Priority Pricing of Interruptible Electric Service with an Early Notification Option*

*Todd Strauss*** and Shmuel Oren****

Priority pricing of interruptible electric service induces each customer to self-select a rationing priority that matches the rank order of its interruption loss. This paper extends the theory by considering the possibility of early notification, an option offered by many electric utilities. The proposed tariff structure allows a customer to choose either early notification and pay a fixed fee, or select no advance notification along with a level of compensation when interrupted. The chosen compensation determines customer service priority and corresponding price. Service priority is interpreted as an externality component of the marginal cost of system shortfall.

INTRODUCTION

More than 70% of investor-owned electric utilities in the United States offer some form of voluntary interruptible or curtailable electric service (Ebasco, 1987). Interruptible electric service refers to any customer load that is subject to partial or complete elimination for a period of time upon adequate notice from the electric utility. Typically, "adequate notice" ranges from 10 minutes to one full day. This is the *notification time* a customer receives prior to actual interruption. The conditions of notification time are included in

The Energy Journal, Vol. 14, No. 2.

*This research has been partially funded by National Science Foundation grant IRI-8902813, the University of California Universitywide Energy Research Group, the California Public Utilities Commission, and the Electric Power Research Institute.

**Yale University, School of Organization and Management, Box 1A, Yale Station, New Haven, CT 06520.

***University of California at Berkeley, Berkeley, CA 94720.

interruptible service tariffs, which may also specify the maximum number of interruptions allowed per day, per month, or per year; the maximum duration of any particular interruption; and the maximum number of interrupted hours per year. Christensen Associates (1988) describes interruptible service in more detail.

Analysis of interruptible service tariffs has focused on varying the demand charge for customers with different interruption losses, or *outage costs* in the terminology of the industry. Marchand (1974), Tschirhart and Jen (1979), Woo and Toyama (1986), Chao and Wilson (1987), and Viswanathan and Tse (1989) consider one-dimensional models that differentiate on reliability, i.e., probability or frequency of interruption. The models of Panzar and Sibley (1978), Hamlen and Jen (1983), and Woo (1990) include the amount of interruption as well as the frequency. Oren and Doucet (1990) model customers differing with both interruption cost and location on the distribution network. Using a load duration curve model, Chao, Oren, Smith, and Wilson (1986) consider both frequency and duration, but not amount. Smith (1989) and Oren (1990) consider two-dimensional models that incorporate both frequency and duration of individual interruptions.

None of these models includes notification time, an important element of actual interruptible service programs and tariffs such as Niagara Mohawk's Voluntary Interruptible Pricing Program, New England Electric Service's Cooperative Interruptible Service Program, and Southern California Edison's I-3 tariff schedule. In the analysis presented here, notification time is included through a simplified two-period model. Customers may be notified in the first period of interruption in the second period, or interrupted in the second period without advance notification. We first analyze the socially optimal notification and rationing plan under perfect information about customer interruption losses. Then we consider the implementation of that allocation through customer self-selection, and derive a simple tariff schedule that will achieve this goal. The tariff is interpreted as marginal cost pricing in the presence of an externality.

Our model is a case of priority pricing for two attributes—interruption cost, and benefit from early notification. By including the dependency of outage costs on early notification and customer heterogeneity with respect to the effects of such notification, we make some novel additions to the priority pricing literature.

MODEL DESCRIPTION AND NOTATION

Different customers and end uses incur different losses when an interruption of electric power occurs. However, no customer prefers a longer interruption to a shorter one, and no customer prefers a sudden, unexpected interruption to an interruption with some advance notification. Customers ordinarily assume that they will receive electric power, so a customer's interruption loss or *outage cost* is an "additive adjustment to the surplus...derived from its normal electric power consumption" (Smith, 1989). Similarly, the prices discussed in this paper are additive adjustments to customer bills "for avoided [or contracted] interruptions, as opposed to consumption."

With a control and metering technology that is able to separate end uses, each kilowatt (kW) of demand may be addressed separately. Consequently, each customer is considered to have one kW of demand; alternatively, each kW of demand is regarded as an independent decision-making unit. Furthermore, demand is non-stochastic.[1] Supply is stochastic, so the system *shortfall* is the difference between N, the number of customers, and the realized value of supply.

Each customer may then be characterized by (i) its loss if interrupted suddenly and unexpectedly, and (ii) its benefit from advance notification of an impending interruption. As modeled here, any and all notifications to customers may be issued only once, at a fixed length of time before the impending shortfall, say one period.

A customer's outage cost depends upon the time at which the customer makes an irrevocable decision not to use electric power at time T. Customers utilize early notification of impending interruptions to take advance actions in order to mitigate the effect and cost of an interruption. For example, customers may cancel shifts, reschedule production processes, or fire up backup generators. These irrevocable actions result in customer costs that are effectively sunk and irremediable, akin to unit commitment costs in traditional supply-side electric operations. Once a customer takes such an irrevocable action and incurs the commitment cost, the customer's marginal value of receiving power (at the posted energy cost) is assumed to be zero.

It is assumed that customers make immediate use of early notification. Once notified at time S of an impending interruption at time T, a customer takes some irrevocable action at time S to mitigate the interruption loss. Thus, a customer receiving early notification incurs its interruption loss minus its early notification benefit. This net loss is herein referred to as a customer's *early cost* (c_S). Because such a customer has no marginal benefit of receiving power, it is

1. Both unit demand and its non-stochastic nature are common assumptions of priority pricing research. For example, see Wilson (1989). Non-stochastic demand may be a reasonable assumption for large industrial and commercial customers with high load factors (average load divided by peak load). Residential customers tend to have low load factors and residential loads tend to be much more sensitive to ambient weather conditions. If only industrial and commercial customers are considered interruptible, the stochastic demands of residential customers may be lumped with stochastic supply into the random variable that is shortfall magnitude.

socially efficient to interrupt this customer. In other words, notifications lead irreversibly to interruptions.

At the time notifications are issued (S), the magnitude of an impending shortfall is uncertain, but it is fully revealed when the actual interruptions commence (T). If the magnitude of the actual shortfall exceeds the number of customers that were notified, this difference between supply and demand is satisfied by interrupting additional customers, without notification (or on very short notice), at time T. The interruption loss suffered by a customer interrupted without notification is referred to as its *late cost* (c_T). Any customer not notified at time S is on *standby*, and may be interrupted without notification at time T.

The customer population is heterogeneous with respect to c_S and c_T and characterized in terms of a distribution over c_S and c_T in the domain $0 \leq c_S \leq c_T$. Late costs are bounded above by C. For notational convenience, customer costs are scaled such that C equals one.

The electric utility must decide which customers should be notified and which should be on standby, and which standby customers should be interrupted, if necessary. This is the allocation problem faced by the utility. However, interruption losses are private information. Each customer knows its own particular outage costs, but the utility does not know the outage costs of any particular customer, only the aggregate distribution of outage costs in the customer population. Thus, a socially efficient tariff structure must induce customers to reveal their true outage costs through their selections from a menu of service options. The electric utility then uses the revealed preferences to allocate advance notifications and interruptions. Because this scheme calls for customer choice from a menu of service options, the tariff designer must recognize that each customer chooses from such a menu so as to maximize its own consumer surplus. This is the self-selection criterion (Mussa and Rosen, 1978).

The problem at hand is to design an interruptible service tariff. There are two aspects to this problem: allocating notifications and interruptions among customers (the allocation problem), and implementing this allocation through pricing, to account for customer self-selection. The goal is to minimize total expected customer interruption cost.

Shortfall duration is suppressed; this may be interpreted either as assuming that all shortfalls have a fixed duration or as taking expectations over duration. For notational convenience, both the distribution of interruption costs in the customer population and the distribution of shortfall magnitude are assumed to be continuous. If these distributions were discrete or mixed, clumping of customers might occur, but the sense of the analysis would remain unchanged.

We use the following notation:

N	=	Number of customers.
$D(x,y)$	=	Population distribution. Percentage of customers with *late cost*—cost of being interrupted without notification—less than or equal to x, and with *early cost*—cost of being interrupted with one period advance notification—less than or equal to y.
$d(x,y)$	=	Population density.
S	=	Time when advance notifications are issued.
T	=	Time when interruptions commence.
$G(q)$	=	Distribution of short-term uncertainty in shortfall magnitude at time S. Probability that the magnitude of the shortfall will be no greater than q.
$g(q)$	=	Density of short-term uncertainty in shortfall magnitude.

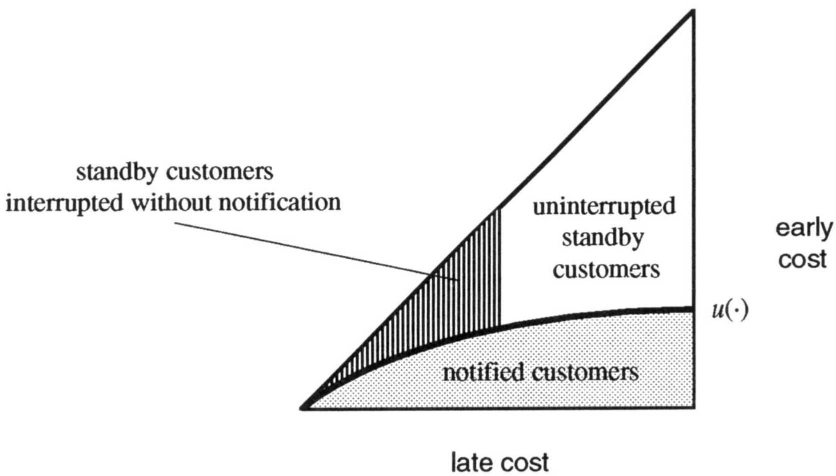
ALLOCATION PROBLEM

The utility's objective is to minimize the total expected social cost of an impending shortfall. We first consider the case where the utility has complete information about each customer's costs. The utility must decide which customers are notified and which are standby, and among standby customers, which to interrupt, if any, in order to balance supply and demand at time T.

The optimal notification policy is derived in the Appendix. Figure 1 qualitatively illustrates the optimal policy. The customer population is partitioned by a decision curve u. All customers under the decision curve u are notified at time S, while the balance of the shortfall is met at time T by interrupting standby customers in increasing order of late costs.

The standard model of priority rationing without considering early notification corresponds to the decision curve $u(x) = 0$, that is, no customers receive early notification. The policy of notifying all customers corresponds to the decision curve $u(x) = x$.

Figure 1. Optimal Notification and Interruption Policy



Now we compute the total expected interruption cost, at the decision time S, in terms of the decision curve to be optimized, u. The total interruption cost of all customers that are notified is (1).

$$N\int_0^1\int_0^{u(x)} yd(x,y)dydx \tag{1}$$

The number of standby customers interrupted is a random variable that depends on the short-term uncertainty in the magnitude of the shortfall. The probability that a standby customer with late cost v will not be interrupted is the probability that the shortfall magnitude will be less than $h(v;u)$, as defined by (2). The first term on the right side of (2) is the number of customers receiving early notification; the second term is the number of standby customers with late cost less than or equal to v.

$$h(v;u) = N\int_0^1\int_0^{u(x)} d(x,y)dydx + N\int_0^v\int_{u(x)}^x d(x,y)dydx \tag{2}$$

Given that v is the largest late cost among standby customers actually interrupted, the total interruption cost of standby customers is (3).

$$N\int_0^v x\int_{u(x)}^x d(x,y)dydx \tag{3}$$

Taking expectations over v yields (4), the total expected cost of interrupting standby customers.

$$\int_0^1\left(N\int_0^v x\int_{u(x)}^x d(x,y)dydx\right)dG[h(v;u)] \tag{4}$$

The sum of (1) and (4) is the total expected interruption cost due to a shortfall, when the notification policy described above is employed. Finding the optimal curve u is a problem in the calculus of variations. The derivation of the optimal curve appears in the Appendix. The formula for the optimal curve may be described by (5).

$$u(v) = v\bar{G}[h(v;u)] + \int_0^v z\ dG[h(z;u)] \tag{5}$$

Equation (5) has an intuitive interpretation. The first term on the right side of (5) is the expected interruption cost of a standby customer with late cost v. The second term is the expected marginal cost to the system when some standby customers are interrupted, but not standby customers with late cost v. Hence, $u(v)$ is the total expected marginal cost to the system of a standby customer with late cost v. The second term on the right side of (5) is an externality cost imputed on standby customers with late costs less than v by standby customers with late cost v. The externality is *reliability*: customers with higher priority impose interruptions on customers with lower priority.

In terms of social cost, the electric utility is indifferent between notifying and not notifying customers on the boundary, $[v,u(v)]$, $0 \leq v \leq 1$. However, a customer on the boundary prefers standby status to being notified, since its interruption loss is only the first term on the right side of (5). These results are discussed further in the section on pricing.

An Illustrative Example: Uniform Distribution of Costs and Shortfall Magnitude

To illustrate the above result, we consider the case where customer outage costs are uniformly distributed. This is represented by $d(x,y) = 2$, $0 \leq y \leq x \leq 1$. We also assume the shortfall magnitude to be uniformly distributed, which may be interpreted as a non-informative forecast. This is represented by $g(z) = 1/N$ for $0 \leq z \leq N$.

With these uniform distributions in effect, the total expected interruption cost may be represented as (6).

$$N\int_0^1\left\{[u(v)]^2+4\left(\int_0^v x[x-u(x)]dx\right)[v-u(v)]\right\}dv \tag{6}$$

From the derivation in the Appendix (A.6-8), the appropriate optimality conditions for the curve u are (7-9).

$$u'' - 2u = -2v \tag{7}$$

$$u(0) = 0 \tag{8}$$

$$u'(1) = 0 \tag{9}$$

The solution to (7-9) is (10). Figure 2 displays the solution graphically.

$$u(v) = v - \frac{\sinh\sqrt{2}v}{\sqrt{2}\cosh\sqrt{2}} \tag{10}$$

In this example, 46% of all customers receive early notification.[2] Interrupted standby customers are an additional expected 15% of all customers. The total expected interruption cost is $0.15N$, and 45% of this cost is borne by notified customers. Table 1 compares the results for this scenario under three models: priority rationing with early notification, as presented here; priority rationing without early notification, the standard priority pricing model; and random rationing. Priority rationing with early notification results in the smallest social cost, although more customers are interrupted. As indicated in the column labelled *relative cost*, priority rationing without early notification has total expected interruption cost 81% greater than the model presented here, while random rationing costs more than twice as much.

2. Since each customer has unit demand, 46% of all units of demand receive early notification.

Figure 2. Optimal Decision Curve for Uniform Assumptions

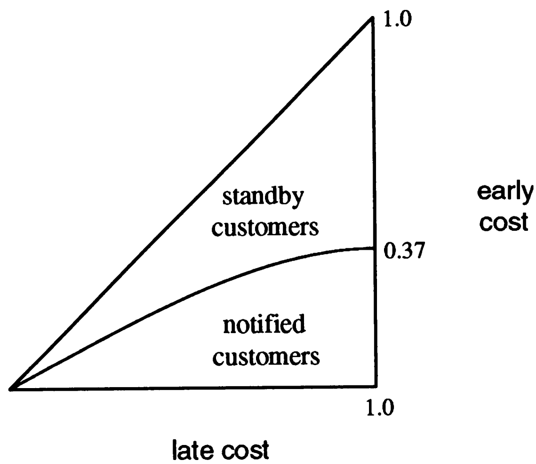


Table 1. Interruption Amounts and Costs for Uniform Scenario

Model	Amount Interrupted	Total Expected Interruption Cost	Relative Cost
Priority Rationing with Early Notification	$0.61N$	$0.15N$	100%
Priority Rationing without Early Notification	$0.50N$	$0.27N$	181
Random Rationing	$0.50N$	$0.33N$	226

OPTIMAL PRICING FUNCTION

If each customer's outage costs were known to the utility, notifications would be allocated in accord with the optimal decision curve, u, described above. Any additional demand that needed to be interrupted without notification would come from the remaining customers, interrupted in order of their late costs. However, outage costs are private information; while the utility knows the distribution of outage costs in the population, and hence can determine the function u, it cannot identify the outage costs of particular customers, which are needed to administer centrally the selective notifications and interruptions called for by the optimum rationing plan.[3]

The objective of the tariff designer is to develop a price schedule that will induce customers to self-select a service option that will result in the optimal allocation prescribed by the solution to the central planning problem. To achieve this objective, we will consider a tariff schedule in which a customer can choose whether to receive early notification of an impending shortfall or to be standby, and also chooses a priority level if opting for standby status. If choosing early notification, the customer pays k. If choosing standby, the customer pays $p(\theta)+k$ for priority level θ, and receives compensation θ if subsequently interrupted without notification. $p(\theta)$ may be interpreted as a demand charge or an insurance premium. Interruption insurance for electric power is discussed in Chao and Wilson (1987) and Oren and Doucet (1990).

By design, customers opting for notification status will always be interrupted at time T, regardless of the actual magnitude of the shortfall. This contract specification is intended to discriminate among customers with respect to their valuation of early notification, and is essential in order to make the contract enforceable. As with many indirect price discrimination mechanisms, there may be some "dead weight loss" that must be born in order to achieve market segmentation. In short, not supplying power to notified customers is a specification that characterizes the contract structure proposed here.

Each customer chooses the service option that minimizes its expected net interruption cost. A customer with late cost a_T and early cost a_S will sustain a cost a_S if choosing early notification, while if choosing standby at priority level θ, the customer's expected net cost depends on its probability of being interrupted. To allocate interruptions, the utility will use the revealed priority levels as chosen by the customers. It is assumed that customers know this when they choose service options. In other words, for each priority level in the standby price menu, either the utility provides the probability of interruption along with the price, or customers correctly forecast that probability. In addition, each customer behaves as though all other customers were induced to reveal their true priority levels, and its own decision will not affect the utility's allocation scheme. This is an assumption of rational expectations.

With this scheme, the customer's problem is represented as (11).

3. Utilities employ customer surveys to ascertain the aggregate distribution of outage costs.

$$\min_{\theta} \left\{ a_S + k,\ \bar{G}[h(\theta,u)](a_T - \theta) + p(\theta) + k \right\} \tag{11}$$

The first expression (a_S+k) represents the customer's cost for choosing the early notification option; the second expression, the customer's expected cost for choosing standby status at compensation level θ. The curve u in (11) refers to the optimal decision curve described in the previous section. Since the constant k is added to both expressions, the value of k does not affect customer choices. The constant k may be positive or negative. Its value may be set to achieve one or more equity goals, including (i) *revenue neutrality*, to ensure that the electric utility does not profit more from the priority pricing tariff than from random rationing, or a previously existing tariff; and (ii) *Pareto superiority*, to provide all customers at least as better off under the priority pricing tariff as under random rationing, or a previously existing tariff. The constant k may also be set such that the price of interruptible service is competitive with other supply options, such as installing additional peaker resources.

The first-order necessary condition for the optimization problem (11) is (12); the second-order necessary condition is (13).

$$0 = p'(\theta) - \bar{G}[h(\theta)] - g[h(\theta)]h'(\theta)(a_T - \theta) \tag{12}$$

$$\begin{aligned} 0 \le\ & p''(\theta) + 2g[h(\theta)]h'(\theta) + g'[h(\theta)][h'(\theta)]^2(a_T - \theta) \\ & + g[h(\theta)]h''(\theta)(a_T - \theta) \end{aligned} \tag{13}$$

The first-order condition has the standard economic interpretation that at the customer's optimum, marginal cost equals marginal benefit. The marginal cost to the customer is associated with the first term of (12), $p'(\theta)$. The customer's marginal benefit has two components. The second term of (12) is associated with the marginal revenue from selecting a higher compensation level. The third term of (12) is associated with the marginal benefit from being interrupted less often (since standby customers are interrupted in order of increasing compensation level).

The tariff schedule

$$p(\theta) = u(\theta) \tag{14}$$

satisfies (12-13) at the value $\theta = a_T$. Thus, (14) supports the existence of a rational expectations equilibrium. However, this equilibrium may not be the unique one, and the tariff schedule (14) may yield other equilibria that are not socially optimal. By announcing the corresponding probability of interruption

(A.10) along with the tariff at each priority level, the electric utility may facilitate customer formulation of (11), resulting in truthful revelation and the rational expectations equilibrium.[4]

This scheme compensates customers for their true interruption losses. Therefore, customers on the planner's optimal decision curve—those with late cost v and early cost $u(v)$, $0 \leq v \leq 1$—will be indifferent in net expected interruption cost between selecting the early notification option, paying k, and sustaining interruption loss $u(v)$, and selecting the standby option at compensation level v and paying $u(v)+k$. Any customer with late cost v choosing the standby option pays $u(v)+k$. From (5), $u(v)$ is the sum of its expected interruption loss and an externality cost. The compensation scheme reimburses the customer for its own actual interruption losses, resulting in a net payment that is the externality cost plus the constant k.

From the representation of u in (5) and the ensuing interpretation, we see that the result (14) is a manifestation of marginal cost pricing under an externality. By choosing the standby option, a customer with late cost v increases the probability that standby customers with smaller late costs will be interrupted. A customer choosing standby at compensation level v is therefore charged both for its own expected interruption cost and for the extra interruptions it imposes on other customers. The monopoly power of the regulated electric utility enables collection of the priority charge for the externality, so customers self-select the social optimum.

EXAMPLE: UNIFORM SCENARIO, TWO SERVICE OPTIONS

The analysis above utilizes a continuum of priority levels. As shown by Wilson (1989), several discrete priority levels are often enough to capture most of the efficiency gains of priority pricing. Here, the uniform scenario described earlier is further analyzed. Customer outage costs and shortfall magnitude are uniformly distributed, as before.

Two service options are offered: a customer may choose either (i) to be notified and pay price A, or (ii) to be standby, pay price $A + B$, receive compensation θ if interrupted without early notification, and be interrupted without early notification with probability r. If the shortfall exceeds the number of notified customers, the balance between supply and demand is achieved by

4. A customer selecting the early notification option reveals only that its notification and late costs are such that, in the feasible domain, they fall below the optimal decision curve, u. A customer selecting the standby option reveals the actual value of its late cost, say v, as well as the fact that its early cost is greater than $u(v)$.

randomly selecting standby customers to be interrupted without early notification.

The tariff design problem has parameters A, B, θ, and r. Parameter A is the constant that may be used to achieve revenue neutrality, Pareto superiority, or some other goal. Parameter B is the price of choosing the standby option rather than the early notification option. Since only two service options are offered, customers choosing standby do not select a compensation level; instead they receive fixed compensation level θ, as determined by the planner. Probability r is an endogenous function of the self-selection requirement and the constraint that demand cannot exceed supply; hence, r depends on B and θ.

The customer's problem is to choose the service option that minimizes its expected interruption loss. A customer with late cost a_T and early cost a_S faces the problem represented by (15).

$$\min\{a_S + A,\ (a_T - \theta) r + A + B\} \tag{15}$$

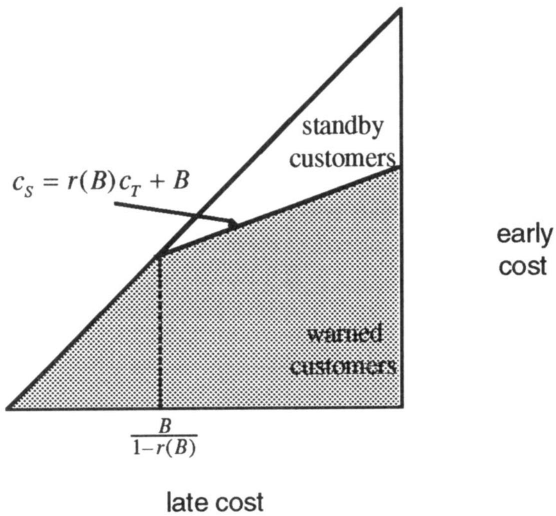
However, the compensation level θ can be set to zero without affecting the optimal social allocation, individual expected interruption losses, or total social welfare. Parameters B and θ determine whether individual customers select early notification or standby. The socially optimal allocation fixes the expected interruption cost of a particular standby customer at some value. To retain the socially optimal allocation, increasing θ will result in a corresponding increase in B. Thus, money is transferred among standby customers, to those standby customers actually incurring an interruption. The realized interruption losses of individual customers are changed, but individual expected losses and total social welfare remain the same. In broad terms, determining compensation level θ is an issue of equity and risk aversion rather than efficiency. To emphasize the efficiency aspects of this problem, compensation level θ is set to zero.

Hence, customers will choose service options according to (16). Figure 3 illustrates customer self-selection. If $B + r(B)$ is greater than one, all customers choose the early notification option.

$$\begin{array}{ll} \text{Choose early notification:} & c_S < r(B)c_T + B \\ \text{Choose standby:} & c_S > r(B)c_T + B \\ \text{Indifferent:} & c_S = r(B)c_T + B \end{array} \tag{16}$$

The allocation problem is to choose the value of B that minimizes total expected interruption loss (17).

Figure 3. Customer Self-Selection



$$N\int_0^{\frac{B}{1-r(B)}}\int_0^{c_T} c_S d(c_T,c_S)\,dc_S\,dc_T + N\int_{\frac{B}{1-r(B)}}^1 \int_0^{r(B)c_T+B} c_S d(c_T,c_S)\,dc_S\,dc_T + N\int_{\frac{B}{1-r(B)}}^1 \int_{r(B)c_T+B}^{c_T} r(B)c_T d(c_T,c_S)\,dc_S\,dc_T \tag{17}$$

Substituting the appropriate uniform densities into (17) yields (18), an expression for total expected interruption loss as a function of B and $r(B)$.

$$N\left(B^2 - \frac{2B^3}{3[1-r(B)]} + \frac{2r(B)}{3} - \frac{[r(B)]^2}{3}\right) \tag{18}$$

Before the allocation problem is solved, $r(B)$ is explicitly written as a function of B. Under self-selection (16), the number of customers notified, $Q(B)$, is given by (19).

$$Q(B)=N\int_0^{\frac{B}{1-r(B)}}\int_0^{c_T} d(c_T,c_S)dc_S dc_T + N\int_{\frac{B}{1-r(B)}}^{1}\int_0^{r(B)c_T+B} d(c_T,c_S)dc_S dc_T \tag{19}$$

Substituting the appropriate uniform densities into (19) yields (20).

$$Q(B) = N\left(r(B) + 2B - \frac{B^2}{1-r(B)}\right) \tag{20}$$

Given random rationing within the class of standby customers, the probability that a standby customer is interrupted without early notification is given by (21).

$$r(B) = \int_{Q(B)}^{N} \frac{y-Q(B)}{N-Q(B)} g(y)dy = \frac{N-Q(B)}{2N} \tag{21}$$

Substituting the right side of (20) into (21) yields (22), an equation for $r(B)$.

$$r(B) = \frac{1}{2}\left(1 - r(B) - 2B + \frac{B^2}{1-r(B)}\right) \tag{22}$$

Expression (22) can be rewritten as an equation quadratic in r. Since $B + r(B)$ must be no greater than one for (19) to be valid, the relevant root is (23).

$$r(B) = \frac{1}{3}\left(2 - B - \sqrt{1+2B-2B^2}\right) \tag{23}$$

With $r(B)$ written as an explicit function of B, the right side of (23) can be substituted into (18). This messy expression is not written, but is solved numerically for the optimum B. The results are as follows:

B^*	$= 0.172$	$r(B^*)+B^*$	$= 0.403$
$r(B^*)$	$= 0.232$	Total Expected	
$Q(B^*)$	$= 0.537$	Interruption Loss	$= 0.162$
$\dfrac{B^*}{1-r(B^*)}$	$= 0.223$	Ratio	$= 1.096$

Ratio compares the total expected interruption loss for this menu with only two options to the optimal solution yielded by a continuous menu.[5] The efficiency loss (the increase in interruption loss because of the simpler menu) is less than 10%. Fewer customers are notified when the simpler menu is used (54%, compared with 61% from Table 1). For the uniform scenario, the menu with only two options appears to gain in simplicity much more than it loses in efficiency.

Wilson (1989) derives the result that a menu with k options yields an efficiency loss of $\mathcal{O}(1/k^2)$. This result is valid for a one-dimensional model; with respect to the two-period model discussed in this paper, it is as if there were no early notifications at time S, only interruptions at time T. Nevertheless, since only standby customers are segmented under the two-period model, while customers notified at time S are dispatched alike, it seems that Wilson's result would apply here. Real priority menus thus offer a handful of choices, yet retain most of the social welfare benefits indicated by our analysis.

CONCLUSIONS

The priority pricing menu developed here extends the models in the priority service literature by allowing an early notification option. If no notification is allowed, the decision curve u disappears, and the standard one-dimensional priority service model applies. Compared with the standard model, our model results in smaller expected social cost due to capacity shortfalls.

Our analysis is for a single shortfall. An electric utility may experience a number of shortfalls during an operating year. The exact number of shortfalls, and exactly when each shortfall will occur, is not known in advance. Furthermore, shortfalls may be of different types: a number of short-term uncertainty distributions, G, may exist; shortfall k may have distribution G_k. This may complicate solving for the optimal allocation and implementing this allocation through priority pricing.

One of our assumptions is that notified customers incur their full interruption loss minus the notification benefit; in other words, once issued, notifications cannot be voided. Allowing the electric utility to void notifications reduces the impact of short-term uncertainty in shortfall magnitude, resulting in greater social welfare. Allowing notifications to be voided may change the optimal allocation and the resulting priority pricing menu.

Finally, our analysis of interruptible electric service with an early notification option may apply to other services with reservation options. Travel

5. Refer to the uniform scenario example in the section on the optimal decision curve.

services such as airlines, car rentals, and hotels all offer differential prices that depend on advance reservation by the customer. Profit maximization distinguishes those situations from the analysis presented here.

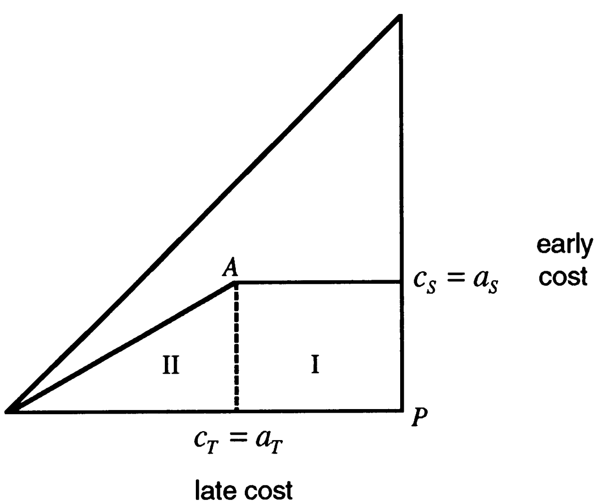
APPENDIX: DERIVATION OF OPTIMAL DECISION CURVE

Existence of Utility's Decision Curve

We first show that the optimal set of customers to be notified is connected, and that the boundary between the set of notified customers and the set of standby customers may be expressed as a (smooth) curve u. Then, at time T, standby customers may be interrupted as needed, in increasing order of late costs.

Consider customer A, with late cost a_T and early cost a_S. Suppose w, the set of notified customers under the optimal notification policy, contains A. Then any customer with costs $c_S \leq a_S$ and $c_T \geq a_T$ (region I of Figure 4) should also be notified; otherwise, the total expected interruption cost can be reduced by notifying such a customer instead of A. In this sense, region I *dominates* A.

Figure 4. Regions I and II Dominate *A*



The fact that region I dominates A yields three properties of w and its boundary:

1. w is connected.

   The line segment joining A and P—the point (1,0)—is a continuous path lying in w. Hence any two points A and B in w may be connected by the path along the line segments $\overline{AP}$ and $\overline{PB}$. Therefore, w is connected.

2. The boundary of w can be described by a function $u(\cdot)$ with c_T as its domain and c_S as its range.

   Suppose $u(a_T) = a_S$. All customers (a_T,c_S), $c_S < a_S$, lie in region I and should be notified, while any customer (a_T,c_S), $c_S > a_S$, cannot be notified, lest all customers between (a_T,a_S) and (a_T,c_S) also be notified, in which case $u(a_T) = c_S$ and not a_S.

3. u is nondecreasing.

   Suppose $A \in w$. For $c_T > a_T$, (c_T,a_S) is in region I, hence, in w. Therefore, $u(c_T) \geq a_S$.

Region II also dominates A. If $A \in w$ but some customer Q with late cost $c_T < a_T$ and notification benefit factor $c_S/c_T \leq a_S/a_T$ (region II of Figure 4) is not notified, then the total expected interruption cost can be reduced either by notifying Q instead of A or by notifying Q in addition to A.

The fact that region II dominates A yields continuity of u:

4. u is continuous.

   Suppose $0 < \delta < \epsilon$. From Region II,

   $$u(a_T+\delta) \leq (a_T+\delta)\, u(a_T)/a_T = u(a_T) + u(a_T)\, \delta/a_T.$$

   Hence, $0 \leq u(a_T+\delta)-u(a_T) \leq \delta\, u(a_T)/a_T \leq \delta < \epsilon$.

This completes the development of the utility's planning curve. Assuming u to be twice differentiable facilitates using the calculus of variations to derive the optimal decision curve. The derivation is described next.

Derivation of Utility's Optimal Decision Curve

We begin with (A.1), the total expected interruption cost. The first term is (1) in the main text; the second term is (4).

$$N\int_0^1\int_0^{u(x)} y d(x,y)dydx + \int_0^1 \left(N\int_0^v x\int_{u(x)}^x d(x,y)dydx\right) dG[h(v;u)] \tag{A.1}$$

Applying integration by parts and the substitution

$$w(v) = N\int_0^v\int_0^s\int_0^{u(x)} d(x,y)dydxds \tag{A.2}$$

yields total expected interruption cost (A.3).

$$\int_0^1 \left\{ \begin{matrix} uw'' - N\int_0^{u(v)}\int_0^x d(v,y)dydx \\ + \\ \dfrac{\left(NvD(v,v) - vw' - Nv\int_0^v D(s,s)ds + w\right)}{\dfrac{g\left(ND(v,v) + w'(1) - w'\right)}{\left(N\int_0^v d(v,y)dy - w''\right)}} \end{matrix} \right\} dv \tag{A.3}$$

Let I represent the integrand in (A.3) and ξ denote the difference between the left and right sides of (A.2). With the Lagrange multiplier $\lambda(v)$ attached to constraint ξ, the Euler equations are (A.4) and (A.5). Boundary condition is (A.6) and transversality condition is (A.7).[6]

$$0 = I_u + \lambda\xi_u = w'' - N\int_0^{u(v)} d(v,y)dy + \lambda(v)\left(-N\int_0^v\int_0^s d[x,u(x)]dxds\right) \tag{A.4}$$

6. The appropriate transversality condition is

$$0 = \left[I_{w'} - \frac{dI_{w''}}{dv}\right]_{v=1}$$

This reduces to (A.7).

$$0 = I_w + \frac{d}{dv}\left(\frac{dI_{w''}}{dv} - I_{w'}\right) + \lambda \xi_w \tag{A.5}$$

$$u(0) = 0 \tag{A.6}$$

$$u'(1) = 0 \tag{A.7}$$

Differentiating (A.2) twice with respect to v yields w'' equal to the second term on the right side of (A.5), with the result that the Lagrange multiplier $\lambda(v)$ is zero for all values of v. Algebraic manipulation of the partial derivatives of I yields (A.8).

$$u'' = \frac{d}{dv}\left(\frac{dI_{w''}}{dv} - I_{w'}\right) \tag{A.8}$$

The second term of (A.5) is replaced with (A.8), while I_w is written explicitly, transforming the Euler equation (A.5) into (A.9), an expression in terms of the original function u.

$$0 = \frac{dG[h(v;u)]}{dv} + u'' \tag{A.9}$$

Equation (A.9) indicates that u is not positive, hence u is concave.

Integrating (A.9) yields (A.10), an implicit expression for the desired curve u in terms of the given functions D and G. The constant of integration is zero, as determined by the transversality condition (A.7).

$$u'(v) = \bar{G}[h(v;u)] \tag{A.10}$$

The expression on the right side of (A.10) is the probability that a standby customer with late cost v will be interrupted. After integrating (A.10) with respect to v, applying the boundary condition (A.6), and performing some mathematical manipulation, one obtains (A.11).

$$u(v) = v\bar{G}[h(v;u)] + \int_0^v z dG[h(z;u)] \tag{A.11}$$

Equation (A.11) is a functional equation. Corresponding to a particular distribution D and distribution G is a particular function u that satisfies the optimality conditions (A.5-7). Equation (A.5) reduces to a second-order ordinary differential equation, hence a general analytic solution may not exist, although analytic solutions may be found for some particular D and G. When an analytic solution cannot be found, (A.5) may be solved numerically for the function u.

REFERENCES

Chao, H.P., S.S. Oren, S.A. Smith, and R.B. Wilson (1986). "Multilevel Demand Subscription Pricing for Electric Power." *Energy Economics* 8: 199-217.

Chao, H.P., and R. Wilson (1987). "Priority Service: Pricing, Investment and Market Organization." *American Economic Review* 77: 899-916.

Christensen Associates, L.R., Inc. (1988). *Customer Response to Interruptible and Curtailable Rates. Volume 1: Methodology and Results.* Electric Power Research Institute Report EM-5630.

Ebasco Business Consulting Company (1987). *Innovative Rate Design Survey.* Electric Power Research Institute Report RP2381-5.

Elsgolc, L.E. (1962). *Calculus of Variations*, Reading, Massachusetts: Addison-Wesley.

Hamlen, W.A., Jr., and F. Jen (1983). "An Alternative Model of Interruptible Service Pricing and Rationing." *Southern Economic Journal* 49: 1108-1121.

Marchand, M. G. (1974). "Pricing Power Supplied on an Interruptible Basis." *European Economic Review* 5: 263-274.

Mussa, M., and S. Rosen (1978). "Monopoly and Product Quality." *Journal of Economic Theory* 18: 310-317.

Oren, S.S. (1990). "Pricing of Interruption and Resumption Priorities for Triangular Shortfalls." Chapter 10 of *Service Design in the Electric Power Industry*. Electric Power Research Institute Report P-6543.

Oren, S.S., and J.A. Doucet (1990). "Interruption Insurance for Generation and Distribution of Electric Power." *Journal of Regulatory Economics* 2: 5-19.

Panzar, J., and D. Sibley (1978). "Public Utility Pricing Under Risk: The Case of Self-Rationing." *American Economic Review* 68: 888-895.

Smith, S.A. (1989). "Efficient Menu Structures for Pricing Interruptible Electric Power Service." *Journal of Regulatory Economics* 1: 203-223.

Tschirhart, J., and F. Jen (1979). "Behavior of a Monopoly Offering Interruptible Service." *Bell Journal of Economics* 10: 244-258.

Viswanathan, N., and E.T.S. Tse (1989). "Monopolistic Provision of Congested Service with Incentive Based Allocation of Priorities." *International Economic Review* 30: 153-174.

Wilson, R. (1989). "Efficient and Competitive Rationing." *Econometrica* 57: 1-40.

Woo, C.K. (1990). "Efficient Electricity Pricing with Self-Rationing." *Journal of Regulatory Economics* 2: 69-81.

Woo, C.K., and N. Toyama (1986). "Service Reliability and the Optimal Interruptible Rate Option in Residential Electricity Pricing." *The Energy Journal* 7(3): 123-136.