Discussion of Electricity Price Reforms and Other Regulatory Options To Effect Efficient Consumption

Regulatory Assistance Project
August 2003

I. Introduction

As noted in the our policy paper, "China's Regulatory Challenges Relating to Power Shortages,"

> Power shortages present SERC with a special challenge and opportunity. SERC has a critical role in solving an electricity crisis and in preventing future problems. Power shortages frequently lead to poor decisions and lost opportunities. There are many examples of power shortages leading countries to commit to too much of the wrong kind of power supply, with undesirable economic and environmental consequences.
>
> The most effective, quickest, lowest cost, and cleanest options to address the shortage are energy efficiency, renewables, and innovative pricing reforms. SERC focus on these options is needed because these options are not well known and may not be attractive options to the utility companies.

The purpose of this paper is to discuss the third category of options – innovative pricing – and to identify promising reforms that will effect positive changes in consumer behavior ("demand responsiveness"), thus improving the reliability of, and reducing costs on, China's electric system. For China and SERC, our discussion and recommendations are shaped by three conditions:

1. SERC does not have pricing authority (though it may in the future). It does, however, have other authority to adopt market rules and other innovative programs that can achieve many of the benefits of improved pricing. SERC also has the authority to recommend prices to NDRC.

2. Some retail prices currently are close to market prices, but prices for many of the largest and most energy-intensive industries are well below cost. These large, energy-intensive industries are the consumers that are best able to increase their end-use efficiency in response to more economically accurate prices.

3. Innovative pricing practices can be combined with other policies aimed at to increasing China's end-use energy efficiency.

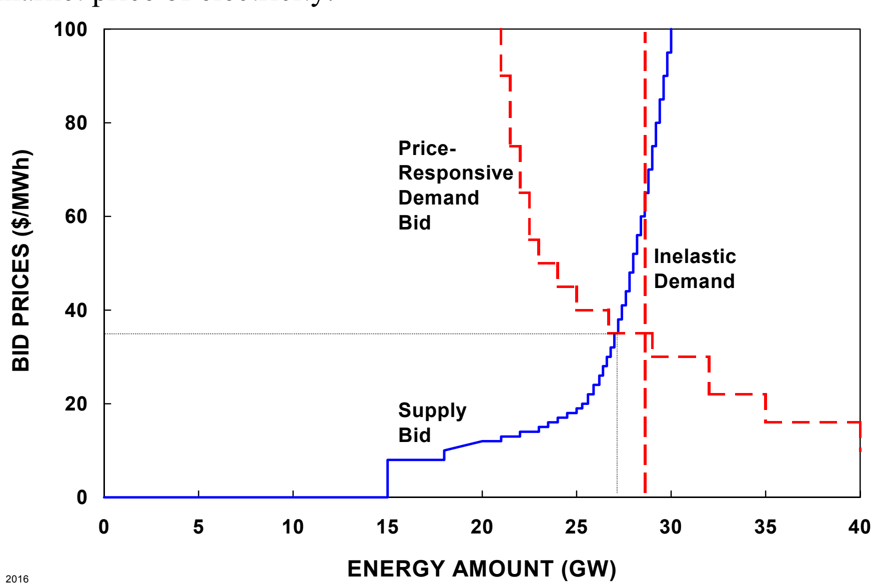
Our basic conclusion is there is a significant amount of efficiency that can be achieved through better economic signal given to consumers though better pricing and other innovative programs. We recommend that SERC evaluate and adopt pricing and related policies that will most cost-effectively lead to increased efficiency. These reforms should be implemented in ways that are consistent with other stated objectives, such as consumer protection, economic efficiency, equity, and environmental protection.

II. Statement of the Problem

With or without the threat of power shortages, China can achieve significant efficiency improvements by reforming electricity pricing and through other innovative programs that give consumers the incentive to invest in more efficient buildings, appliances, motors, and processes. The electric industry reforms now being implemented, and the creation of new power markets give China and SERC special opportunities.

Most commodity markets work through a steady feedback loop that relates price, demand, and supply. Prices are the language that the feedback loop depends on. In electricity, typically prices have only the barest connection to the cost of production. In electricity, retail prices are typically set to recover the average cost per kilowatt-hour. The consequence of this is inefficient consumption – too much consumption at times when the production cost exceeds the retail price and, in very limited circumstances, too little when price exceeds cost.[1] This can lead to degraded reliability, higher system costs, higher prices, and shortages.

Consumer demand responds to prices and other economic signals. There are many times when consumers – residential, commercial, and industrial – are willing to reduce their demand when prices are high or when they receive benefits of reducing or delaying consumption. This benefits consumers and the electric system in two ways. First, the consumer that reduced demand saves money. Second, all consumers benefit because the reduction in demand lowers the price of electricity experienced in wholesale generation markets. Figure 1 shows how, in one of the regional US wholesale markets, the willingness of customers to reduce their demand results in significant decreases in the market price of electricity.



[1] Moreover, to the extent that retail prices do not reflect the external environmental costs of electricity production, the commodity is undervalued and may be over-consumed.

Figure 1. Wholesale electric market supply and demand curves. Revealing customers' willingness to pay yields a small reduction in demand (from 28 to 27 GW), and a large reduction in the market clearing price (from $60/MWh to $38/MWh).[2]

Broadly speaking, there are three means of achieving demand reductions of these kinds.

1. Short-term customer load curtailments in response to direct payments from utilities or the system operator.

2. Short-term customer load curtailments in response to changes in retail prices.

3. Long-term improvements in end-use energy efficiency in response to retail prices or other incentive programs.[3]

III. Rate Designs to Improve the Efficiency of Electricity Consumption

1. There are two basic ways to give consumers better economic signals. The first is through changing the level and structure of retail prices. For example raising on-peak energy prices.

2. The second approach is through a wide range of market rules, incentive and penalty schemes, and other innovative programs. For example, a market-based payment to consumers who reduce demand during peak periods.

Both approaches can produce a more efficient result.

Changing the level and structure of retail prices may be the traditional regulatory approach, it may be more difficult in China because pricing authority is not a SERC function and because significant changes in electricity prices must be phased in gradually.

The second approach can be a more targeted ways of sending better economic signals to suppliers and customers, and are currently within SERC's authority. Moreover, consumers tend to be more accepting of these approaches than they are to significant changes in prices. We describe both approaches below with our recommendations.

---

[2] "The few examples that have been observed indicate that when supply is scarce relative to expected demand a reduction in demand of 2-5 percent could reduce prices by half or more." M. Rosenzweig, et. al., "Market Power and Demand Responsiveness: Letting Customers Protect Themselves," *Electricity Journal*, May 2003, p. 15.

[3] A harsher method for reducing customer demand, not discussed in this paper, is simply to turn off particular distribution and transmission circuits, i.e., rolling blackouts of the sort used in California during the crisis of 2001. This, of course, should be avoided.

Electric service is priced in a variety of ways. Pricing policy, whether set by firms or regulators, is influenced by a number of factors and objectives. Among these are economic efficiency, fairness, revenue stability, as well as certain practical considerations, such as simplicity, customer acceptance, continuity, and the availability and costs of metering and communications technologies to support those policies. When viewed in this light, pricing structures run along a continuum that marks the trade-offs between innovative and more complex pricing on the one hand and information needs and ease of administration on the other. The further one deviates from average prices, the more "dynamic" the rate structure becomes.[4] That continuum can be roughly divided into three broad segments:

1. *Energy-only pricing.* Rate designs that do not require special metering capability beyond that of the traditional revenue meter, which measures energy consumption only and is typically read once a month: flat, seasonal, block, etc.;

2. *Multi-part and time-of-use pricing.* Rate designs that depend upon more sophisticated metering – multi-part (energy and demand) and time of use (TOU) – but are still mostly read only monthly; and

3. *Real*-time *pricing.* Rate designs that send customers different prices on short notice for different hours of the day and for different days, to in some way reflect changing conditions in the short-term market – *e.g.*, real-time pricing (RTP) – and make use of sophisticated metering and communications systems that link them to their suppliers.

In determining whether a potential rate design is appropriate, the regulator must consider its potential effects. 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? Who benefits, who loses? What kinds of special metering and information management systems are required to support particular rate structures, and what are their costs? Regulators must apply their judgment when making these decisions. Seemingly small changes in a rate design can have very significant consequences for different customers.

---

[4] "Dynamic pricing" is a term used to describe any rate design that aims to give customers a truer signal of the economic costs of meeting their demand than simple average cost rates. Thus, a shift from average rates to time-of-use rates to demand and energy charges or to the various forms of real-time prices is considered a move toward more dynamic pricing. Others hold a more narrow definition: dynamic pricing "is any electricity tariff that recognizes the inherent uncertainty of supply prices." Stephen S. George and Ahmad Faruqui, Charles River Associates, *The Economic Value of Dynamic Pricing for Small Consumers*, presentation at the California Energy Commission Workshop on "Achieving Greater Demand Response in the California Electricity Market," March 15, 2002.

There are a variety of approaches to retail pricing that will evoke changes in customer behavior. Whether the changes can be relied upon for managing system loads in the short run depends on the degree to which the rates reflect the real-time variability of wholesale prices. While time-of-use and seasonally differentiated rates will have positive long-term impacts on system load factor, resource needs, and efficiency, they provide little incentive to adjust load in response to actual hourly or daily prices. The challenge facing policymakers is to develop rate structures that meet a variety of objectives – among them economic efficiency (in both the short and long runs), fairness, ease of administration, simplicity, and so on.[5]

The following is a menu of price reform options.

- *Time-of-use rates*. These daily energy or energy and demand rates are differentiated by peak and off-peak (and, possibly, shoulder) periods. One variation is the overlay of a real-time "critical" peak period, in the manner of the Gulf Power AEM program. Another is to identify critical days, rather than simply hours, during which consumption is priced to reflect the very high market costs.

- *Seasonally differentiated*. Those months during which consumption drives system peak see rates that reflect, in some measure, the costs of the capacity (generation, transmission, and distribution) needed to serve that peak. Seasonal differentiation can be applied not only to simple energy-only rates, but also to TOU and multi-part rates.

- *Multi-part rates*. These rates separate the charges customers pay for energy and capacity. Historically, demand charges were linked not to coincident system peak but simply to the customer's peak demand during the billing period. Multi-part rates are also often differentiated by season and daily time of use.

- *Block rates*. These are typically energy-only rate designs in which the unit price for incremental consumption changes as defined thresholds of usage within a period are passed. For example, the first 200 kilowatt-hours of usage might be priced at $0.10/kWh, the next 400 kWh at $0.08/kWh, and all succeeding usage at $0.065/kWh. This would be an example of declining block rates, but they could as easily be inclining. While these rates do give customers some idea of the cost of incremental production, it is approximate at best, since there is an imperfect relationship between the rate charged and the time of use (coincidence with system peak or other constraint).

- *Distribution-only service*. In restructured industries where commodity sales are separated from delivery service, the design of rates for distribution remains a regulatory responsibility. Treating distribution as if its costs do not vary with the time or amount of usage (which, in the long run, they do) can lead to the adoption

---

[5] Bonbright, James C., *Principles of Public Utility Rates*, Public Utilities Reports, Inc., Columbia University Press, New York, 1961, at 291.

of large fixed, recurring rates that are unavoidable, regardless of changes in demand. This can inhibit customer willingness to take otherwise cost-effective demand reduction actions.

- *Real-time pricing*. RTP links hourly prices to hourly changes in the day-of (real-time) or day-ahead cost of power. One option is "one-part" pricing, in which all usage is priced at the hourly, or spot, price, adjusted as appropriate for delivery, congestion, line losses, and other relevant costs. Unlimited in this fashion, they place all price risk on the customer and, consequently, few customers have taken service under them. Providers have developed risk mitigation (risk sharing) products to address this concern: for example, price caps and floors, options for locked prices for limited periods, and triggers (where the spot price is paid only when it exceeds a specified minimum for a specified period). A second approach is "two-part" pricing. There is an "access" charge for using a pre-determined baseline quantity (*e.g.*, baseline kWhs * embedded rate/kWh), and spot prices (or credits) for variations from the baseline. The baseline is often set on a customer-specific basis. The two-part RTP rate is a more common form of price-risk sharing, and it provides a certain measure of revenue certainty for both the provider and the customer.

IV. Recommended Reforms

Based on conditions in China, we suggest focusing on the following reforms. The first group of reforms is aimed at changing retail electricity prices. These reforms would require approval by NDRC. The second group consists of non-price reforms that can produce efficiency gains similar to the price reforms. The second group of options are within SERC's authority.

A. Price Reforms

Inclining block rates for residential and small commercial customers.

The initial block should be set so as to cover the average minimum consumption of a consumer in each rate class. The second block (often called the "tailblock") would be priced at a higher rate to discourage inefficient demand. The intent is not to penalize customers for demand that they cannot avoid (for example, for basic services such as refrigeration and lighting), but to encourage them to turn off end-uses that they do not value highly. The inclining block rate design could be limited to only peak season (i.e., summer) months.

Real-time prices for large industrial customers

Subjecting some part of large commercial and industrial users' consumption to hourly market prices could have substantial impacts on their demand. If they are put at risk for high peak prices, they will have a strong incentive to manage their loads more efficiently.

Typically, the customers who are willing to take service under such rates are those who can shift significant parts of their load to low-cost, off peak times of the day and night.[6]

Interruptible programs for large commercial and industrial customers

A two-part credit may be the most effective form of payment for these programs. The first part would be a periodic (monthly) bill credit that the customer receives in return for enrolling in the program – that is, in return for being *available* to be interrupted. This credit reflects the value of the capacity that the customer will provide if called upon to interrupt all or part of its load. The second part would be payments for actual interruptions, based on the value of the energy saved. This two-part payment structure encourages customer investment in on-site generation. Moreover, the program can be designed so as to require participants to improve the efficiency of their end-uses.

Lower rates for efficient end-uses

Customers that achieve specified levels of end-use efficiency would be eligible for tariffed rates that are lower than those paid by less efficient customers. The economic justification for this is that efficient customers impose fewer costs on the electric system than do inefficient customers. Critical to this kind of rate design is that the efficiency standards are clear and not easily "gamed."

B. Non-Price Reforms

The following options give SERC the ability to improve the economic incentives seen by consumers without directly changing retail electricity prices. These options provide targeted approaches to encouraging more efficient use of electricity and greater investment in energy efficiency and load management. We recommend that SERC test all of these options.

Demand buy-back

A demand buy-back programs gives (in the form of wholesale market rules that give customers the right to sell demand reductions back to the utilities or the operator of a wholesale market). The price paid for the demand reduction varies depending on the system conditions at the time of the reduction. These programs are an overlay on any of the other rate designs. The payments to the customer can be structured in ways that encourage the customer to install distributed generation or invest in improved end-use efficiency.

Credit or rebate programs

These include programs that encourage greater investment in end-use efficiency and distributed generation. They can be targeted to meet particular needs, for example, to acquire demand reductions at peak times or in constrained areas of the transmission and

[6] Beginning on August 1, 2003, large commercial and industrial customers in New Jersey that do not take service from competitive suppliers will be required to pay hourly prices for generation. It is expected that these customers will either respond favorably to the new rate design or find competitive suppliers to provide hedges against the price risk.

distribution network. For example, customers can be rewarded for achieving specified reductions in their electricity consumption, when compared to their usage in the same period (month) of the prior year. In California during the crisis of 2001, customers who reduced their summer (peak) usage by 20% automatically received a 20% discount off their electric bill (the discount was in addition to the bill savings they received by simply reducing consumption). This was called the "20/20" program. Also during that time, a similar program was adopted in the state of Washington. Both programs were extremely effective in getting customers to reduce load, mostly by investing in end-use efficiency measures.

A variation on this would be a program in which customers who improve the efficiency of their end-uses by a specified and demonstrable amount would be eligible for lower rates (this program could remain in effect after any potential power shortages are averted). In addition, improved building and appliance standards could be set, and a customer's failure to meet those standards would result in higher rates or bill surcharges.

Rate credits for efficiency and distributed generation in high-cost areas of the transmission and distribution network

This program would call for the utility to offer financial credits for efficiency or distributed generation installed in specified, high-cost areas of the network. The aim of the program is to encourage the installation of customer resources that will avoid the need for more costly investment in wires. The credit amount would be based on the cost savings that the efficiency or distributed generation could provide.

On-peak public benefit charges for large energy intensive industries classes

Typically TOU rates are designed to be "revenue neutral" in comparison to the pre-existing flat rates. In other words, the on-peak rates will be greater than the average rate (reflecting the higher costs of production in those hours), and the off-peak rates will be lower than the average (reflecting the lower costs of production in those hours); thus, both rate designs will produce the same amount of revenue for a given level of demand.[7] In China, prices for these customers are already well below cost. Adopting typical TOU rates may produce some system benefits but average prices for these customers may fall further below cost. A better, more targeted approach may be to impose an energy efficiency charge during on-peak hours. The additional revenue that is collected during off-peak periods (that is, the difference between the average rate and the off-peak rate) should be used to fund end-use energy efficiency programs. The effect of such a program will be to make these industries more competitive and reduce the subsidy they currently receive

[7] In reality, however, the TOU rates change customer behavior, and utility revenues may differ from expectations. Later adjustments in the TOU rate design may be necessary to correct such problems.

Stable, longer term "green" pricing

Customers should be permitted to purchase stably priced "green" electricity for extended periods of time (e.g., several years). In return for the non-dynamic price, the customer agrees to commit to the green (renewable) product for a specified duration. This kind of pricing will allow customers to hedge against price risk, and will encourage the development of additional, non-polluting resources to help meet increasing demand for power.

Changes in rate design may require changes in the metering of customers' usage. The appendix that follows discusses metering issues in general and also gives a brief description of traditional regulatory approaches to utility revenue setting and retail rate design.

Appendix:Metering and Methods for Setting Rates

I. Metering

Changes in rate design will, in many cases, require changes in metering technology. Currently, metering in China consists primarily of meters that measure only the number of kilowatt-hours that are consumed in a period, without regard to the actual hours in which the consumption takes place. This kind of metering is limited in its capability to support more dynamic rate designs. It reveals nothing about the customer's usage patterns and it is available for review only after manual meter reading, typically long after the fact of consumption. These shortcomings constrain providers to energy-based rate structures that are not time-differentiated within billing periods, but they do allow for certain consumption-based structures (*e.g.*, inclining and declining blocks). In addition, seasonal differentiation is possible, so long as the rate changes correspond to the beginnings and ends of billing periods.

Rate structures that vary by periods that are shorter than the billing period require a meter that can differentiate between consumption in the several (typically) daily periods. Given the higher costs of metering and administration for TOU rate structures, these meters have been limited primarily to the higher usage consumers.

Multi-part rates require metering that can record both energy usage and peak customer demand during the billing period and, where applicable, time-of-use differentiations.

Interval meters, as their name suggests, record and store usage data for each interval, generally an hour, though often shorter periods are possible (even down to one minute). Most utilities in the United States collect hourly usage data from their larger commercial and industrial customers, although the data are typically retrieved only once a month. The infrequent collection inhibits the utility's ability to offer more dynamic pricing options.[8] Interval metering can support any of the rate designs discussed in this paper; however, in the case of real-time pricing, additional equipment that communicates the hourly price of electricity to the consumer must also be installed.

---

[8] But it should be noted that dynamic pricing is not the sole or even primary justification for interval meters. The data provided by interval metering improve billing accuracy, support the more accurate assignment of costs to customers, give utilities better tools with which to manage their customers' loads, and support rate design generally, all of which provide significant value to companies and customers.

The following table relates rate designs to the metering technologies necessary to support them.

Rate Designs	Type of Meter System	System Features	Capabilities for Rate Design
• Energy-only • TOU • Demand and energy • Seasonally differentiated	Conventional Manual / Electronic Keypad	• Requires meter reader to cover a fixed route • Meter values key-entered or electronically downloaded via port to hand-held recorder	• Typically limited to a single kWh usage value each billing cycle • TOU meter for TOU rates • Demand meter required for multi-part rates • Cannot economically or logistically support the collection of time varying kW interval data • Data only available once each billing cycle or with special read
• Energy-only • TOU • Demand • Seasonally differentiated	Remote Meter Reading	• Requires meter reader to cover a fixed route • Van-based drive by or hand-held systems that use low power radio to transmit meter reading over short distances	• Can support the collection of multiple kWh register values used in standard TOU rates • Demand meter required for multi-part rates • Communication methods cannot economically or logistically support the collection of time varying kW interval data • Data only available once each billing cycle or with special read
• All of the above • Real-time pricing	Automated Meter Reading	• Meters connected to a data repository by telephone, PCS, paging, satellite, fiber, or other communication technology • Stored meter reading can be collected on a fixed schedule or on demand	• Preferred methodology for collecting interval data • Full complement of interval and other meter data generally available on demand • Accessibility varies by technology

II. Cost-Based Rate Design

Rates should be set so as to enable a utility a reasonable opportunity to recover prudently incurred expenses (including investment) and a fair return on the remaining cost (the un-depreciated portion) of investment.

A. The Mechanics of Traditional Rate Setting

The general mathematical formula for determining rate levels for monopoly services begins with a computation of total revenues (revenue requirement) necessary to meet demand for service, as follows:

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

where:

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

The period of time under examination is called the "test year." In many places, rates are set using a *historic test year*, adjusted for ***known and measurable*** changes. The exercise yields an *adjusted test year* cost of service that is meant to be a predictor of a company's revenue needs during the period rates will be in effect.

The simplest way to set rates would be to divide the revenue requirement by sales volume (kWh), as follows:

$$\text{Rates} = \text{RR/Volume of sales}$$

Although actual rate-setting is somewhat more complicated than this (for example, customers are grouped according to their usage patterns, and the revenue requirement is allocated among those classes, according to principles of cost causation), but the essential mathematical relationship holds: the product of rates and sales is the revenue requirement.

A very important point must be kept in mind: this exercise assumes that there is a direct relationship between a utility's revenue requirement and the rates it should be allowed to charge. This is, of course, true, but bear in mind that regulators have traditionally set *rates*, not revenues. The revenue calculation is merely a tool for doing so. But, because rates are set to cover costs, regulators devote a good deal of attention to the constituent elements of a company's cost of service.

B. Rate Design: Pricing for Regulated Services

To a regulator, rate design is 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 charges.

The general objectives of economic regulation inform the rate design process. More specifically, we want to set economically-efficient prices (*i.e.*, prices which reflect, to the greatest extent possible, the long-run marginal costs of service), while simultaneously enabling the regulated firm a reasonable opportunity to recover its legitimate costs of providing service (including return on investment).

The particular problem faced by regulators in this exercise is that the legitimate historic (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 forward-looking long-run marginal costs that form the basis of economically efficient prices. The reconciliation of the need to cover historic costs with the desire to set economically efficient prices, and then to meet other objectives of regulation (such as fairness and low-income protection), requires much judgment. The several and sometimes competing rate design goals can be categorized as follows:

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;
- 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:
- A rate design should be, to the extent possible, simple, understandable, acceptable to the public, and easily administered.

a. Embedded Costs

As previously, rates are intended to recover the prudently incurred, embedded costs of service the costs that the utility actually pays. These costs are allocated among customer classes, consumer groupings typically formed according to their patterns of usage. Similar usage causes similar costs, thus enabling class-specific assignment of those costs. Among the costs to be identified and functionalized are energy and capacity, transmission, distribution, customer service, and others. The methods for cost assignment can be complex, but in the end the objective is to have those customers who cause the costs pay the costs.

Of course, not all costs can be easily categorized (for example, the joint and common costs that are necessary to the overall operations of the firm but are not directly necessary to the provision of any particular service), and so apportioning them among customer classes becomes an exercise in judgment. Regulators may decide in certain instances to allocate a cost according to a class's share of total energy usage, and in others according to class coincident demand for capacity. Notions of reasonableness and fairness when making these decisions guide regulators.

Once the cost of service is allocated among customer classes, rates can be set according to the mathematics already described. Each customer class has its own revenue requirement and expected volume of sales. Typically, however, not all of the costs of service are collected in energy charges, some (usually small) portion of them may be recovered through fixed, recurring fees called customer charges. These are billed whether the customer uses any electricity or not; the charges are intended to cover the costs of utility activities that are unrelated to usage, for example, metering, billing, and collection. In the main, however, the majority of costs are recovered through charges that vary with a customer's usage. There are two categories of these: energy and demand.

Energy charges collect revenues on a per-kWh basis. Demand charges collect revenues on a per kW basis. It is common for low-usage customer classes to pay energy-only charges, and included in those fees are the costs of capacity needed to serve that customer group. High-usage customers often are billed on both an energy and demand basis; their capacity costs are separated from their energy costs. While the costs of metering for this kind of service are higher than energy-only metering, the savings (for both the customer and utility) that flow from the customer's ability to respond to the clearer price signals invariably exceeds those costs.

b. Marginal Cost Pricing

The marginal cost of service is the cost incurred to serve an additional unit of consumption at a particular time, and it represents the cost to society to satisfy that incremental demand. By the very nature of monopoly, however, it is unlikely that at any particular time marginal cost will equal embedded cost (which is, in large measure, an average historic cost), and thus setting prices strictly equal to marginal costs will fail to generate the appropriate level of revenues for the company. Whether they are too high or too low will depend on the relationship of the utility's historic costs to the current costs of fuel and new technology.

The task of identifying and functionalizing the utility's costs for the purpose of determining its marginal cost of production at specified times is, in many ways, quite similar to the work done for embedded costs. Unlike an embedded cost study, which in effect calculates the average cost per unit of demand for each class and period under examination, a marginal cost study measures the cost of producing a defined increment of demand for each class and period specified. Total cost is only relevant insofar as marginal cost is a measure of the change in total cost as demand changes. In certain cases, particularly at times of peak demand when additional capacity may be called for, marginal cost will often exceed average cost; at other times, marginal cost may be significantly less than average cost, since typically the only costs incurred to serve incremental demand off peak are variable fuel and maintenance costs.

Once calculated, marginal costs are then treated as prices and are multiplied by expected units of demand in the various periods under study. This yields the expected total revenue that the company would collect under a marginal-cost pricing regime, which can then be compared to the embedded cost revenue requirement. How prices should then be adjusted depends on whether the marginal cost revenues are greater or less than the embedded.

There are a variety of ways to reconcile marginal cost prices with an embedded revenue requirement. Rates differentiated on the basis of time of day, week, or year of use are quite common, and often are designed to reflect marginal costs at times of peak demand (when costs are high) and average costs at other times. In this way, the utility's risk of revenue shortfall is lessened, and consumers see the important cost signals at times of capacity constraints. Inclining or declining tail-block rate structures are another option. With these, price changes (inclines or declines) as volume demanded during a time period (say, a month) increases. These may not send as accurate a price signal as will time-of-use rates, but they are generally seen as an improvement over flat, average rates.

In the end, regulators must apply their expertise and judgment when designing rates. Considerations that can inform their discretion include fairness, differences in demand elasticities (willingness to pay), and other public policies (such as low-income support and the pricing of environmental externalities). Distortions that hinder economically efficient outcomes will inevitably creep into prices; this disjunction between marginal and average costs is an unavoidable aspect of natural monopoly. What distortions, and in what magnitudes, then are acceptable? This is one of the central dilemmas of regulation, and there are no easy answers.