers can simply choose alternative suppliers who do not exact similar charges.

Long distance telephone service is available to consumers from a seemingly endless array of providers who offer an almost equally endless array of products and pricing options. Historically, customers were charged per-minute usage charges that differed by time of day, distance, and destination (in-state or out). Regulators justified these differences in pricing by reason of both cost causation (peak, off-peak) and value of service (for daytime users, mostly businesses). As the long distance market became (and becomes) increasingly competitive, as technology evolved, and as usage patterns changed, the justifications for such pricing policies diminished, and new pricing schemes have emerged. The central feature of any long distance product remains, however, the unit price per minute, and carriers compete primarily on this basis. But the services are packaged in a variety of ways, giving customers choices about how to best meet their needs: minimum monthly charges accompanied by guaranteed per-minute charges, volume discounts, or time-differentiated usage-only charges. While some of these offerings may require a minimum monthly payment from a consumer if her usage does not exceed a specified level, the consumer is not constrained in her choice of product or carrier; and most, if not all carriers, offer service priced on a per-minute only basis.[13]

The fact that telecommunications is a capital-intensive industry makes it an interesting case study.[14] Airline travel is also an interesting study. Also highly capital intensive, the airline industry covers its costs through usage, per trip, prices. While ticket prices vary widely by duration, time of week and year, routing, cabin section, and even time of purchase, no airline requires its potential passengers to pay a fixed periodic charge simply for the opportunity to later purchase travel services. The idea, of course, is ludicrous. As it would be for automobiles,

---

12. (...continued)
Suppliers go to great lengths to differentiate their products from others', to create in effect niche markets for their particular goods, thus giving them some degree of monopoly power. The differentiation of one brand of soap from another, of Ford automobiles from Chevies, and of Sam's Club from Safeway are all examples of this kind of behavior. But the essential point remains that, where there is product choice, consumers can avoid costs.

13. In fact, this kind of service is heavily marketed by a number of carriers these days. "No monthly fees!" is a central element of their sales campaigns.

It is true that a number of the costs that long distance carriers incur to provider service are set by state and federal regulators — per-minute access charges, per customer fees, etc. — but they are free to recover those costs in rate designs of their choosing, which consumers are equally free to adopt or reject.

14. Recent developments in the wholesale pricing of the components of the local exchange network may have applicability to the question of electric distribution pricing. A number of states and the Federal Communications Commission have adopted a pricing policy that is based on the total element (or service) long-run incremental cost methodology (TELRIC). TELRIC calls for the calculation of an average incremental cost, given the *total* demand for a particular good or service (or network functionality) over the planning horizon. For a network industry characterized by economies of scale, this method will capture all the incremental costs, both fixed and variable, caused by demand for the functionality.

gasoline, shoes, package delivery, and the thousands of other goods and services that households and business purchase every day.[15]

Perhaps the most interesting, and most apposite, example of pricing in a competitive market is that of the electric generation sector in those states that have restructured their industries. Here we find an industry in which a high proportion of costs are fixed, but yet are recovered in prices per unit of output – kilowatt-hours. In New England, for example, the independent system operator (ISO-NE) manages a power exchange into which suppliers "bid in" electricity (often, power in excess of their contractual commitments) on an hourly basis.[16] The bids are made a day in advance, for the hours in which a supplier expects to have power available for sale. In each hour, generators are dispatched by the grid operator (ISO-NE, in this case) on the basis of their bids, lowest first, followed by successively higher ones until demand in that hour is satisfied. The bid price of the last (*i.e.*, most expensive) unit dispatched in that hour becomes the price that all output in that hour is paid. The power market in the United Kingdom works in a similar fashion.

A supplier is free to bid whatever price he chooses. Whether his generation will be dispatched depends on where in the resource "stack" his bid has put him. Too high, and his unit (or units) will not be called upon. For this reason, this kind of bidding and dispatch system will generally cause generators to bid their variable costs of operation (for the economist, the short-run marginal cost). In some cases, a supplier may bid less than variable cost — even zero. Such bids assure that a unit will be dispatched; but, since it is the bid price of the last needed unit in that hour that determines what generators are paid, the zero-price supplier can be confident that he will be paid some amount of money for his output in that hour.

The question then is how do generators recover the capital costs of their investments? As load increases from one hour to the next, the dispatch (market clearing) price also increases, since more costly units are needed to meet that increasing load. At the very highest loads, the last units being dispatched (typically, combustion turbines) have operating costs per kWh that exceed not only the operating costs of most of the other generation, but often also their average total costs. At these times, therefore, generators lower in the "stack" receive additional revenues that cover some or all of their capital costs. Figure 1 shows graphically how units whose operating costs are

---

15. The variations in prices for what is ostensibly the same product – airline travel – suggest that, in some ways at least, airlines engage in value-of-service pricing. To the extent that there is not a meaningful alternative to the product – travel at a particular time, with specified originating and terminating points, and with a desired level of associated services – then an airline may be in a position to charge in excess of an otherwise competitive price. To the extent, however, that the travel products are distinguishable – by the period of advance notice given, cost of itinerary changes, and need for and timing of a return trip – then other cost factors come into play that justify different prices.

16. It doesn't matter whether dispatch is based on generators' bids or on their actual variable operating costs. The analysis that follows is applicable to either case. (If the market functions properly, bids will mostly equal operating costs.)

less than the system's marginal cost (*i.e.*, the marginal bid that sets the market clearing price in the hour) receive extra revenues to offset their capital costs.

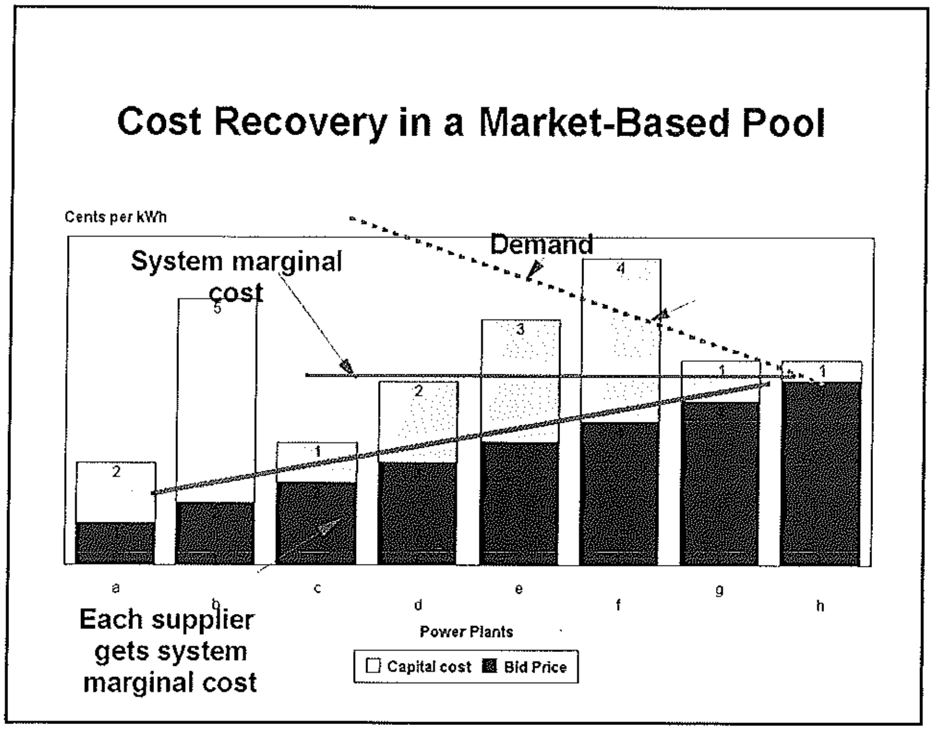


Figure 1

It is worth examining products whose price structures seem to belie the contention that competitive markets offer usage-based pricing. Membership fees for health or country clubs typically are fixed by period, regardless of the number of times a member makes use of the facilities during the period. This raises the question of whether the market for the product is in fact competitive. If the good is physical fitness or rounds of golf, a consumer has plenty of alternatives to paying fixed monthly charges — other clubs, other exercise regimens, other golf courses — if she wishes to avoid such fees. If, however, the club has unique attractions or in some other way possesses some measure of market power (*e.g.*, lack of other providers in the area), then the premise that the good or service is competitively provided is mistaken, and we should not be surprised when the firm sets prices in a way that exploits that power.

Cable television (CATV) is another example of such a uniquely situated industry. While CATV service is but one of the scores of alternative forms of entertainment (movies, videos, and reading to name only a few), it nonetheless possesses substantial market (if not monopoly) power in one sub-segment of that market: the market for the delivery of television programming. Here, as with certain health and country clubs, the consumer pays for access, and usage is essentially free. Often there are different packages of service — minimum or "basic" programming, expanded, and then premium (movie channels), as well as pay-per-view offerings (concerts, prize fights, and other one-time events). Satellite technology has begun to compete with CATV, but there is some question as to whether it has yet injected much discipline into the market. The levels and structures of prices for satellite service typically mirror those of wireline companies.

Similar observations can be made about use of the Internet. In return for a fixed monthly fee, Internet service providers (ISPs) typically provide unlimited access to the worldwide web. Usage is free (not including any incremental phone charges that might apply). The monthly fees vary little from provider to provider (in the range of $20, depending on the kinds of ancillary services are offered), but many of the ISPs (AOL, Earthlink, AT&T, CompuServe, IBM) also offer low-use options (for example, $5.00/month for five hours). Many of the ISPs supplement their revenues through the sale of advertising services, and at least several (Juno, Net4Free, NetZero) rely only on advertizing and other sales for their revenues, while providing Internet access to customers at no cost. The competitive nature of the web is changing.

The home heating fuel business offers insights as well. Fuel oil is typically priced on a per-gallon basis, and the unit-price decreases as volume increases. This is simple arithmetic: the delivery, marketing, and billing costs of lower-volume customers must be recovered across fewer units sold. In certain cases (often where low-volume price premiums are less common), suppliers may require a minimum purchase (say, 100 gallons) in order to cover the delivery costs. Propane is similar, although there is the added twist of tank ownership. In order to assure that on-premises pressure tanks are properly and safely maintained, the propane company will retain ownership. Although the tanks are removable, and companies frequently do so in order to sign up new customers (often only on the condition that they will remain customers for some minimum period of time), it nevertheless constitutes an added barrier to the free and efficient operation of the market. Even so, customers pay only for the amounts they use.

The general conclusion to be drawn from these examples is that in a competitive market a consumer can, if she so chooses, purchase goods and services in any of a variety of ways, including on a unit basis, or, if she chooses not to buy, spend no money at all. Competitive markets are by their very nature hostile to the imposition of unavoidable charges upon consumers; such charges are only sustainable, by themselves, when a firm can exercise some degree of market power. Competitive markets provide goods and services in all sorts of ways, with an almost infinite variety of product offerings and pricing structures: consumers are given meaningful choices and are thus able to avoid costs either by not consuming or by finding substitutes. And the availability of goods and services on a price per-unit-purchased basis is a feature common to them all.

III. AN OVERVIEW OF UTILITY PRICING

No discussion of appropriate pricing policies can be meaningful without an understanding of the objectives of those policies. What interests are to be served? What are the intended outcomes? This chapter provides a general background into these questions.

A. The Goals of Economic Regulation

There are two broad, fundamental justifications for governmental oversight of the utility sector. The first is the widely-held belief that the sector's outputs are essential to the well-being of the society — its households and businesses — and the second is that its technological and economic features are such that a single firm often can serve the overall demand for its output at a lower total cost than can any combination of more than one firm. Competition cannot thrive under these conditions and, eventually, all firms but one exit the market. This is called "natural monopoly," and, like monopoly power in general, it bestows upon the surviving firm the power to restrict output and set prices at levels higher than are economically justified. Economic regulation is seen then as the necessary and explicit public or governmental intervention into a market to achieve a public policy or social objective that the market fails to accomplish on its own.

In light of the economic and public welfare characteristics of utilities, certain purposes for price regulation emerge. They can be generalized in the two goals of *economic efficiency* and *fairness* (or equity), which can then can be further broken down as follows:

- *Economic Efficiency.* Since electric utilities generally do not operate in competitive markets that would impose cost discipline upon them, regulation must fulfill that function. This objective is promoted by setting rates that reflect, to the greatest extent possible, the long-run marginal costs of production.[17]

---

17. This statement deserves fuller proof, which we give in Appendix A. Sufficient for our purposes here, we can say that the economically efficient price is that which reflects the actual (or *marginal*) cost incurred when more or less of a good is produced, and when the costs of other goods and services (*i.e.*, other consumption opportunities) also reflect their marginal costs of production. Only then can fair comparisons be made and resources appropriately allocated. In this way, consumers, who make purchasing decisions based on the relative values that they assign to alternative uses of their own resources (income and wealth), will make consumption decisions that allocate society's resources to uses to which they assign greater relative weights. If a good is underpriced (priced below its marginal cost), then some quantity of the good will have been produced at a cost that exceeds its value to society, and the resources that were given to its production could have been allocated to better (more highly valued) uses elsewhere. The converse is true of over-priced goods. Competitive markets will drive the prices of goods down (or up) to marginal cost.

We speak of economic efficiency generally, but it is important to understand that there are a variety of aspects to it. There are allocative and productive efficiencies — those associated with the allocation of goods (efficiencies in consumption) and those associated with the allocation of inputs to the production process — as well as X-efficiencies, arising from how well management maximizes output for a given level of inputs. "Static"

(continued...)

- *Fair prices*, for consumers and investors. Price regulation is intended to guard against the reaping of supra-normal profits (called economic "rents") while still enabling the utility to generate revenues adequate to cover prudent expenses and investment and to provide a reasonable return on that investment. Prices should also be fair to competitive providers or, more accurately, the competitive process. Their distortional effects on the economy should be minimized.

- *Non-discriminatory access* to service for all consumers.
- *Adequate quality and reliability*. Because electricity is an essential service, reliability is critically important.
- *Other stated public policy objectives* (*e.g.*, environmental protection, universal service, low-income support, energy efficiency, etc.).[18]

For goods and services that can be provided by competitive markets, the markets by themselves will go a long way toward meeting these goals.[19] Thus it can be said that economic regulation is intended to achieve outcomes that competition, were it feasible in the market in question, would otherwise achieve.[20] Also, prices in regulated industries naturally affect prices in competitive ones, and vice versa, and affect therefore the overall efficiency of the economy: all the more reason then to adopt utility rate designs that most closely resemble price structures in competitive markets.

B. Goals of Rate Design

The general goals of economic regulation inform the rate design process. More specifically, the object is to set economically efficient and fair prices, while simultaneously giving the regulated firm a reasonable opportunity to recover its legitimate costs of providing service (including return of, and on, its investment).

---

17. (...continued)
efficiency refers generally to efficiency in allocation of inputs and outputs at a given point in time. "Dynamic" efficiency denotes efficiency (including innovation and technological development) that arises over time. The differences between the short-run and long-, and the general preference that rates cover at least the long-run marginal costs of service, are discussed in more detail in Appendix A.

18. Bonbright, pp. 25-41; Pierce, Richard J., Jr., and Ernest Gellhorn, *Regulated Industries*, 4th Ed. (West Group, St. Paul, MN, 1999), p. 11; Kahn, Vol. I, pp. 20-25, 69-70, and Vol. II, pp. 243-246.

19. This is not to say that competitive markets will, by themselves, satisfy all, or fully any, of the welfare-enhancing objectives that a society embraces. Transactions costs, externalities, lack of information, the preexisting distribution of wealth and income, to name a few, all affect the operations of markets in ways that often call for some form of governmental intervention into the market, for the benefit of the public overall. Content labeling, performance requirements, health standards, labor, anti-trust and anti-discrimination laws, financial requirements, etc., are all examples of governmental actions taken to assure that other, highly-valued outcomes (such as equity) are achieved.

20. Kahn, Vol. I, p. 17; Bonbright, p. 372; Pierce and Gellhorn, pp. 2, 47-48, 94-95.

The particular problem faced by regulators in this exercise is that the legitimate historical (accounting or "embedded") costs that a utility incurs are to be recovered in rates, but these costs may only bear a passing resemblance to the marginal costs that form the basis of economically efficient prices. The reconciliation of the need to cover historical costs with the desire to set economically efficient prices, and then to meet other objectives of regulation requires much judgment. Bonbright dedicated five chapters and 120 pages to the subject, beginning with a catalogue of the several and sometimes competing objectives of rate design. It remains today the comprehensive synthesis upon which regulators rely. Paraphrased, Bonbright's principles are:

Revenue-Related Objectives:

- Rates should yield the total revenue requirement;
- Rates should provide predictable and stable revenues; and,
- Rates themselves should be stable and predictable.

Cost-Related Objectives:

- Rates should be set so as to promote economically-efficient consumption (static efficiency);
- Rates should reflect the present and future private and social costs and benefits of providing service (*i.e.*, all internalities and externalities);
- Rates should be apportioned fairly among customers and customer classes;
- Undue discrimination should be avoided; and,
- Rates should promote innovation in supply and demand (dynamic efficiency).

Practical Considerations:

- Rates should be simple, certain, payable conveniently, understandable, acceptable to the public, and easily administered.
- Rates should be, to the extent possible, free from controversies as to proper interpretation.[21]

C. Rate Design and Recovery of the Costs of Service

Those features of monopolies that render it difficult, if not impossible, to rely upon competitive markets to set the prices of their services, that give them market power, also preclude regulators from simply setting prices at marginal cost (short-run or long-), since to do so would, in most cases, be to condemn the companies to financial hardship. So long as their marginal costs of production decrease as output increases, their average total cost to serve will remain greater than

21. Bonbright, p. 291; Bonbright *et al.*, pp. 384-385.

their marginal cost; and price set at marginal cost will fail to generate sufficient revenues to cover their legitimate costs of service.[22]

Faced with this dilemma, regulators must necessarily set average price above marginal cost, if their utilities are to continue to provide safe, adequate, and reliable service (on the assumption that there is no other source of revenue — *e.g.*, taxation — available to make up the shortfall). The simplest way to set rates would be to divide the revenue requirement by sales volume (kWh).[23] While this would produce rates that would (ignoring their demand effects) likely generate revenues sufficient to cover costs, it is not clear that societal welfare overall would be maximized.[24] But the assumption upon which such a rate design is based, that marginal cost equals average cost, is flawed. Typically, prices equal to average cost would, at certain times, exceed marginal cost and, at others, be less than marginal cost. Production and consumption would hardly be efficient, or fair, under such circumstances.[25]

1. Reconciling Marginal Cost Rates and the Revenue Requirement

A few jurisdictions use marginal cost studies to inform the rate design process, to discover how far embedded cost rates deviate from marginal cost rates. Since it would only be by coincidence that rates set at marginal cost would generate funds sufficient to cover the revenue requirement,

---

22. "The usual assumption is that, if the incremental costs of all services, separately measured, were added together, they would fall materially short of covering total costs — an assumption based on the belief that most public utility enterprises operate under conditions of decreasing costs with increasing output. When this assumption is valid, it implies that a public utility company cannot cover its total revenue requirements without charging *more* than incremental cost for at least some of its services." Bonbright, p. 299.

Note that the term "total costs" is synonymous with "revenue requirement," that amount of money needed by a utility to meet it total cost of service in defined period. The simple formula for a revenue requirement under rate-of-return rate-making, the one upon which most regulatory commissions rely, is as follows:

$$RR = E + d + T + [r * (V - D)]$$

where:

RR =	Revenue requirement, or total revenues
E =	Operating expenses
d =	Annual depreciation expense
T =	Taxes
V =	Original book value of plant in service
D =	Accumulated depreciation
	Note: (V - D) = "Net rate base"
r =	Weighted average cost of capital

23. Mathematically: Rates = (Revenue Requirement)/(Volume of sales)

24. The validity of this conclusion and the magnitude of the consequences depend on consumers' price elasticity of demand and on whether the excess of price above marginal cost fully reflects the "external" costs of production. See Chapter V, Section A.3., and the Appendix.

25. Some regulators have professed skepticism of the assumption that consumers respond rationally to prices, thus making deviations from marginal cost pricing less of a concern.

some method for reconciling the two is called for.[26] Many considerations inform these pricing decisions, among them fairness and the desire to inhibit economic efficiency as little as possible.[27] Economists argue that any deviation from marginal cost pricing harms efficiency to some degree and they counsel therefore that the least distorting method is most desirable. As a general matter, while economic efficiency is a legitimate objective, it is not in all cases necessarily the overriding one. Rates should equal at least the long-run marginal costs of service. After that, how historical costs in excess of marginal costs should be recovered should not be dictated by efficiency considerations alone; other factors, such as shifts in revenue burdens among customers and customer classes, effects on universal service, and changes in risk allocation, also should be weighed.

There are several general approaches for recovering additional revenues when marginal cost prices by themselves do not cover total costs.[28] One is simply to effect a "lump-sum" transfer from customers to the utility, to bridge the gap between revenues generated by marginal cost prices and the overall revenue requirement. The difference between the two is simply collected as a surcharge (or a rebate in times when marginal costs exceed average) on customers' bills. For reasons of equity (proportionally higher costs or refunds fall upon low-volume users) and unpredictability of application, it is a rarely used technique.

The inverse-elasticity rule, or Ramsey Pricing (named for the economist who first proposed it), calls for imposing the greatest mark-ups (or discounts) to marginal cost prices on those customers whose demand is least elastic, that is, on those customers who place a higher value on the service and are thus less willing to forego it. In this way, goes the argument, demands (relative consumption levels) are altered least, and the overall allocation of resources conforms more closely to that which would flow from straight marginal cost pricing. Applying the rule requires some knowledge of demand elasticities for electric consumers and products.[29]

There is no one method for implementing the inverse-elasticity rule, but rather a host of them, each with different distributional impacts. Adjusting the customer charge component of a tariff is a typical approach; energy and demand charges are set at marginal prices while the customer charge (a fixed, periodic fee, unaffected by demand or usage) is set to collect the revenue shortfall. Like the lump-sum transfer, however, it has obvious equity problems. Another

---

26. Bonbright, p. 367; Uhler, Robert G., "Electric-Utility Pricing Issues: Old and New," in Danielson, Albert L., and David Kamerschen, eds., *Current Issues in Public Utility Economics: Essays in Honor of James C. Bonbright* (Lexington Books, Lexington, MA, 1983), p. 79.

27. Bonbright, p. 300.

28. Electric Power Research Institute (EPRI), *Rate Design: Traditional and Innovative Approaches* (CU-6886, Research Project 2343-4, Palo Alto, CA, July 1990), pp. 7-3 to 7-11; Kahn, I, pp. 123-158; National Association of Regulatory Utility Commissioners (NARUC), *Electric Utility Cost Allocation Manual*, Washington, D.C., 1992, pp. 147-165.

29. Elasticities for electricity, like other utility products, are often difficult to estimate because demand for the product, which is merely an input necessary to meeting some other want (*e.g.*, lighting, heating, television, industrial processes), must be inferred ("derived") from consumer demand for ultimate end-use or good. Bonbright, p. 382.

approach is to adjust demand charges, on the assumption that they are less elastic than energy charges. A difficulty here arises in those rate classifications for which there are no demand charges. A third technique is to increase (or decrease) all rate components for all rate classes equally. The presumption in this case is that demand elasticity is constant across demand, energy, and customer classes. There is no reason to think that this is likely.[30]

Setting rates to recover total costs in those instances when marginal cost-based rates will by themselves fail to do so is problematic at best. There is no "objective standard of rationality" for doing so.[31] Moreover, since "no rate relationships can be made completely nondiscriminatory as long as all or some of the rates must be set above marginal costs in order to yield adequate revenues,"[32] commissions regularly struggle with apportionments of overhead and other non-marginal costs among customers and classes, according to vague notions of fairness and reasonableness. (Indeed, most commissions rely on embedded cost studies to allocate costs and design rates, and effectively avoid the marginal cost question altogether.) In the world of distribution utility regulation, commissions will continue to face this age-old challenge.

30. Use of Ramsey, or "value of service," pricing has other, more profound drawbacks as well. In particular, the many preconditions necessary for application of the rule to produce its desired effects rarely exist simultaneously and, in the absence of one or several, the outcomes produced by the rule may in fact do more harm than good. See Bonbright *et al.*, pp. 533-542. ". . . Ramsey pricing, because of its unsatisfactory formulation, remains for the present more of a theoretical curiosity than a workable regulatory rule." *Id.*, p. 534. See, also, Sheehan, Michael, "Why Ramsey Pricing is Wrong: The Case of Telecommunications Regulation," *Journal of Economic Issues*, Vol. XXV, No. 1, March 1991, pp. 21-32.

In one form or another, all states have prohibitions against *undue*, or *unjust*, discrimination amongst utility consumers. For the most part, regulators and courts have found that differentiating among consumers on the basis of the costs to serve them is not undue discrimination. The most common example of such differentiation is the grouping of customers, on the basis of their expected load profiles, into rate classes, *e.g.*, residential, commercial, and industrial. To the economist, this is one form of *price discrimination*. Another form is differentiation according to consumers' willingness to pay — that is, Ramsey pricing — and some advocates argue that it, unlike more strictly defined cost-based distinctions, may run afoul of undue discrimination statutes.

31. Bonbright, p. 338.

32. *Id.*, 342.

IV. THE COSTS OF DISTRIBUTION SERVICES

A first question to be answered when designing rates is what does it cost to provide the service? What are the causes and magnitudes of the relevant costs? It's helpful to observe that the costs recovered by distribution-level rates have historically extended far beyond the distribution system. Are there other costs, not directly related to distribution services, that distribution rates are expected to recover? What follow here are an overview of utility costing methodologies and a discussion of some practical considerations to keep in mind when determining rate structures.

A. Utility Plant Costing Methods

Utilities and regulatory commissions use a variety of methods for determining and allocating cost responsibility among customers and customer classes. There are two general types of cost study, embedded and marginal. Embedded, or fully distributed, seeks to identify and assign the historical, or accounting, costs that make up a utility's revenue requirement. Marginal, as the name connotes, aims at determining the change in total costs imposed on the system by a change in output (whether measured by kilowatt-hour, kilowatt, customer, customer group, or other relevant cost driver). Each commission around the country uses these studies in its own way to inform the rate design process; in the end, most commissions rely on embedded cost studies for ultimate allocations and price levels, constrained as they are by a legal requirement to set rates that offer the prudent utility a reasonable opportunity to earn a fair rate of return on its assets used in service to public.[33] The allocations, however, are often structured to reflect at least relative differences in the marginal costs of providing a company's various services.

1. Cost Causation

There is broad agreement in the literature that distribution investment is causally related to peak demand. Numbers of customers on the system and energy needs are also seen to drive costs, but there is less of a consensus on these points or on their implications for rate design. In addition, not all jurisdictions employ the same methods for analyzing the various cost components, and there is of course a wide range of views on their nature — marginal, embedded, fixed, variable, joint, common,[34] etc. — and thus on how they should be recovered in rates.

33. NARUC, p. 32.

34. The costs of multiple products or services supplied by the same plant or process are either "common" or "joint." Common are those that generally do not vary with changes in output. The classic example is the president's desk, which is needed to run the firm as a whole but is incremental to the provision of no particular good or service. Another example is that of an airline flight, the majority of whose costs are incurred in a single lump and do not vary with the number of passengers carried. Put another way, common costs are those for which "the unit of production (the single flight), which is the basis of cost incurrence, is larger than the unit of sale (a single ticket to a single

(continued...)

Numbers of customers, usage, and demand, however, are only part of the story. Other factors also play an important role: geography (particularly population density), system design (*e.g.*, aerial versus underground lines), and the utility's business practices (for example, the extent of expenditures on billing, answering customers' questions/complaints, etc.). The implications of such factors on rate design is unclear, however: one can charge for services on the basis of numbers of customers, usage, and demand, but not on the basis of other such factors.[35]

2. Embedded Costs

a. Cost Classification: Customers, Demand, and Energy

Traditionally, customer costs are those that are seen to vary with the number of customers on the system — service drops (the line from the distribution radial to the home or business), meters, and billing and collection. Some utilities and jurisdictions also include some portion of the primary and secondary distribution plant (poles, wires, and transformers) in these costs, on the ground that they also are driven more by numbers of customers than by demand or energy. Similar reasoning leads to the designation of the costs of customer service and customer premises equipment as customer-related. But, since the system and its components are sized to serve a maximum level of anticipated demand, the notion that there are any customer costs (aside from perhaps metering and billing) that are not more properly categorized as demand can be challenged (see Subsections 3 and 4, below).

Utilities classify significant portions of their embedded distribution investment as demand-related, reasoning that it is designed and installed to serve a customer or group of customers according to their contribution to some peak load (system, substation, etc.). Substations are a typical example of such costs, but so too may be a significant portion of the wires and related facilities, since they are "sized," at least in part, to serve a peak demand.

There are a number of methods for differentiating between the customer and demand components of embedded distribution plant. The most common method used is the "basic customer" method, which classifies all poles, wires, and transformers as demand-related and meters, meter-reading, and billing as customer-related. This general approach is used in more than thirty states. A

---

34. (...continued)
passenger)." Kahn, Vol. I, p. 77. If services produced in common can be produced in varying proportions, it may then be possible to identify separate marginal production costs for each.

Products that are produced in fixed proportions (*e.g.*, cotton fiber and cottonseed oil, beef and hides, mutton and wool) are characterized by joint costs. For that aspect of their production process that is joint, the products have no separately identifiable marginal costs. *Id.*, p. 79. See also Bonbright, pp. 355-360.

35. These other cost factors can have huge effects on prices. Three distribution utilities in the American south, owned by the same holding company and using the same costing methodology, recently proposed new metering, customer service rates, and delivery rates. The rates, designed as a combination of monthly per-customer and per-kW of peak demand charges, vary from company to company by ratios ranging from 1.25 to 1.9.

variation is to treat poles, wires, and transformers as energy-related – driven by kilowatt-hour sales – but, though it has obvious appeal, only a small number of jurisdictions have gone this route.

Two other approaches sometimes used are the "minimum size" and "zero-intercept" methods. The "minimum size" method operates, as its name implies, on the assumption that there is a minimum-size distribution system capable of serving customers' minimum requirements. The costs of this hypothetical system are, so the argument goes, driven not by customer demand but rather by numbers of customers, and therefore they are considered customer costs. The demand-related cost portion then is the difference between total distribution investment and the customer-related costs. The "zero-intercept" approach is a variation on the "minimum size." Here the idea is to identify that portion of plant that is necessary to give customers access but which is incapable of serving any level of demand. The logic is that the costs of this system, because it can serve no demand and thus is not demand-related, are necessarily customer-related.[36] However, the distinction between customer and demand costs is not always clear, insofar as the number of customers on a system (or particular area of a system) will have impacts on the total demand on the system, to the extent that their demand is coincident with the relevant peak (system, areal, substation, etc.).

Any approach to classifying costs has virtues and vices. The first potential pitfall lies in the assumptions, explicit and implicit, that a method is built upon. In the basic customer method, it is the *a priori* classification of expenditures (which may or may not be reasonable). In the case of the minimum-size and zero-intercept methods, the threshold assumption is that there is some portion of the system whose costs are unrelated to demand (or to energy for that matter). From one perspective, this notion has a certain intuitive appeal — these are the lowest costs that must be incurred before any or some minimal amount of power can be delivered — but from another viewpoint it seems absurd, since in the absence of any demand no such system would be built at all. Moreover, firms in competitive markets do not – indeed, can not – price their products according to such methods: they recover their costs through the sale of goods and services, not merely by charging for the ability to consume, or "access."

Other assumptions are of a more technical nature. What constitutes the "minimum" system? What are the proper types of equipment to be modeled? What cost data are applicable (historical, current installations, etc.)? Doesn't the minimum system in fact include demand costs, since such a system can serve some amount of demand? The zero-intercept method attempts to model a system that has no demand-serving capability whatsoever, but what remains is not necessarily a system whose costs are driven any more by the number of customers than it is by geographical considerations, whose causative properties are neither squarely demand- nor customer-related. Does use of an abstract minimum system place a disproportionate share of the cost burden on

36. It is called "zero-intercept" because it relates "installed cost to current carrying capacity or demand rating, creat[ing] a curve for various sizes of the equipment involved, using regression techniques, and extend[ing] the curve to a no-load intercept." NARUC, p. 92.

certain customers or classes, in certain cases even resulting in double-counting? The answers chosen to these and other questions will have impacts upon the respective assignments (by type and customer class) of costs.[37]

Historically, the investment decisions of system planners in vertically-integrated utilities were constrained by the least total cost objective: simply, that they would make that combination of investments that were expected, given their assessments of risk, to meet expected demand for service over some reasonable planning horizon. Given the inability to store electricity and the typical obligation to serve all customers *on demand*, a utility was required to have sufficient capacity available to meet peak demand. And, if its only obligation were to meet peak demand, then it would install only the most inexpensive capacity. However, it had also to serve energy needs at other times, and it is a general characteristic of electric generation technology that as capacity costs decrease variable operating costs increase. There is, therefore, a trade-off between capacity and energy costs that system planners considered when building (or purchasing) new capacity, if they hoped to minimize total costs. Put another way, significant portions of generating capacity were purchased not to meet demand, but to serve energy, when the fuel cost savings that the more expensive generation would produce were greater than the additional costs of that capacity. These incremental capacity costs were therefore correctly viewed as energy costs.

A similar kind of analysis can inform the design of distribution systems, as it also does transmission. The question is whether there is some amount of capacity in excess of the minimum needed to meet peak demand that can cost-effectively be installed. The additional capacity — larger substations, conductors, transformers — will reduce energy losses; if the cost of energy saved is greater than that of the additional capacity, then the investment will be cost-effective and should be made.[38] For the purposes of cost analysis and rate design, these kinds of distribution investments are rightly treated as energy-related.[39]

b. Cost Allocation

As a general matter, distribution facilities are designed and operated to serve localized area loads. Substations are designed to meet the maximum expected load of the distribution feeders radiating from them. The feeders are designed to meet at least the maximum expected loads at the primary

---

37. Sterzinger, George, "The Customer Charge and Problems of Double Allocation of Costs," *Public Utilities Fortnightly*, July 2, 1981, p. 31; see also Bonbright, p. 347-348.

38. Losses vary with the square of the load. We note also that there is some minimum amount of losses that cannot be avoided, and that conductors must be sized such that the losses can be absorbed while still meeting peak load. To this degree, losses impose a capacity, rather than energy, cost.

39. An unhappy consequence of separating distribution and transmission planning from that of generation in restructured markets is the potential loss of this capacity-versus-energy consideration when making new investment. Certainly, without some sort of regulatory or legislative requirement, wires-only companies have no generation cost-savings motive to guide their planning decisions.

and secondary service levels. (As noted above, some investment in distribution capacity may be seen as reducing energy losses rather than serving peak demand.) For costing purposes it is the relevant subsystem's (substation, feeder, etc.?) peak that matters, but these peaks may or may not be coincident with each other or with the overall system's peak. There can be significant variation among them. Consequently, one practice is to allocate the costs of substations and primary feeders (which usually enjoy relatively high load factors) to customer class non-coincident peaks and to allocate secondary feeders and line transformers (with lower load factors) to the individual customer's maximum demand.[40] In addition, costs are allocated according to voltage level; customers taking service at higher levels are typically not assigned any of the costs of the lower-voltage systems that do not serve them. Costs are then allocated among customer rate groups (or classes) which requires, among other things, information and judgments about coincidence of demand when customers of different classes share facilities, as is often the case.

3. Marginal Costs

For the reasons stated earlier, it is the long-run marginal cost that is most relevant to designing rates. It can be described as the cost of that "lumpy," geographically dispersed set of investments that a utility must make if demand continues to grow after the distribution system has initially been built out.

a. Demand and Energy

As already noted, the drivers of distribution costs are typically seen to be peak demand (itself driven by both customer demand and numbers of customers) and energy needs.[41] For the purposes of marginal cost analysis, it is also necessary to identify investments that are not made to serve incremental demands, but are made for some other purpose — reliability, replacement of existing systems, etc. The costs of these investments are generally not included in marginal cost calculations, although, in certain cases, there may be legitimate arguments to the contrary.[42]

---

40. Class non-coincident peak may not be the best measure of cost causation, since much of the system serves a variety of customer classes. Chernick, Paul, *From Here to Efficiency: Securing Demand-Management Resources*, Vol. 5, 1993, p. 81. Ideally, the object is to design rates that reflect the costs of customers' contributions to the relevant peak.

41. It is worth noting that, in the short run, distribution costs vary more closely with numbers of customers than with load (except in capacity-constrained areas). For rate design, with its focus on the long run, this fact need not be a distraction. It does, however, have implications for setting revenue requirements. We address this question in Chapter V, below.

42. For instance, at the time that an investment to replace existing facilities (whose loads, let us say, are not expected to change over some extended period) is being contemplated, there are costs that can potentially be avoided. In the extreme, replacement would be unnecessary if all customers served by the facility were to decide to go "off grid." Other, more likely alternatives involve combinations of end-use efficiency, distributed generation, and smaller, more efficient distribution technologies. On these bases, the marginal or, more reasonably, the larger

(continued...)

Many of the same cost classification and assignment questions that pervade embedded cost analyses also recur in marginal cost studies, although their answers have different analytical effects. Whereas an embedded cost study strives to identify and assign total historical costs to classes of service (on the basis of any of a number of principles, including cost causation and fairness), a marginal cost analysis aims to determine the cost consequences of changes in output and thus the value of resources that must be used to serve incremental demand. Therefore, costs that are unaffected by changes in output (which describes all common and many joint costs) are excluded from the costs under examination.[43]

The study period for a marginal cost analysis is forward-looking and should be of sufficient duration to assure that all incremental demand is related to the investments forecast to serve that demand: a mismatch of timing and investment could result in significantly over- or understated costs. Those incremental costs are then discounted to their present value and annualized over the planning horizon. This has the effect of smoothing out the "lumpiness" of investment in relation to changes in demand.[44] This analysis relates changes in total costs to changes in demand (aggregating demand increases caused by the addition of customers with those caused by increases in demand per customer).[45] Since new customers create additional demand, this approach is not unreasonable.

Even so, some jurisdictions consider certain costs customer-related and treat them separately for the purpose of marginal cost analysis. Customer premises equipment – that which is dedicated specifically to individual customers and unrelated to variations in demand (meters and perhaps service drops) – are probably the only distribution costs that can be directly assigned to customers (except in the cases of customers who have additional facilities — transformers, wires, even substations, dedicated solely to their needs).[46] Some jurisdictions also consider other

42. (...continued)
incremental costs of distribution can be calculated. If replacement of the particular component of the system is forecast for some time in the future, then its expected future costs would need to be discounted appropriately to yield a present-value incremental cost.

43. Because marginal cost is defined as the change in total cost arising from a change in output, all costs are, strictly speaking, included in the analysis. It just happens that most are netted out, to reveal those that are caused by the change in output. As a practical matter, however, an analyst may simply identify the costs that vary with output and exclude the rest. It is this second approach, however, that raises debates about the nature of costs and whether they should be included in the analysis. Are they joint or common? Do they vary with demand, energy, customers, or not at all? Resolving the issues usually requires large doses of judgment.

44. An alternative approach is to calculate the cost (savings) of advancing (deferring) by one year the planned stream of investments to meet the increment (decrement) in demand. This approach yields a cost that is equal to the value of the marginal investments for one year (which is the same as the economic carrying charge on those investments). This method is often used, for example, to determine an annual cost per kW of generating capacity.

45. For sizing much of the distribution system, demand is the critical factor. One customer contributing six kilowatts to peak demand has the same impact as two each contributing three kilowatts.

46. After the meter, the customer service drop is typically seen as the least demand-related component of the system: it is sized to exceed any realistic maximum demand that the consumer might impose and it will last a very long time. However, although it is true that no investment would be made unless a customer were present, it is also
(continued...)

facilities (line transformers, secondary level conductors) in some measure customer-related, but, to the extent that they are jointly-used to serve more than one customer, it may be difficult to establish that the addition or loss of any one customer will affect the costs of those facilities.[47] In any event, if some costs are deemed marginal customer costs (which means that they are avoidable only at the time of hook-up), it by no means follows that they should be recovered in recurring monthly fixed fees (see Section V.A.5., below).

Other approaches sometimes used to resolve the cost-causation question are the "minimum system" and "zero intercept" methods. Here, instead of using embedded cost data, the distribution system is modeled to determine the cost (in current dollars) of a hypothetical system that could serve all customers' minimum demand or (in the case of zero-intercept) that could provide voltage but not power.[48] This cost would be deemed customer-related and separated from the total incremental cost previously determined, to identify the demand (or, more properly, the demand- and energy-related) portion. For the reasons stated earlier, we challenge the wisdom of these approaches.[49]

Other methodological difficulties may also arise. By definition, joint and common costs are not marginal, but occasionally they creep into the analysis, when, for example, they make use of what are in effect *average*, not *marginal*, investments and expenditures.[50] And, as with embedded costs, marginal costs are typically broken out by customer class. Here, again, the analysis requires reasonable assessments of the coincidence of demand, when customers of different classes share facilities.

46. (...continued) true that the amount of the initial investment increases as the customer's forecasted load increases. Thus, customer investments can be seen as demand-related, as can investments farther up the system – transformers, wires, and substations – whose sizing depends on expected peak demand. Bouford, James D., "Standardized Component Method for the Determination of Marginal and Avoided Demand Cost at the Distribution Level," Central Maine Power Company, (unpublished and undated), pp. 3-4.

47. NARUC, p. 136.

48. A handbook published by the National Economic Research Associates (NERA), which is often cited in support of the "minimum system" distribution cost classification, states that only the labor costs necessary to put together a minimum system – and no conductor and transformer costs – are customer-related. NERA, "How To Quantify Marginal Costs: Topic 4," (prepared for the Electric Utility Rate Design Study, March 10, 1977), pp. 76.

49. California, for instance, has rejected the "minimum system" approach to marginal costs, favoring instead a method which uses the weighted average of the costs of continuing to serve existing customers and the costs of initiating service to new customers.

50. See, *e.g.*, NARUC, p. 127, which notes that, because calculating marginal distribution and customer costs can be difficult, "it is still common for analysts to use some variation of a projected embedded methodology for these elements, rather than a strictly marginal approach." This tack is justified by the sweeping assumption that projected embedded distribution costs are a reasonable approximation of marginal costs. The assumption is, however, contestable. FERC accounting requirements, which form the basis of most embedded cost analyses, include in distribution certain, and often substantial, administrative and general (A&G) costs (Accounts 920 to 935). A&G is not caused by the provision of distribution service.

Another dimension of cost, and perhaps most revealing, is the geographic. There are several aspects to it. First are the topographical and meteorological characteristics of the area over which the distribution system is laid. Elevations, plant life, weather, soil conditions, and so on all have effects on costs. So too demography, which is captured partly by demand and numbers of customers, but also affecting costs is the density of customers in an area (sometimes expressed as customers per mile). These influences combine in assorted ways, with themselves but also with changes in load and rates of investment, to produce variations in costs from one area of the distribution system to another. It is not unusual to see marginal distribution costs varying greatly from one place to another, even when the distances between the different areas is comparatively short. Table 1 describes the significant variations in costs for incremental distribution investments in a large mid-western utility.

	Average System Marginal Costs per kW	Area Specific High-Low Marginal Costs per kW	Annual Cost @ 15% Capital Cost Recovery Factor	Average Marginal Costs per kWh @ 20% Load Factor[51]	High Marginal Costs per kWh @ 20% Load Factor
Transmission	$230	NA	$34	$0.02	$0.04
Distribution Lines	$960	$1,575 - 0	$140	$0.08	$0.135
Distribution Transformers	$60	$300 - 0	$9	$0.0015	$0.025
Total	$1,250	$1,875 - 0	$183	$.1015	$0.20

Table 1

Differentiating marginal costs along these lines will tell a utility where investment (whether in new facilities, end-use efficiency, or distributed generation) is needed and what the minimum value of that investment is. Whether for rate-making purposes this information is useful – should distribution rates be geographically "deaveraged"? – is a tougher question. We take it up in Chapter V, below.

51. This is estimated load factor for the incremental distribution investment alone, not for the entire distribution system altogether. Incremental investment to meet peak needs typically manifests low load factors; 20% is a conservatively high estimate.

4. Key Concern in Determining Costs: Follow the Money

The occasionally technical and arcane matters taken up in embedded and marginal cost studies are, of course, important, but it is perhaps more important to bear in mind that, in rate design cases, what is fundamentally at issue is who should bear what revenue responsibilities. In the interplay between cost allocation and rate structures, the debate over money is played out. First is the question of what costs will be categorized as *distribution*, as opposed to transmission or generation in the case of vertically integrated utilities, or perhaps competitive services in other instances. This is no small matter, since significant portions of a firm's joint and common costs (typically, administrative and general) are often attributed to the distribution business, even though there is no causal relationship between them. Then there is the designation of a cost as either *customer* or *demand*, which will affect both how costs are divvied up among classes and who within each class will pay them (*i.e.*, both inter- and intra-class allocations). While there is a touch of cynicism in the observation that there is no shortage of academic arguments to justify particular outcomes, it is nevertheless largely true. Always be aware of the revenue effects of a particular rate structure. Who benefits, who loses? Fixed prices, because they recover revenues by customer rather than by usage, invariably shift a larger proportion of the system's costs to the lower-volume consumers (residential and small business). The positions that interested parties take with respect to rate design should, in part, be considered in light of their impacts on class revenue burdens and on the profitability of the utility. Here the admonition to be practical cannot be stressed enough. Seemingly small changes in a rate design can have very significant consequences for different customers.[52]

---

52. Consider the following example (the hypothetical rates cover distribution services only). A residential customer using 500 kWh per month and paying $0.05 per delivered kWh and a monthly customer charge of $5.00 sees a monthly bill of $30. If rates were revised so that residential customers paid a fixed charge of $20 per month plus $0.02 cents per kWh, a customer using 500 kWh would receive the same total bill of $30. For this customer, the rate redesign is "revenue neutral." However, for a customer using 300 kWh/month, the monthly bill under the original rate structure is $20 and, under the new rates, is $26 – a 30% increase, even though there is no change in usage. For a customer using 700 kWh/month, the original bill is $40 and the revised bill is $34, a 15% reduction.

Consider again the customer using 500 kWh/month. If, under the original rate structure, she reduced her electricity use to 300 kWh per month (whether by load reduction, demand-side management, the installation of a rooftop solar electric system, or some combination of these options), she would reduce her bill by $10. However, under the revised rate structure, she would only reduce her bill by $4.

Whether the impacts of a rate design change are immediate and substantial depends, of course, on a variety of factors. The extent to which class cost allocations are altered will determine whether particular customers' total bills (all else being equal) will go up or down. Even those changes that are meant to be class "revenue-neutral" will affect individual customer bills: as already noted, shifts from usage-based to fixed charges recover disproportionately higher revenues from low-volume users and then, more subtly, there are the effects (both positive and negative) on bills and revenues that flow from demand responses to the changes in rate structure.

5. Usage Sensitivity: What's Avoidable?

a. Peak Demand and Sizing the Wires

Distribution investment is made to serve an expected level of demand over a period of time, often determined by the useful life of the equipment. To the extent that, once a network (or component of it) is built, there is excess capacity in it, the marginal cost of using that excess capacity will be quite low (possibly very close to zero, insofar as there is little in the way of variable cost). It is this phenomenon – that the short-run marginal cost of delivering a kilowatt-hour is zero – that underlies the argument that there should be no per-kilowatt-hour charge for doing so.

As peak load grows, it will press up against the capacity limits of the system. At the time of constraint, the marginal cost of delivering a kilowatt-hour is, in fact, significantly greater than zero: at a minimum it is the cost of the additional investment needed to carry that marginal kilowatt-hour to end-users.[53] At that point, presumably, the new investment is made, and it is sized to minimize the total costs of delivery over the long term and thus, as before, there is suddenly "excess" capacity causing once again the marginal cost to fall to almost zero.

This "non-linearity" of investment with demand is a characteristic of much of the distribution system, the closer one gets to the end-user. To the extent that there are not an infinite number of equipment sizes to enable precise matching of investment and demand, excess capacity is almost necessarily built into the system, from substation facilities to feeders, transformers, customer service drops. But this has less to do with the finitude of equipment options than it does with the least total cost planning objective (optimizing total construction and operations costs over the investment horizon). The analytical key is to view the system over a time period long enough to smooth out the "lumpiness" of investment in relation to changes in demand.[54]

What emerges from such analysis is the recognition that there are costs associated with load growth, savings generated by reductions in load growth, and savings flowing from reductions in existing load. These values, not necessarily equal to each other, reflect in part the fungibility of significant portions of the system (*e.g.*, substations and feeders). Capacity unused, or freed up, by one customer can be used by others.[55]

Sometimes cited as an interesting and somewhat anomalous characteristic of some distribution investment, specifically that closest to customers (such as the service drop) is its manifestation of positive marginal costs with load growth but seemingly zero marginal (or avoided) costs with

53. And it may indeed be greater, if the value to consumers of that marginal delivery is greater than the cost of the additional investment. See Appendix A.

54. The justification for analyzing costs over the long run, and for setting prices on that basis, is discussed in Appendix A.

55. Chernick, Vol. 5, p. 68.

load reductions. This is because, so the argument goes, load reduction makes no capacity available for alternative uses, that did not already exist. This not so, however, because the inability to re-use capacity does not mean that there is no value to not using it. At the very least, future replacement costs can be deferred and the equipment installed on replacement can be down-sized, thereby reducing costs for all users.[56]

The differences in costs and savings associated with load growth, reduced growth rates, and reductions in existing load may leave some room for debate about their implications for rate design; but, given the declining-cost nature of the distribution system, these differences will probably have less of an impact than will the need to recover an embedded revenue requirement. The critical point here is that distribution costs vary primarily with load over the longer term.

b. Energy: The Costs of Throughput

As discussed earlier, to the extent that distribution investments are made to offset energy needs, there are necessarily costs associated with avoiding those investments. Losses, heat build-up, frequency of overloads, etc., are aspects of energy use that affect distribution investment and operations; thus there are marginal energy costs in distribution. Whether avoiding those costs make alternatives to distribution cost-effective is an empirical question. But, for purposes of rate design, it is sufficient to say that these marginal costs should be understood and appropriately reflected in rates. They are unquestionably volumetric in nature.

B. Conclusion: The Costs of Distribution Services

Cost studies are intended to provide useful information about the causes and magnitudes of costs, to inform a rate design process that is guided by the general principle that those who cause a cost should pay that cost. However, the usual "drivers" ascribed to distribution costs (both embedded and marginal) describe only part of the story, and the force-fitting of square "costs" into round "drivers" can lead to rate designs that will not best promote long-run dynamic efficiency. This is especially true of embedded cost studies, in which a central objective is to assign or allocate costs to particular services or classes of customers, even though many of those costs cannot be assigned unequivocally according to the principle of causation. By their very nature, many utility costs are joint or common to two or more services; consequently there can be no unshakeable assertion that any one service in fact caused a cost and, therefore, that a particular rate element should recover it. And marginal cost studies often suffer from this deficiency as well. This means that regulators should be very careful before relying upon what are essentially (though not

56. *Id.*, pp. 68-71. Also affected is the magnitude and cost of over-sizing equipment in order to serve forecast demand. See also NERA, pp. 17-18.

necessarily unreasonable) arbitrary cost assignments for the purposes of designing rates.[57] Too great a dependence on cost studies is to be captured by their underlying assumptions and methodological flaws. Utilities and commissions should be cautious before adopting a particular method on the basis of what may be a superficial appeal. More important, however, is the concern that a costing method, once adopted, becomes the predominant — and unchallenged — determinant of rate design.

Marginal cost analysis demonstrates that distribution costs vary with load in the long run. This has important implications for rate design. Embedded cost analysis, though it relies on *a priori* assumptions about causes (and allocations therefore) of historical costs, is useful in rate design at least insofar as it informs the process of reconciling marginal cost-based rates with revenue requirements.[58] We recognize that there are honest disagreements over approaches to both kinds of analysis.[59] But what is important here is for regulators to be aware of the fundamental relationships between costs and demand for electric service, in order to devise rates that best serve the objectives they seek.

57. "To ensure that [embedded distribution plant] costs are properly allocated, the analyst must first classify each account as demand-related, customer-related, or a combination of both. The classification depends upon the analyst's evaluation of how the costs in these accounts were incurred." NARUC, p. 89. Interestingly, the manual, in a table on page 34, acknowledges that there is an energy-related component to embedded distribution costs, but is otherwise silent on the question.

58. Bonbright, pp. 366-367. Bonbright expresses some skepticism as to the usefulness of most embedded cost studies for rate design, on the ground that they often ignore the relationship between cost causation and apportionment. "One may suspect that the choice of [allocation] formula depends, not on principles of cost imputation but rather on types of apportionment which tend to justify whatever rate structure is advocated for non-cost reasons." *Id.*, p. 368.

59. See, *e.g.*, Chernick, Vol. 5, pp. 58-83, and NARUC, pp. 86-104 and 137-146.

V. SETTING RATES FOR DISTRIBUTION SERVICES

A. Distribution Pricing To Satisfy the Long-Standing Objectives of Rate Design

Prices in both regulated and competitive markets serve, in large measure, the same functions. In light of those objectives and in recognition of the unique cost characteristics of the wires network, how should prices for distribution services be set?

A distribution network is built mainly to serve the forecasted peak load and, to a lesser degree, the energy needs and numbers of customers it will serve. However, an acknowledgment that there are several dimensions of cost causation does not by itself dictate the most appropriate structure for rates (particularly when the rates will also recover costs that have no causal relationship to the service). Nor does the fact that a particular rate design may satisfy more or less a particular rate design objective: regulators must settle on rate structures that, in their view, best balance the competing goals they seek to achieve.

It also important to bear in mind what consumers seek. They want electricity delivered reliably. Therefore they purchase commodity delivery services, and they justly expect to pay a price per-unit delivered (either bundled in the commodity price or not), as they do for letters, packages, gasoline, and just about everything else traded in the economy. The conversion of the capital costs of delivering electricity into a cost per unit delivered has its analogues in virtually all other markets.

1. Efficiency and Fairness

It is our firm conclusion that usage-based unit prices, primarily energy-based and at least equal to the long-run marginal costs of providing service, best serve the several objectives of rate design. By recognizing the variability of distribution costs over time, such rates promote long-run, dynamic economic efficiency and, insofar as they eliminate cross-subsidies among consumers, they are fair. Moreover, they cause those who use more of the service to cover proportionately more of its costs – a second measure of fairness.

Since distribution costs are related more closely to demand rather than energy, it might be argued that demand charges are more appropriate. Such charges, in order to promote efficiency, would have to be linked closely to the relevant peak and allow for the savings from reductions in demand to accrue to customers. Demand rates that charge according to a customer's own peak, which may or may not bear any relation to the relevant peak, may reflect only an indirect link between customer demand and the need for capacity.[60] Another drawback of demand rates is that often they are "ratcheted" (monthly kilowatt demand is calculated as a percentage of the customer's annual peak), which tells the customer that it is annual demand, not monthly, that matters. Ratchets have the effect of fixing charges for an extended period, thereby muting the price signal and consumer response to it. This is not true of competitive generation markets, nor it is necessarily so of distribution service. It is better to encourage month-by-month reductions in usage through methods – more efficient equipment, limited heating and cooling settings, etc. – which, in turn, are likely to limit the need for distribution system reinforcement and expansion.

Demand charges, though variable, are by themselves less preferable to at least a combination of demand and energy charges. The rationale for two-part rates, that they distinguish between those costs that vary with peak demand (capacity) and those that vary with energy needs, has certain attractions. Such rates have long been in effect for high-usage customers of vertically-integrated utilities, and they may very well make sense for certain higher-volume (primarily commercial and industrial) purchasers of distribution services, who are better able to adjust their demand to improve their load factors. This rate design also confers benefits upon the utility, so long as the demand being measured and charged for is the customer's contribution to the relevant coincident peak (areal or substation in the case of distribution).

For lower-volume consumers, energy-based charges, differentiated as appropriate by time of use to reflect capacity constraints, offer the same economic incentives as demand or two-part rates, are immediately avoidable, and are administratively simpler, particularly for lower-volume users.[61] Indeed, energy-only charges for residential and small business users have predominated throughout the century, and there is nothing about the separation of distribution from generation that suggests that a radical shift in rate design policy is warranted.[62] In any case, whether one-part or two-part in structure, it is important that the rates not be fixed irrespective of consumption

---

60. A similar complaint can be made about declining block energy rates, which may be justifiable on the basis of incremental energy costs, but send no relevant signals about the costs of consumption on peak. (This drawback of declining block rates in distribution can be distinguished from volume-based discounts in competitive industries, such as home heating oil and propane, insofar as the rates in the competitive markets do not mask a capacity constraint in the delivery of the commodity.)

61. Two-part demand and energy rates have, of course, a long history in vertically-integrated firms. Kilowatt and kilowatt-hour charges were implemented chiefly to reflect the costs of generation capacity and operations, although of course distribution and transmission costs were covered as well. What's interesting in the shift to competitive generation is that spot markets are primarily energy-based. Kilowatt-hours, not kilowatts, have emerged as the key commodity, and their prices vary on an hourly (or half-hourly) basis, reflecting the value of capacity at times of increasing load. The apparent unsustainability of two-part rates in competition shouldn't be surprising. "Since multi-part rates have the effect of making the same product available at different unit rates, they can only be charged by a firm with monopoly power for a product that cannot practically be resold. Many regulated firms sell under these conditions." Pierce and Gellhorn, p. 205. Competition inhibits such behavior.

62. In US competitive generation and retail markets, the costs of bulk delivery – transmission – are generally included in the prices end-users pay for electricity, which is to say that in many cases transmission costs are recovered on a commodity basis, as we are recommending here for distribution. In New Zealand, retailers purchase distribution services at wholesale, and recover those costs in the electricity prices they charge their retail customers.

levels: saving energy, improving the efficiency of energy use, should be rewarded by the avoidance of charges.[63]

Fixed recurring charges fail to satisfy the fairness and efficiency objectives. They do not recognize either cost or value-of-service differences among customers, particularly as the differences pertain to usage volumes. Fixed charges discourage customer investments in cost-effective alternatives to consumption; in effect, they say to customers "Nothing you do matters at all." Since any signal that there are costs to additional consumption is lost, use of the network appears to be free. Consumption will uneconomically increase, until available capacity is exhausted, thus creating the need for new investment that, under a volumetric pricing regime, would have been avoided. This pattern will then repeat itself, as the costs of the new investment are averaged across all consumers through fixed charges. In addition, such fees shift a disproportionate share of the network (and other) costs onto lower-volume users, as well as from on-peak to off-peak users. In these ways, fixed charges violate both the efficiency and fairness criteria of a sound rate design.

Fixed prices will remove utilities' incentive to promote usage and will also remove their disincentive to promoting cost-effective end-use energy efficiency and conservation, since changes in usage levels will not affect companies' revenues. However, these virtues are not by themselves enough to recommend implementation of such rates, given their other drawbacks and given that there are other ways to address the incentives problem without sacrificing the consumer and economic benefits of usage-based prices. See Subsection A.2., below.

a. Reliability

The failure of pricing that leads to over-investment in distribution assets also threatens, in several ways, system reliability. There are many dimensions to reliability. Some reliability problems are distribution level failures attributable to localized overloads. Typically, distribution rates (whether or not bundled with generation and transmission) are averaged across peak and off-peak hours and therefore do not well inform consumers of the limits of capacity. Such rates wrongly suggest that on-peak consumption is less costly than in fact it is. This problem will only be exacerbated by a shift to a fixed pricing regime, which offers consumers no information at all about the scarcity and costs of distribution capacity. Reliability will be further degraded. As Table 1 (see Chapter IV) showed, the costs of maintaining reliability, simply through additional construction, can be very high. Sensible rate design can go a long way to avoiding those costs.

Improper distribution rate design can have adverse effects that extend beyond the local wires network. To the extent that distribution rates do not cover at least the marginal costs of delivery,

63. To the extent that some small amount of revenue is, or continues to be, recovered in monthly customer charges, these should reflect the long-run marginal customer costs. As stated earlier, we believe these to be at most related to service drops, meters, and billing.

the total costs – from the consumer's perspective – of energy service (generation, transmission, and distribution) will be understated, and will lead to uneconomic additional consumption. Depending on the coincidence of such consumption with other system constraints, the potential for deterioration of overall system reliability is increased.

2. Revenue-Related Criteria

What impacts will a particular rate structure have on company revenues? How does it affect the allocation of risk between customers and the utility? Whether rates yield the total revenue requirement, provide predictable and stable revenues, and are themselves stable and predictable has less to do with their structure (usage-based or fixed) than it does with rate levels themselves, utility management, and the regulatory environment in which the utility operates.

As a general proposition, usage-based rate designs reward firms for increased sales.[64] This is because, in the short run, a company's costs do not vary directly with kWh consumption. Revenues from incremental sales go to the bottom line. This provides a powerful incentive for the utility to promote sales, even when it is economically inefficient and environmentally damaging. Steps should be taken by regulators to remove that incentive, to break the link between sales and profitability.

There are other nuances to the recent debate over fixed prices for distribution services. For the utilities, the central issues are revenue growth and risk reduction. While it is true that, without some sort of regulatory mechanism that "de-couples" profits from sales, a company will benefit from increased sales, it may, under certain circumstances, be more profitable to link revenues instead to customer growth. And, even where it may not be, a utility may perceive that its overall level of business risk is reduced by a fixed pricing structure.

Do fixed recurring rates really solve the revenue problem that some distribution utilities think they have? What effect upon revenues, and the need for rate cases, will a fixed recurring charge have for a company whose sales are growing steadily? Consider the case of the utility whose usage per customer is declining, but whose number of customers is increasing. If usage is declining by 2% and customers are increasing by 2%, then with usage-based pricing revenues remain constant. With fixed prices, however, revenues – and profits – go up. If the opposite is occurring – sales increasing while customers decrease – then volumetric pricing will provide the higher revenues. And, of course, there are equivalent down-side revenue risks.

The freedom of a company and the commission to seek rate adjustments as needed is the traditional means of mitigating these risks. Another approach is a performance-based, per-customer revenue cap. It rewards a firm for increases in efficiency, while making it, at the very least, indifferent to the volume of throughput over its wires. Since, in the short-run, a

64. Moskovitz, David, *Profits and Progress through Distributed Resources*, NARUC, 2000, pp. 16-18.

distribution company's costs vary more closely with numbers of customers than with load growth, a per-customer revenue cap would produce revenues that more closely track annual costs. To the utility, a per-customer revenue cap looks just like a fixed-price rate structure, and it renders the company at least indifferent to customer installations of efficiency and other distributed resources. However, the revenue cap enables prices for end-users to be set on a usage basis, thereby enabling them to make consumption decisions and alternative energy investments that are, in the longer term, most efficient.[65] In addition, whereas under traditional regulation fixed prices remain in place until changed by commission order, a performance-based rate-making scheme can adjust rate levels automatically to encourage and capture efficiency increases over time, thus returning some measure of savings to consumers.[66]

3. Cost-Related Criteria

We have already established that usage rates set, at a minimum, at long-run marginal cost, differentiated if appropriate by time of use, serve the twin objectives of economic efficiency and fairness. Fixed prices do not reflect customers' contributions to peak demand or energy needs, the primary drivers of distribution investment. As for apportionment of the overall cost-of-service among customers (insofar as marginal-cost-based rates do not generate sufficient revenues) and the avoidance of undue discrimination, no method of rate design can by itself assure fair outcomes. What is appropriate, or fair, can only be determined by regulators, in light of the facts in each case.

a. Environmental Externalities

Ideally, rates should reflect the present and future private and social costs and benefits of providing service (*i.e.*, all internalities and externalities should be included in rates). What this means, of course, continues to be the subject of much debate. But worth considering in the context of distribution rate design, particularly in those markets that have introduced competitive generation and retail electric services, is the question of whether prices fail to serve this objective and, if so, what steps can be taken to correct the problem.

To the extent that generators incur costs to reduce the environmental effects of their production processes (*e.g.*, emissions controls, waste disposal, etc.), some measure of the cost of damage is included in the electricity prices. However, there remain significant environmental damages from electricity generation that go unpriced; they have costs that are paid, not in prices, but

65. *Id.*, pp. 20-22.

66. Whether and how this works depends, of course, on the mechanics of the performance-based regulatory plan, specifically on the relationships among the various adjustment factors (inflation, productivity, exogenous).

through their effects (reduced public health, acid deposition, etc.).[67] Put another way, the marginal environmental costs of generation, which are largely associated with fuel consumption and therefore are directly correlated to kilowatt-hour production, are not reflected in current prices for electricity.

Because generation markets do not internalize all the costs of production, it falls to regulators and policymakers to correct the failure. Volumetric pricing for distribution services, appropriate for the reasons already stated, is also justified on the ground that there are incremental kilowatt-hour costs that commodity prices fail to capture; in this way, the mark-up on usage-based distribution charges needed to cover the embedded revenue requirement serves as a proxy for some portion of the environmental damage costs of production. Whether the mark-up is sufficient to "cover" those damage costs and whether additional mitigation efforts are warranted remain, of course, questions policymakers must grapple with.

4. Practical Considerations

Usage-based rates are well-understood by consumers. They are, for the most part, uncomplicated and can be easily administered. Fixed prices share these attributes.

5. Other Issues

a. Customer Charges

One kind of fixed charge has long been a fixture of utility pricing: the monthly (or daily) customer charge. In most jurisdictions, recurring periodic rates designed to cover at least the costs of metering and billing serve to generate a stream of revenues that does not vary with usage and thereby provides some measure of financial risk mitigation for the utility. For residential customers, these charges range from as little as a dollar to ten dollars or more per month. For commercial and industrial customers, they can be considerably greater.[68]

The current debate about pricing for distribution services really comes down to a simple question: should customer charges be increased and usage charges decreased (or even eliminated) and, if so, by how much? Our inquiry concludes that, for the most part, the answer is no, and even suggests that it may be appropriate in certain cases to reduce customer charges. Of course, decisions taken by regulatory commissions will be based on the particular facts of each case; our

67. Competitive commodity markets for electricity do not capture these costs in prices; nor are they typically reflected in marginal cost studies in those states where the industry remains vertically integrated.

68. One variation of the customer charge is the "minimum bill" approach, such as that used by Central Maine Power (see Section II.C.3.), which requires payment of a monthly charge, but with it also comes a specified number of "free" kilowatt-hours of delivery service. Delivery in excess of the allowance is billed on a per-unit (kWh) basis.

intention here is to examine the various policy considerations and potential consequences of different actions.

We do not foresee an outright elimination of customer charges, although, as competition in the industry grows and alternatives to grid-provided power become more cost-effective, we believe that they will become less and less tenable. The rate-making principles that counsel against the imposition of fixed charges also discourage radical and immediate changes in rate design. Nominal customer charges have been around a long time. They are well-understood by consumers, and they provide some revenue stability for utilities. Any change in rate design should be deliberate, to minimize potentially deleterious impacts on customers and companies.

In evaluating proposals for redesign of distribution rates, commissions may be asked to consider structures that call for some blend of customer and usage charges, weighted so as to increase the revenue share of the fixed rate elements (in relation to historical allocations). Although much of the discussion in this paper has been cast in "either-or" terms (usage-based vs. fixed rates), its general prescriptions apply no less to any intermediate proposal: the magnitude of a shift from usage-based to fixed rate elements will have predictable effects on consumer demand, utility revenues, and long-term dynamic efficiency. As one moves along the continuum of rate designs from usage-based to fixed, the benefits of the former give way more and more to the difficulties of the latter. This is the kind of trade-off that commissions are often faced with balancing: our analysis concludes that the balance strongly favors a rate structure that allows consumers to avoid charges, when there cost-effective alternatives that they value more highly. Usage-based rates fit this bill; so do "hook-up fees" (see the following section).

b. Customer Costs and Hook-Up Fees

In recognition of the dedicated nature of customer-related facilities (meters and service drops), regulators might consider an alternative rate structure for recovering their costs. As discussed earlier, marginal customer investment costs can be distinguished from other utility marginal costs of service, insofar as they are only avoidable at the time that the facility is installed or replaced. In a competitive market, a customer would pay the prevailing price of purchasing the hook-up at the time of installation, which would approximate marginal cost. This is the way in which consumers purchase many durable goods which are affixed to their premises and have no other uses apart from the premises (curtains, ceiling insulation, etc.). Consequently, it may be more economically efficient to recover the costs of access equipment in the form of a customer "hook-up" fee.

The revenue impacts of this charge should be carefully considered . If hook-up fees are to be implemented, it is critical that double-counting of costs be avoided. Regulators must be careful to assure that these costs, if recovered in a hook-up fee, are not also included in other distribution charges.

c. Stand-By Rates

For customers who provide much of their own power but who remain connected to the grid for reliability purposes, the question of how to set "stand-by" rates for distribution service arises. As with distribution generally, the nature of the costs incurred becomes the point on which the discussion turns. Proposals for fixed rates are based on the argument that self-generators merely need "access" (an unvarying cost) and that, in the absence of any meaningful throughput over the wires, the capital costs will go unrecovered.[69]

Standby charges originally were intended to cover the direct economic costs to a utility of maintaining the otherwise-unutilized capacity necessary to provide service in the event that a customer's on-site generating facility experiences an unanticipated outage. Stand-by rates should, like all rates, reflect at least the costs of providing service. To the extent that customer generation avoids costs, those savings should offset the costs of stand-by. What is relevant here is the likelihood of the self-generator demanding service at times that will contribute to an increase in distribution capacity needs. It is a straight-forward enough process to calculate rates that reflect the probability of a self-generating customer contributing to peak needs (*i.e.* causing costs), in much the same way that rates for interruptible service are determined.[70] Such rates can be differentiated by time of use, on either an energy or capacity basis.

d. High- and Low-Cost Areas: Geographically De-Averaged Buy-Back Rates

Ultimately, prices are in some manner averaged, at the cost of some efficiency, but with the benefits of administrative simplicity and added fairness (in much the same way that we consider flat pricing-per-unit for letters and postcards to be fair and societally beneficial). But the distribution costs that a utility faces at any point in time can vary significantly from area to area.

---

69. Morey, Matthew J., "Distributed Generation: Is It the Wave of the Future," Edison Electric Institute, presentation at the mid-year meeting of the National Association of State Utility Consumer Advocates, June 5, 2000.

70. It would be appropriate to treat self-generating "customers" as a distinct class for the purposes of setting standby charges. Calculating the probability of coincident peak demand would, therefore, be done on a group, rather than individual customer, basis. Whether self-generators should then be further differentiated along traditional lines (*e.g.*, residential, commercial, industrial) would depend upon whether such groupings are predictive of particular and distinguishable usage characteristics.

Defining standby customers may, in certain cases, be a tricky matter. It is, arguably, inappropriate to deem owners of generating facilities that produce power intermittently (*e.g.*, wind, solar, etc.) to be standby customers, since their fluctuations in demand are often comparable to the fluctuations in demand of non-generating customers. It is the utilities' business to respond to such usage patterns, which they routinely do under the terms of their existing tariffs and rate structures.

In areas with excess capacity, the marginal costs of distribution may be very low, but in constrained areas the marginal costs of incremental consumption can be exceedingly high.[71]

Where the marginal costs of distribution are high, the utility has a strong incentive to invest in less costly means of providing service: end-use efficiency, distributed generation, and load management, for instance. This is particularly true where, as in most areas, the retail rates for service are averaged, and marginal on-peak costs exceed marginal revenues. In such circumstances, utilities have a very palpable profit motive to reduce costs. Customers, in contrast, do not. They are not being given price signals that reflect the full marginal costs of service, at least at times of peak, and consequently their incentives to invest in distributed resources are muted. And, if they are paying fixed prices, the incentives are non-existent altogether.

One response is to de-average distribution prices, according to location. However, assuming that the geographic de-averaging of prices is not possible, alternative approaches for promoting economically efficient outcomes must be developed.[72] One such approach is the geographically de-averaged "buy-back" credit. The utility would establish financial credits for distributed resources installed in a given area. The credit amount would be a function of the distribution cost savings generated by the distributed resources. Credits would be limited in duration and magnitude, in order to match the timing and need for distribution system reinforcements. For example, credits might be available to the first 20 MW of distributed resources installed in the next year because, after that period, loads are expected to have grown to the point that distribution line upgrades are unavoidable. The dollar amount of the credits should, at most, equal the value (savings) derived from deferring the distribution upgrade. Credits would also vary by location of the distributed resources. Credits would be highest in areas of greatest need and would be as low as zero in low-cost areas.[73] For example, customers in an area with 20¢ distribution costs might be offered a 15¢ credit.[74] This would certainly produce a strong economic incentive for customers and others to invest in distributed resources. Because the credit

71. In fact, at times of total capacity utilization (or overloading), they can exceed the cost of building new facilities. This is because the loss of value that consumers who are prevented from consuming at that time may be very much greater than the cost of new investment. This of course is true also of competitive markets; indeed this characteristic of supply and demand is a critical determinant of new investment.

72. To the economist, differentiating prices according to geographic cost characteristics is no different than doing so according to time of use. However, in light of the potentially very great differences in rates from area to area, the administrative complexity of the rate structure, and universal service considerations, we are unlikely to see geographically de-averaged rates any time soon.

73. Variations of the de-averaged distribution credits could be a sliding scale standby rate or a hook-up "feebate." For example, stand-by rates could be on a sliding scale ranging from high to negative. Negative stand-by rates, which look like distribution credits to customers, would be charged in high-cost areas. A hook-up feebate would be a revenue-neutral charge that collects from customers installing distributed resources in low-cost zones and pays to customers who install distributed resources in high-cost zones.

74. Demand-side resources are so much less costly that the winning bid prices would likely be far below 15¢.

is 15¢ instead of the 20¢ the utility would incur to upgrade facilities, there is an opportunity for savings to be shared.[75]

B. Conclusion: Setting Rates for Distribution Services

In sum, we urge regulators to adopt pricing and rate-setting policies that will serve the longer-term public interests: fairness, economic efficiency, competitive provision and innovation, and environmental protection. In the distribution system, this calls for usage-based pricing – primarily volumetric (energy-based) but, where appropriate, both demand- and energy-based. Also, we recommend that policymakers implement revenue-cap performance-based regulatory schemes for distribution companies. In so doing, the firms would obtain much of the benefit that they seek through fixed pricing, while consumers and the economy overall would still retain the benefits that usage-based, competitive pricing provide. We note, however, that a revenue cap or similar mechanism should not be seen as a necessary prerequisite to the usage-based pricing structure.

75. Moskovitz, p. 24.

APPENDIX. THE ECONOMICS OF REGULATION

Most people who have ever tried their hands at designing rates for regulated utilities invariably say that it is more art than science, an aphorism that has become almost trite from overuse. That rate design is an endeavor that depends heavily upon the collective judgment of regulators, utility officials, advocates, and other policymakers makes it an art without doubt, but also it relies in strong measure upon the judicious application of certain scientific principles – those of economic theory. Those principles emerge from a comparative analysis of competitive markets and monopolistic ones, and they reveal the opportunities for, and limits of, particular pricing policies. It is useful, therefore, to survey the fundamental economics of competitive and regulated industries, to uncover the bases of those policies.[76]

There are two broad, fundamental justifications for governmental oversight of the utility sector. The first is the widely-held belief that the sector's outputs are essential to the well-being of the society – its households and businesses – and the second is that its technological and economic features are such that a single firm often can serve the overall demand for its output at a lower total cost than can any combination of more than one firm. Competition cannot thrive under these conditions and, eventually, all firms but one exit the market. This is called "natural monopoly," and, like monopoly power in general, it bestows upon the surviving firm the power to restrict output and set prices at levels higher than are economically justified.[77] Economic regulation is seen then as the necessary and explicit public or governmental intervention into a market to achieve a public policy or social objective that the market fails to accomplish on its own.

1. Allocating Society's Scarce Resources

Economists are interested in discovering the elements and conditions of economic activity that will yield the greatest level of societal welfare (or "satisfaction"), given an *a priori* distribution of

76. "The unique set of tools that economics can contribute to the regulatory process is the familiar body of microeconomic theory, which purports to explain and predict the behavior of the individual consumer, investor, worker, firm, and industry under various circumstances. Like all other scientific models and the generalizations that emerge from them, the models of microeconomic theory are simplified, describing causal relationships involving a limited number of variables." Kahn, Alfred E., *The Economics of Regulation*, Vol I. (1988: The MIT Press, Cambridge, MA), pp. 16-17.

There are a number of insightful treatises examining regulation and rate-setting, a number of which will be referred to herein. Probably the preeminent among them are James Bonbright's *Principles of Public Utility Rates* and Alfred Kahn's, *The Economics of Regulation*. Bonbright's book, first published in 1961, was reprinted in 1988 in a revised second edition, authored by Albert L Danielsen and David R. Kamerschen (Public Utilities Reports, Inc., Arlington, VA). The two volumes of Kahn's study were first published in 1970 and 1971 (John Wiley & Sons, Inc., New York) and again in 1988, in a single bound volume with a new introduction.

77. See Kahn, Vol. I, pp. 11-12. Kahn adds that there may be reasons other than natural monopoly that prevent competition in a particular industry from working well.

wealth and income. Societal welfare is increased by maximizing economic efficiency: namely, by assuring that scarce resources are put to their most highly valued uses and are used most efficiently in production.

At any particular moment, an economy's productive capacity available to meet an increase in demand is fixed — additions to the existing set of assets (land, buildings, facilities, etc.) cannot be quickly procured, constructed, or in some other way altered — and thus that new demand must be met by increasing other inputs (labor, fuel, materials, etc.) to be combined with the existing capacity. The question then becomes how to make best — most efficient — use of that capacity. "The basic economic problem, in short, is the problem of choice."[78] In deciding to allocate resources to the production of particular goods and services, we are necessarily deciding *not* to allocate resources to the production of other goods and services, and "[i]t follows that the cost to society of producing anything consists, really, in the other things that must be sacrificed in order to produce it: in the last analysis, 'cost' is opportunity cost — the alternatives that must be foregone."[79]

In America, we prefer generally that individual consumers make their own consumption decisions: if those decisions are to yield the greatest welfare, then consumers must see prices that accurately reflect the opportunity costs of the goods and services available — that is, they must understand the value of the goods and services foregone in order to satisfy their particular demand, and decide whether the sacrifice is worth the purchase. Consumption thus informed will usher the economy's scarce resources into their most highly valued uses, that is, make most efficient use of them and thereby maximize welfare.

What then is the opportunity cost of a particular good or service? How is it measured? One of the great insights of economic theory is its equation of opportunity cost with the marginal cost of production. Because demand is responsive to price (which is to say that, as price changes, consumers' willingness to purchase also changes[80]), consumers' decisions to purchase more or less — a *marginal* amount — of a good will only be efficient if the price of the good reflects the actual — the *marginal* — cost incurred when more or less of the good is produced, and when the costs of other goods and services (*i.e.*, other consumption opportunities) also reflect their marginal costs of production.[81] Only then can fair comparisons be made and resources

---

78. *Id.*, p. 65.

79. *Id.*

80. This is referred to by economists as "own price elasticity of demand." It is the ratio of the proportional change in quantity demanded to the proportional change in price. Consumers whose demand is very sensitive to changes in price are said to have a high elasticity of demand.

81. The percipient reader will see this last condition as a very big "if," one that, given the realities of market imperfections, taxes, government spending, and so on, is never satisfied. It raises the question of "the second best:" how, in a world where not all prices equal marginal cost, should prices should be set so as to minimize the inefficiencies attending the distortions? Recognition of the problem, however, should not constrain policymakers to inaction; inoptimal organization elsewhere in the economy should not, except in limited and well-analyzed circumstances, deter pursuit of marginal-cost-based pricing policies in the utility sector. Kahn, Vol. I, pp. 69-70.

appropriately allocated. In this way, consumers, who make purchasing decisions based on the relative values that they assign to alternative uses of their own resources (income and wealth), will make consumption decisions that allocate society's resources to uses to which they assign greater relative weights. If a good is underpriced (priced below its marginal cost), then some quantity of the good will have been produced at a cost that exceeds its value to society, and the resources that were given to its production could have been allocated to better (more highly valued) uses elsewhere. The converse is true of over-priced goods.

A corollary to the conclusion that pricing at marginal cost will optimize society's exploitation of scarce resources is that it is also equitable. By definition, marginal cost pricing assures that each consumer bears full responsibility for the costs that her demand causes; insofar as this means that other consumers are not covering the costs to serve her, this is deemed to be fair (given a pre-existing distribution of wealth and income in the society).[82]

Firms will continue to produce so long as the income generated by the additional production covers the additional costs incurred by that production. Under perfect competition, where a firm has no power to set the market price — it is a price-taker — the market price describes the incremental revenue that the firm will receive for each additional unit of output. These two propositions, taken together, yield a second sharp insight of economic theory: namely, that a firm will continue to produce until its marginal cost equals price. By its own working, the market should produce goods at prices that reflect their opportunity costs, and efficiency is served.

2. Price in Competitive Markets

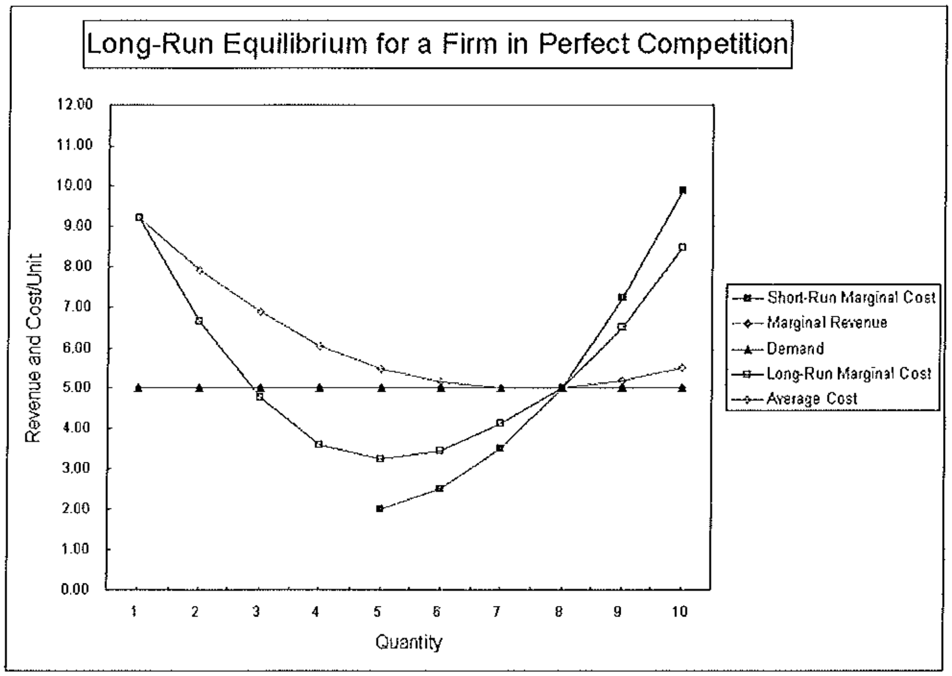
Firms act to maximize their own profit and consumers act to maximize their own welfare. In perfect competition, price is set by the market and, in equilibrium, it is that price at which producers are willing to supply a defined amount of a good, and only that amount, that will meet total demand for that good at that price. As price increases, producers are willing to supply more units of the good, but consumers are willing to purchase fewer units. Thus, there is only one price that satisfies the preferences of both suppliers and consumers simultaneously, and it is often referred to as the market *clearing* price (all goods produced at the price will be demanded).

Because no firm or consumer has market power (which is to say that the production or consumption decisions of any one firm or consumer will have no effect on overall supply or demand and, therefore, no effect on price), firms and consumers in competition are *price-takers*. Put another way, the relationship between price and demand that describes the behavior of consumers in the overall market for the good (namely that as demand increases, the price

82. Under at least one notion of fairness. This is not to say that there are not other, equally legitimate ideas of what is fair, and in designing rates it may be perfectly appropriate for regulators and policymakers to take such considerations into account (for example, the needs of low-income consumers). Here we are merely describing the economist's perspective. So long as price equals or exceeds marginal cost, there is no subsidization of one customer (or customer class) by another.

consumers are willing to pay decreases) does not describe the consumer behavior that any one firm confronts: specifically, the unwillingness of any consumer to pay higher than the market price for any of its output. (They would, of course, be perfectly happy to purchase all its output at less than the market price, but under such circumstances it would be less profitable than it would be while selling at the market price: it would be unable to meet the increased demand and simultaneously cover its costs.)

Because firms in competition cannot change the market price, they will instead optimize their factors of production (capital, labor, other inputs) in order to produce that quantity of goods and services which will, at the market price, maximize their profits (or, conversely, minimize their costs). As already noted, they will continue to produce goods until the cost of producing the next (or marginal) unit of output equals the additional (or marginal) revenue that they will receive for that unit, which of course is the market price. At that point they will stop producing, since to produce more will be to incur marginal costs that exceed marginal revenues, and total profits will necessarily fall. Figure 2 describes graphically a firm in a competitive market. Mathematically, marginal cost equals the difference between a firm's total costs when it supplies the incremental unit and its total costs when it does not.[83]



The interaction between supply and demand in an environment where the costs of production increase or remain flat as output increases (which describes the production functions of most industries) has the effect of creating economically efficient outcomes. The increasing-cost nature

83. It shows a market in equilibrium, which of course is a rare occurrence at best. To the extent that a market is not in equilibrium (in particular, that average total cost, ATC, does not equal price), a firm will earn revenues to cover some additional share of ATC when price exceeds SRMC. In the longer run, those oscillations around the LRMC and ATC curves provide sufficient opportunity for the efficient firm to cover its costs. Power exchanges operate quite explicitly in this fashion. See Chapter IV.

of the particular industry invites new producers to enter the market in the hope of producing at a lower cost, thus winning consumers and profits. However, the overall increase in supply caused by the new producers can only be sold (or *cleared*) at a lower market price, since consumers' willingness to purchase more of the good depends on the lowering of its price. This, as a consequence, improves overall societal welfare, since more consumers will then derive value from use of the good. In this way, competitive markets drive down the price of a good to the lowest possible point for a given level of demand.[84]

3. Price under Monopoly Conditions

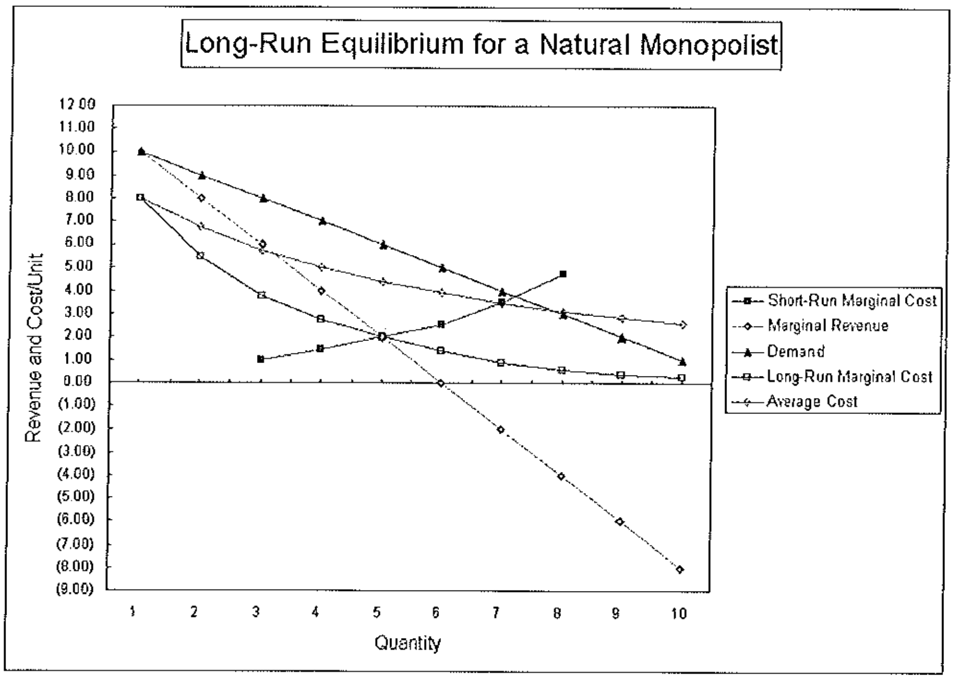
A monopolist, like a competitive firm, will maximize profits at that level of output where its marginal cost equals its marginal revenue. However, for the monopolist, marginal revenue per unit does not equal what would otherwise be the market price for the good. Because a monopolist supplies the entire market for a good, it is not a price-taker. It has the power to set price at that level which maximizes its profits, and is not constrained merely to optimizing its factors of production at a given price. A monopolist's profit-maximizing strategy is to restrict output and raise prices.

Its price-setting power is not absolute, however. The fundamental — inverse — relationship between price and demand still operates. The value that consumers see in a good is a function of its price, and this will determine how much of a good will be purchased at a particular price. Even if the good in question is essential, consumers may nevertheless be willing (or forced) to forego consumption if the price is too high. Ideally, a monopolist would like to charge each individual consumer the highest possible price that he or she is be willing to pay for the good (this is *price discrimination* in the economic, not legal, sense of the term). However, it is prevented from doing this by the threat of emerging secondary markets, wherein consumers would resell the good at prices higher than they themselves paid. This is arbitrage, and the independent attempts by many resellers to engage in it would quickly lower the market price to that originally charged by the monopolist. Thus, all consumers pay the same price for the good.

The effect of this market reality on monopolists is that, as output increases, price falls, but so too does marginal revenue. Consider, by way of example, the monopolist who can sell 100 units of its product at $2.00 per unit, 200 units at 1.50 per unit, and 300 units at 1.00 per unit. In the first

84. There are, of course, numerous assumptions (product homogeneity, firms are free to enter and exit the market, firms are price-takers, complete information is available to consumers, etc.) underpinning this theory. Given the limited focus of this review, they have gone unanalyzed here, but they do not merit unchallenged acceptance. The invalidity of any leads necessarily to a market failure (though the magnitude of that failure may not be particularly significant). One assumption that is rarely met — certainly not in the electric industry — is that *all* the costs of production and consumption are *fully* reflected in marginal cost and, therefore, price. In particular, many environmental costs are often not reflected in price. Consequently, the price signal is insufficient to properly inform consumers about the total costs of their consumption decisions, and of the true opportunity costs of alternative choices.

instance, the firm's total revenue is $200, and its marginal revenue is also $200. If it increases its output to 200 units, its total revenue becomes $300, but its marginal revenue falls to $100. If it again increases its output, this time to 300 units, its total revenue is $300, but its marginal revenue is zero. Unless its cost to make those additional 100 units is also zero (or less!), it is unlikely that the monopolist will produce them.



By itself, this exercise does not tell us what the profit-maximizing price and quantity of output are — before we can determine them, we need to know how the firm's costs change as output increases: we need to know its marginal cost and average costs curves — but it does reveal an important constraint that the price-setting firm faces. For the competitive firm, marginal revenue equals the market price, which does not change as the firm's output changes. But for the monopolist, marginal revenue is always less than price. Thus, because the monopolist (like a competitive firm) will continue to produce until marginal revenue equals marginal cost, the monopolist (unlike a competitive firm) will cease production at a point where price is significantly greater than marginal cost. This is hardly the most efficient level of output — output can be expanded until marginal cost equals price, and society will be better off. Again, whether the monopolist will still be profitable when price equals marginal cost (will it cover its total costs?) depends on the relationship of its average cost curve to its marginal cost curve at that point. But the essential point is that a monopolist's profit incentives do not cause it to act in a way that maximizes societal welfare. Monopoly power, then, is the power to set price above marginal cost (and, of course, above average cost).

4. Natural Monopoly

Monopolies can arise for any of a number of reasons, for example, through possession of legally granted patent or franchise rights or through control over some essential aspect of the production and marketing process. Some industries, however, are characterized by unusual features of production, such that their costs of production actually decrease as output increases. These are commonly described as increasing economies of scale or scope (*scale* in the case of a single product, *scope* when the cost reductions result from the production of two or more goods simultaneously). When this remains true for a broad range of output, it is generally more efficient (less costly) for one firm, rather than two or more, to supply the entire market. The circumstance in which a single firm can produce a desired level of output at a lower cost than any output combination of more than one firm is called "subadditivity of costs" and it leads to what we refer to as a *natural monopoly*.[85] See Figure 3.

Typically, it is an industry's technological characteristics that lead to natural monopoly, and we often see that a common feature of natural monopolies is a high ratio of fixed costs to total costs. Consequently, as output increases, average cost decreases. The technological elements of the electric industry that create natural monopoly conditions are, first and foremost, the transmission and distribution systems. They have high fixed costs and low operating costs. Transmission and distribution exhibit tremendous economies of scale. As for generation, it appears now that we have exhausted (or overcome) most economies of scale — cost no longer inexorably declines as the size of power stations increases. It is this fact that has, in large measure, precipitated various states' and countries' restructuring of their electric industries during the 1990s.

5. Short-Run v. Long-Run

In his seminal 1961 work, *Principles of Public Utility Rates*, James Bonbright warned readers that:

> "[M]arginal cost" is itself a highly ambiguous term, with the result that proposals to base minimum rates or rate relationships on marginal costs mean different things to different people. The most important ambiguity is that suggested by the distinction between "short-run" and "long-run" marginal costs. Indeed, this distinction is of critical import, for most of the really spectacular differences between incremental and average costs of public utility services are those which apply only when the former costs are taken to be of a short-run variety.[86]

At the heart of any analysis of marginal costs is the question of what is being measured. Broadly speaking, firms make use of two types of inputs to production, fixed and variable. Fixed inputs are those that are thought of as not changing as output changes: land, buildings, production facilities, some labor, general overhead, and the like. Variable inputs are, of course, just that — inputs that change as output changes: materials, fuel, labor, etc. And it goes without saying that associated with these inputs are costs. To effect relatively small changes in output, a firm can

85. Strictly speaking, economies of scale are not the essential, or even necessary, feature of natural monopolies. Baumol, William, Yale Journal on Regulation, Vol. 10: 63, 1993, at 67; Bonbright *et al.*, pp. 22-23.

86. Bonbright, p. 318.

alter the proportions of variable to fixed inputs, and with those changes come changes in the costs of production, both total and marginal. The same naturally can be said of large changes in output, except that in such cases the fixed inputs too can be varied — more land purchased, buildings constructed, facilities put in place — again the effects on costs can be calculated.

Economists differentiate between the circumstance in which some inputs (or factors of production) can be varied to meet a change in demand while others remain fixed and the circumstance in which all inputs can be varied as demand varies. The former is referred to as the *short run* and the latter as the *long run*. While we naturally think of the difference between the two as a matter of time, it is not, strictly speaking, correct to do so. But, as a practical matter, time does play a crucial role in the distinction. Marcel Boiteaux stated the issue succinctly more than fifty years ago:

> The theory of marginal cost pricing can be interpreted in many ways. Selling at marginal cost means fixing a price equivalent to the cost of producing one additional unit. This cost obviously differs according to whether it is planned to produce the extra unit once only or to raise by one unit the flow of goods which was turned out before. Production of one additional unit only once would not justify making any changes in plant; on the other hand, a definite increase in the production flow might be inseparable from the adaptation of existing machinery to the new level of production.[87]

The analysis that drives us to the conclusion that price should equal marginal cost also calls for the measurement of marginal cost as a function of very small increments of output. Typically under such conditions, only variable inputs to production can be altered to satisfy incremental demand, and therefore the efficiency imperative (namely, "that of securing the optimum utilization of whatever plant capacity exists at a particular time"[88]) would suggest that price should equal the short-run marginal cost of production (SRMC). In the electric industry, this means that a very small increase in demand is typically met by increasing the output of existing generation. The cost that the firm incurs to meet this demand (*i.e.*, the change in its total costs) consists of primarily additional fuel and, in certain cases, labor costs.[89]

If that increase in demand is not a one-time event, if it can be expected to persist, then the firm can take actions to more efficiently (*i.e.*, at lesser cost) meet that demand, which is to say it can adjust the proportions of all inputs — variable as well as fixed — to produce the desired output

87. Boiteaux, Marcel, "Peak-Load Pricing," 30 *Journal of Business of the University of Chicago*, 1960, p. 157 (translated by H.W. Izzard). This is an update of the author's earlier paper, "La Tarification des demandes en pointe: application de la théorie de la vente au coût marginal," 58 *Revue générale de l'électricité*, 1949.

88. Bonbright, p. 399.

89. This is perhaps an oversimplification. There are other costs that vary in the short run as well and should be recognized in the SRMC: for example, any operations and maintenance costs that can be attributed to the increased production and any additional wear-and-tear (life-shortening physical depreciation) on the fixed assets that is not already reflected in incremental O&M costs.

at a lower total cost. When it is thus free to re-optimize its entire production process, when all inputs are indeed variable, the cost of incremental production is referred to as the long-run marginal cost (LRMC).

Figure 2 describes the functioning of a competitive market when all factors of production in an economy have been optimized, which yields almost magically the confluence of SRMC, LRMC, average cost, and price. Of course, such an event is rare at best, since economies are dynamic things, but its absence does not mean that firms, if indeed pricing at SRMC, will fail to recover their total costs (including returns of and to capital). At certain times, there will be excess capacity in a market, and the SRMC of production — and price — will fall below LRMC and average cost (AC). At other times, there will be capacity shortages and SRMC will rise above both LRMC and AC, reflecting the high, perhaps very high, cost of meeting demand at those times of constraint. Over time, in a well-functioning market, these vicissitudes will operate to cover a firm's total costs by pushing price (that is, SRMC) to LRMC, while simultaneously giving appropriate incentives for firms to enter or leave the market.[90]

Like competitive firms, monopolistic ones will see a convergence of their short- and long-run marginal costs of production; however, this will not occur at a point on the demand curve — that is, at a price — that is equal to these marginal costs. Instead, the points at which the SRMC and LRMC cost curves equal demand (*i.e.*, intersect the demand curve), if they ever do at all, often differ substantially. See Figure 2.[91] It is this feature of monopolies, and natural ones in particular, that makes the question of the short-run or long-run significant for rate design.

Economists have differed on this point. Some argue that close adherence to the strictures of economic theory counsels for SRMC-based rates, but such a strategy would inevitably be attended by great variability in prices from period to period. It would involve tremendous administrative complexities and, for most customers, great confusion and, likely, resentment. But, more importantly, SRMC pricing will have tremendously adverse equity impacts on the various customer classes, and on customers within classes. Relying instead on long-run marginal costs, in an assumed state of equilibrium, as the basis for rates has better attractions: stability, a recognition of the long-term nature of both consumption and investment decisions in the electric sector, and greater administrative simplicity.[92] And this has been the trend in the industry for

90. Kahn, Vol. I, at 73-74, 85; Bonbright, pp. 321, 331-332, 374, fn. 12.

91. Figure 2 shows the SRMC curve rising above the LRMC curve and intersecting the demand curve at a significantly higher price than that which the intersection of the demand and LRMC curves, if shown on this graph, would describe. It could, of course, just as easily happen that the SRMC remains *below* the LRMC curve for a wide range of output, thus intersecting both the marginal revenue and demand curves at different points (if at all) than does the LRMC (again, if at all). The trajectories of those curves, especially in the electric industry, depend upon the particular relationships of capacity to demand at particular times of the day and year.

92. Bonbright, pp. 300-304, 332-334, 396. As previously noted, in equilibrium, short-run marginal cost equals long-run marginal costs. Put another way, when plant investment policy is optimized, SRMC will equal LRMC. Boiteaux, p. 165-167. Consequently, the question of which is preferred is moot. Reliance upon short-run marginal costs, which emphasize energy costs (except in times of capacity constraints) can lead to unwanted, and inequitable,

(continued...)

several decades at least, and it has been generally supported by the theoretical work of Bonbright, Kahn, and others.[93] As the National Economic Research Associates pointed out in 1977, when discussing the proper basis for designing rates:

> It is important to recognize that these marginal costs *do not* represent a form of prospective rate base or average incremental costs, but reflect time-differentiated marginal costs upon which consumption decisions should be based. These costs *do* represent the cost if reproducing the service provided at today's costs and under today's technologies, and are the costs that, in the long run (as defined by the economist) will be saved or incurred in the production and delivery of electric energy. Many people have incorrectly come to view marginal cost as the cost of growth. Economically speaking, this view is wrong because in the long run it will be necessary, even without peak-load growth, to replace old and unreliable facilities at current costs and with current technology. Thus, in the long run, a decision to consume less electricity (as opposed to a decision to continue to consume) will reduce the costs incurred in that replacement.[94]

6. Marginal v. Incremental

This discussion has so far adopted the vocabulary of microeconomics, and has therefore spoken in terms of *marginal* costs. But by its strict definition, marginal cost refers to the cost incurred to produce an infinitesimal increase in output (which is often generalized as a single unit of output). However, as a practical matter, firms can rarely isolate the costs incurred to produce that single additional unit of output (or the costs that are avoided by decreasing production by one unit). Consequently, utilities and regulators often calculate what is sometimes referred to as *incremental* cost, a cost per unit derived from the costs to increase (or decrease) output by some specified amount, significant enough to be measurable and generally corresponding to the way in which additional units of the good are produced. In this sense, the cost derived can be rightly regarded as an *average* incremental cost; but so long as the increment of output under consideration is only as large as it needs to be to yield meaningful results, the distortions that this

---

92. (...continued)
class allocations of costs. This is because embedded costs in excess of SRMC must still be collected, and economic theory offers no sure and fair method for doing so (refer to the discussion on Ramsey pricing in Chapter III). The history of rate design over the past three decades shows consumer advocates favoring SRMC pricing when it benefitted the consumers in question: driven in particular by the advantageous mixing of short- and long-run energy and capacity costs, but also in certain cases by improper application of methodologies. These experiences moved those jurisdictions that use marginal costs as a basis for setting rates to look to LRMC in equilibrium, thus assigning a greater share of costs (including capacity) according to the principle of causation, and covering more of the embedded costs in rates before applying some other principle of cost assignment. See also Bonbright, pp. 318*ff.*

93. Bonbright, pp. 317-336; Kahn, Vol. I, pp. 71-87; Bonbright *et al.*, pp. 467-468. These authors all note, however, that the preference for LRMC is not absolute, that there are times when pricing at SRMC will be both appropriate and possible.

94. NERA, pp. 17-18.

cost estimate introduce will be minimal or even non-existent (because the marginal cost curve over relatively small changes in output is often quite flat).

SELECTED BIBLIOGRAPHY

Awerbuch, Shimon, Leonard S. Hyman, and Andrew Vesey, *Unlocking the Benefits of Restructuring: A Blueprint for Transmission*, Public Utilities Reports Inc., Vienna, VA, Nov. 1999.

Baumol, William, *Yale Journal on Regulation*, Vol. 10: 63, 1993.

Boiteaux, Marcel, "Peak-Load Pricing," 30 *Journal of Business of the University of Chicago*, 1960 (translated by H.W. Izzard). This is an update of the author's earlier paper, "La Tarification des demandes en pointe: application de la théorie de la vente au coût marginal," 58 *Revue générale de l'électricité*, 1949.

Bonbright, James C., *Principles of Public Utility Rates*, Public Utilities Reports, Inc., Columbia University Press, New York, 1961.

Bonbright, James C., Albert L Danielsen, and David R. Kamerschen, *Principles of Public Utility Rates*, Public Utilities Reports, Inc., Arlington, VA, 1988.

Bouford, James D., "Practical Application of the Standardized Component Method for the Determination of Marginal Demand and Customer Costs at the Distribution Level", Central Maine Power Company, unpublished and undated.

Bouford, James D., "Standardized Component Method for the Determination of Marginal and Avoided Demand Cost at the Distribution Level", Central Maine Power Company, unpublished and undated.

Chernick, Paul, *From Here to Efficiency: Securing Demand-Management Resources*, Vol. 5., 1993.

Electric Power Research Institute, *Rate Design: Traditional and Innovative Approaches*, CU-6886, Research Project 2343-4, Palo Alto, CA, July 1990.

Fox-Penner, Peter, *Electric Utility Restructuring: A Guide to the Competitive Era*, Public Utilities Reports, Inc., Vienna, VA, 1997.

Gould, J.P., and C.E. Ferguson, *Microeconomic Theory*, 5th Ed., Richard D. Irwin, Inc., Homewood, Illinois, 1980.

Kahn, Alfred E., *The Economics of Regulation*, Vols. I and II, The MIT Press, Cambridge, MA, 1988.

Morey, Matthew J., "Distributed Generation: Is It the Wave of the Future," Edison Electric Institute, presentation at the mid-year meeting of the National Association of State Utility Consumer Advocates, June 5, 2000.

Moskovitz, David, *Profits and Progress Through Distributed Resources*, NARUC, 2000.

National Association of Regulatory Utility Commissioners, *Electric Utility Cost Allocation Manual*, Washington, D.C., 1992.

National Economic Research Associates, Inc., "How To Quantify Marginal Costs: Topic 4," Electric Utility Rate Design Study, March 10, 1977. (EURDS was commissioned by the National Association of Regulatory Utility Commissioners, and was the joint effort of the Electric Power Research Institute, the Edison Electric Institute, the American Public Power Association, and the National Rural Electric Cooperative Association.)

Neil, Charles E., "Entering the Seventh Decade of Electric Power," *Edison Electric Institute Bulletin*, September 1942.

Pierce, Richard J., Jr., and Ernest Gellhorn, *Regulated Industries*, 4th Ed., West Group, St. Paul, MN, 1999.

Sterzinger, George, "The Customer Charge and Problems of Double Allocation of Costs," *Public Utilities Fortnightly*, July 2, 1981.

Stoft, Steven, Carrie Webber, and Ryan Wiser, *Transmission Pricing and Renewables: Issues, Options, and Recommendations*, Ernest Orlando Lawrence Berkeley National Laboratory, LBNL-39845, May 1997.

Uhler, Robert G., "Electric-Utility Pricing Issues: Old and New," in Danielson, Albert L., and David Kamerschen, eds., *Current Issues in Public Utility Economics: Essays in Honor of James C. Bonbright*, Lexington Books, Lexington, MA, 1983.