Order Number 9316164

Reliability-based pricing of electricity service

Hegazy, Youssef A., Ph.D.

The Ohio State University, 1993

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

RELIABILITY-BASED PRICING OF ELECTRICITY SERVICE

DISSERTATION

Presented in Partial Fulfillment of the Requirements
for the Degree Doctor of Philosophy
in the Graduate School of The Ohio State University

by

Youssef Hegazy, B.Sc., M.Sc.

* * * * *

The Ohio State University

1993

Dissertation Committee:

Dr. Jean-Michel Guldmann
Dr. Philip A. Viton
Dr. Douglas N. Jones

Approved by:

Advisor
Department of City and
Regional Planning

Acknowledgements

I express sincere appreciation to my advisor, Dr. Jean-Michel Guldmann, for his guidance, help and encouragement throughout my graduate studies at The Ohio State University.

I would also like to express my appreciation to Dr. Philip A. Viton for his useful suggestions as member of the reading committee.

I wish to thank Dr. Douglas N. Jones for his support and many useful suggestions during this research.

I am grateful to Dr. Narayan S. Rau and Dr. Robert J. Graniere for their help on many occasions.

For my wife, Eman, I offer sincere thanks for your patience and encouragement.

VITA

1977 . . . . B.SC., Nuclear Engineering Department, University of Alexandria, Egypt.

1978-1984 . . . . . . .Engineer, Research and Planning Department, Egyptian Electricity Authority, Cairo.

1985-1986 . . . . . . .M.Sc., Energy Management and Policy, University of Pennsylvania, Philadelphia.

1987-Present . . . . .Research Associate, The National Regulatory Research Institute, The Ohio State University, Columbus, Ohio.

FIELDS OF STUDY

Major Field: City and Regional Planning.

Minor Fields: Energy Planning, Energy Economics, Power System Analysis and Production Cost Simulation.

Table of Contents

Tables

Figures

Figure | Page

Chapter I

Introduction

I.1 Objectives of the Research

The objectives of this research are (a) to develop a price structure that unbundles electricity service by reliability levels, and (b) to analyze the implications of such a structure on economic welfare, system operation, load management, and energy conservation.

I.2 Need for the Research

Over many decades, progress in power generation and transmission technology has greatly contributed to the efficiency of electric power systems. The development of interconnected, centrally dispatched systems has enabled utilities to achieve further efficiency by pooling loads that vary randomly over time and location and to take advantage of the different operating and cost characteristics of various supply sources. As a result, utilities have been able to provide all customers with a highly reliable service at reasonable prices.

Recently, however, because of uncertainties in demand growth, escalating costs and regulatory uncertainty with respect to the recovery of investment costs, many utilities have become reluctant to commit capital to expand their base load capacity.

As a result, the reliability of supply can no longer be taken for granted. Power shortages may be inevitable. Unexpected interruptions, if they become commonplace, would have an extremely negative impact on customer satisfaction.

Although customers have become accustomed to highly reliable service for all classes of applications, this does not necessarily imply that all applications require it. Recent studies, for example [Billinton, 1986,1987] and [EPRI,1989], have found that the value different customers place on reliability varies widely from several pennies to tens of dollars per kilowatt-hour. Thus, some customer applications could sustain a substantial decrease in reliability in return for appropriate cost discounts, while service with reliabilities higher than those currently offered can be justified for other applications. This diversity indicates that latent customer demand exists for different levels of reliability.

The advancement of metering and control technology makes it feasible for electric utilities to offer new service options to customers [Rosenfield, 1985]. One such class of options is services with explicitly differentiated levels of reliability. In this approach, customers are given a menu of reliability options with correspondingly adjusted charges. The customer then can assign various blocks of load to each of the various options for a specific contract period. Load increments may be identified by individual circuits or by specially designed outlets. Several currently available microcomputer-based metering technologies are capable of implementing this type of service.

Given the capability to offer new types of products, an important issue is how the products should be designed to best serve customers' and utilities' needs. For customers, differentiated products should match the needs of specific applications. For utilities, reliability-based product differentiation should allow them to operate with less reserve margin and thus reduce the average cost of energy over the long term. Interruptible service contracts can substitute for future capacity expansion in some instances. At the same time, cut-rate, low reliability power can encourage sales of unused capacity.

Utilities have many years of experience in offering interruptible service to selected industrial customers. Several experimental programs with broader service offerings are currently under way [Chamberlain, 1985]. Public utility commissions are, in some cases, requesting electric utilities to offer service reliability options to all customers. The Massachusetts Department of Public Utilities ordered Boston Edison (Docket 84-194) to file new interruptible rates to be offered to all customers, stating that customers should be offered a range of services at a range of prices, with low prices having longer and more frequent interruptions than higher prices" [Electric Utility Week, June 17, 1985].

The key aspect in designing a service is pricing. Pricing should reflect the utilities' cost differentials of supplying the various service options so that customers are given correct price "signals" for planning future energy uses [Joskow and Schmalensee, 1986]. The most socially efficient prices in the economic sense are those that maximize total welfare (for example, spot pricing and real-time pricing). Basic economic theory demonstrates that these prices should be equal to marginal costs in a static market equilibrium, but spot market pricing or real-time pricing schemes either fail to recover sufficient revenue to meet the utility's revenue requirement, or overrecover revenue so that they violate regulatory requirements. Schemes to recover the allowable revenue and returns suggested so far, for example, Ramsey-type pricing, result in loss of welfare.

Recently, however, several pricing models have been introduced in the literature that suggest that electric utility customers be allowed to choose different levels of service reliability [EPRI, 1986,1989]. These models argue that such a choice would result in improved efficiency and benefits to customers and utilities. However, some of these schemes introduce economic inefficiencies and fall short of attaining welfare maximization [Rau, 1990]. That has initiated some debate about the usefulness and implementability of reliability-based pricing.

Our review of the reliability-based pricing literature has demonstrated two major shortcomings: (a) supply (power) systems are represented by cost functions that are inaccurate in representing the operations of power systems and the random-binary nature of the systems' components, and (b) these models yield pricing structures that are cumbersome to implement and require highly sophisticated and costly electronic and metering systems.

The supply system is customarily represented by a predetermined cost function independent of any reliability parameter. An electric power network is a highly complex and operationally uncertain system. At any moment there is a possibility of total or partial system outage, leading to immediate changes in system topology and, hence, changes in the form of the cost function. For example, to accurately represent a supply system with N generating units, one needs 2^N different cost functions to actually represent the supply system, each function representing a different possible system configuration, because each unit is either available or out of service. Such a supply system should be stochastically simulated to account for all the possible random changes.

The commonly used paradigm to characterize consumers' choices is that willingness to pay reflects differences in tastes across the population of individuals. This approach, though sophisticated, leads to optimal schedules that are implicitly defined, and to a continuum of options [Chow, 1989]. In this context, priority pricing is an extremely difficult scheme to implement from the dispatching and operational point of view. The ability to monitor and control each increment of demand at each moment is an electronic utopia. However, some differentiation programs that utilities are using now (such as interruptible/curtailable service, demand subscription rates, and direct load control) are successful from the implementation point of view because they offer a limited number of options that the power system can practically provide. The problem with these programs is that they may lead to economic inefficiency, because they are based on an arbitrary rationing scheme.

I.3 Research proposed

It is evident from the above discussion that there is a potential benefit to electric utilities and their customers in differentiating electricity service by offering different options of reliability. It is also evident that there is a need for a price structure that can achieve these benefits. This need gave the impetus for developing this research. As the research proceeded, it became apparent that more effort should be directed toward evaluating the probability distribution of system marginal costs. This evaluation is part of the research presented in this dissertation.

The proposed pricing mechanism combines priority (reliability differentiation) pricing with real-time pricing of electricity. Under this scheme, customers with different reliability preferences are charged different prices. The utility is assumed to be a single welfare-maximizing firm able to set and communicate prices instantly. At times of supply shortages, the utility has direct control over customer loads and follows a rationing method among customers willing to accept power interruptions. Therefore, customers are given the choice either to be served with a high reliability "firm" service, or to be subject to interruption. The first choice means customers are served at all times and circumstances with 100 percent continuous service. The second choice means customers are allowed to select their preferred interruption level (probability and magnitude of interruption) in advance from a priority-price menu provided by the utility.

The utility is obliged to comply with the regulatory revenue constraint. Therefore, we propose a mechanism that makes every customer class pay its share of the costs of service (operating costs plus fixed costs) according to the ratio of their consumption to total demand. The mechanism also makes each customer class, in time of supply shortage, pay (or gain) a sum of money relative to his/her chosen level of reliability. Customers who opt for "firm" service will have to pay for that service. Those who opt to be interrupted at shortage time will be compensated. The payment/compensation concept is designed to minimize social costs and damages due to power outages. A social outage cost (SOC) value is introduced as the mean (average) value of customers' outage costs to make any payment or compensation profitable to its destined customer. It is obvious that SOC as a payment is lower than the outage cost of the high reliability customer, and higher than the outage costs of the low reliability customers. This mechanism assures a subsidy-free cost-sharing allocation.

The proposed model simulates these processes in a stochastic real-time framework. It is essentially a Ramsey-real-time pricing scheme and deviates from Ramsey-real-time only when the probability of having an outage is above zero. It is superior to Ramsey-type pricing because it minimizes the social damage of power shortages. The important task of this research is to compare the implication of the proposed pricing mechanism on economic welfare, system operation and reserve, load management, and energy conservation with other traditional pricing mechanisms.

I.4 Research Performed

We have examined the welfare gain and energy and reserve saving possibilities due to different pricing schemes. The base case for comparison was that of the traditional spot pricing of electricity. Certain assumptions regarding price-demand relationships, outage costs, and consumption patterns were made. The analysis has simulated the effects of different pricing schemes on four customer classes. These schemes are: spot pricing, Ramsey pricing, and the proposed reliability-based pricing. Benefits due to these pricing schemes are compared and analyzed. Spot pricing and Ramsey pricing are each simulated in two different modes: (a) without consideration of outage costs and reliability impacts, and (b) with consideration of outage costs. Also, the proposed reliability-based pricing is simulated in two different cases: (a) two-reliability-options, and (b) three-reliability-options. Each option refers to a different level of reliability. Therefore, we have six different pricing schemes to compare: (1) traditional spot pricing [SPOT1], (2) traditional Ramsey pricing [RAMSEY1], (3) non-traditional spot pricing (outages are considered), [SPOT2], (4) non-traditional Ramsey pricing (outages are considered), [RAMSEY2], (5) two-options reliability-based pricing [RDP-2], and (6) three-options reliability-based pricing [RDP-3].

The results show that reliability-based pricing yields higher economic efficiency (welfare) and energy and reserve savings than both non-traditional spot and Ramsey pricing. Traditional spot and Ramsey pricing assume either an outage-free world or a costless-outage society. Both assumptions are impractical and discredit these pricing schemes, though they misleadingly yield higher economic efficiency.

I.5 Research Assumptions

Several assumptions have been made to conduct this research:

a) Only generation system reliability is considered. Transmission and distribution system reliabilities are ignored.

b) Two parameters of system reliability are utilized in the model as an indication of system reliability. They are the loss of load probability (LOLP), that is, the probability (percentage of time) that the system cannot serve the total demand, and the unserved energy (UE), that is, the expected amount of energy that the system cannot serve due to outages.

c) A stochastic production cost simulation model is implemented, in which the stochastic nature of supply is considered.

d) A customer's choice model is introduced based on the idea that the reliability level of a service represents the risk involved in the customers' benefit from the service.

e) The models referred to in (c) and (d) are integrated under a price-regulation mechanism.

I.6 Research Contribution

The most significant contributions of this research are:

(a) The verification of the feasibility and usefulness of the notion of unbundling electricity service and pricing the service according to the level of reliability associated with it.

(b) The introduction of an accurate and implementable price structure.

(c) The development of a method to estimate the probability distribution of the system marginal cost (lambda). The estimated distribution is not only extremely important to this research, but has several other important uses with respect to system operation, energy pricing, energy exchange, and load management incentive programs.

(d) The development of a consumer's choice model, in which a reliability dimension is introduced to represent the notion of risk involved with customer satisfaction from the (less than reliability-perfect) electricity service.

(e) The interaction between the consumer's choice model and the supply model through the stochastic production cost simulation model. This allows us to account for supply randomness.

I.7 Organization of this Dissertation

Chapter II gives an overview to the key elements related to reliability-differentiated pricing. Chapter III reviews the literature on electricity pricing and related reliability-based pricing. Chapter IV discusses the proposed concepts and illustrates the model and its mathematical formulation. Chapter V presents the data used for the case study, the results, and their analysis. Finally, conclusions, and recommendations are outlined in chapter VI.

Chapter II

Background

II.1 Price Structures

Electric utilities, consumers, and society as a whole have different perspectives on price structures and ratemaking.

Utilities expect to be fully compensated for the cost of providing service; that is, revenue requirements must be met. Revenues must be sufficient to cover capital and operating expenses, and investor-owned utilities want rates to incorporate a reasonable return on their capital investment. Similarly, publicly-owned utilities want to be financially self-sufficient and not rely on subsidization from other revenue sources. From the utility's perspective, ratemaking is also strategic with regard to its ability to provide service using existing capacity and to plan for future capacity additions. Predictable revenues and flexible rate structures are strategically advantageous to the public utility, particularly if it faces any form of competition, including bypass and self-supply.

For consumers, the ratemaking process and resultant rates should be equitable and fair. This usually means that charges to specific types or classes of customers should be based on the costs of serving them and not on arbitrary or discriminatory criteria. Consumers also prefer rates they perceive to be affordable, an increasingly difficult expectation to meet. They also prefer a rate structure that is understandable, which presumably improves consumption decisions. Consumers' understanding and acceptance of utility rates make the task of rate-making much easier.

Society's perspective differs from that of utilities or consumers. Economic or allocative efficiency is a societal goal having to do with costing and pricing. Rates based on efficiency goals encourage appropriate levels of production and consumption and discourage the misallocation of societal resources. In the context of efficiency, society has an interest in conserving resources. Conservation emphasizes the correct valuation and allocation of resources. Ratemaking, however, can send signals about social priorities. Society may place a priority, for example, on electricity for heating use during a severe winter, over electricity for commercial or industrial uses, and this may be reflected in pricing schemes in the form of subsidization. Finally, society may judge rate-making in terms of whether it is just and reasonable, a time-honored standard in utility regulation. Hence, society may want prices that are not unduly discriminatory.

Ratemaking is thus a continual balancing act among the divergent and often competing perspectives of utilities, consumers, and society. Rates that are perceived by consumers to be affordable do not necessarily meet revenue requirements; rates that are equitable are not necessarily efficient; rates that are economically efficient are not necessarily administratively feasible because of practical application issues. In balancing these perspectives, the key objectives of rate regulation emerge. One of the main objectives of this research, then, is to test the hypothesis that reliability-based pricing is economically efficient and implementable.

II.2 Reliability

Reliability of electricity supply means different things to different people. On the supply side, generation planners generally talk in terms of the "adequacy" of the power supply, which can be defined as the ability to meet the load with a certain degree of reliability and without resorting to the use of emergency operating procedures. A shortage in this context is viewed as a condition of inadequacy. The adequacy of a generation system is a function of the generation mix, system topology and interconnections, and the characteristics of demand. Commonly used indices that reflect the degree of adequacy include the reserve margin and the loss of load probability (LOLP)[1]. Another definition of reliable service can be cast in terms of the loss of energy probability (LOEP). This index measures the expected fraction of total energy sales the utility will be unable to serve due to generation shortages. This index would be preferable to the LOLP index if duration and magnitude of outages were the sole determinants of a reliable service level. LOLP is the preferred indicator if frequency of interruption were the sole criterion.

Regulatory agencies, in their public interest guardian capacity, have been entrusted with the legal mandate to ensure that the utilities they oversee provide service to anyone requiring it in a nondiscriminatory fashion, at regulated prices, and

[1] Loss-of-load probability (LOLP): The probability of the system load exceeding the available generating capacity.

with a high degree of reliability. These agencies view reliability in terms of shortage. Accordingly, a shortage is a deviation from a benchmark. Specifically, a shortage is a situation in which supply falls short of demand at currently regulated prices. Thus, the benchmark is the demand at currently regulated prices and at an "acceptable" quality of service. However, in the absence of satisfactory methods available for objectively defining what constitutes reliable service, most regulatory agencies intervene only when consumers complain that the service is "unsafe, improper, inadequate, or insufficient."

From the consumer's perspective, the source or cause of service unreliability is irrelevant; the cost of a service interruption is the main concern. The cost of an outage is a function of the time of day and duration of the outage, the activities affected, the degree to which those activities depend upon electricity, the availability of a backup power source, the ability to resume the affected activity normally after power is restored, the frequency of the outages, and a host of other descriptors. Consequently, the cost of an outage is different for each consumer. From the consumer's perspective, key outage descriptors include:

- Time of occurrence
- Duration of interruption
- Magnitude of outage
- Warning time

and

- Frequency of interruptions

An ideal reliability-based pricing structure should account for all these factors. Because there is no unique reliability index that can capture all the factors involved, some assumptions and compromises have to be made. In this research we assume that the information about outage magnitude and probability is enough to build an implementable price structure. The ultimate test for the proposed structure is to prove its superiority over other pricing schemes in satisfying customer, utility, and societal goals.

II.3 Product Differentiation

Treating electric power service as a family of differentiated products is an approach proposed and tested in a variety of contexts. The economic rationale for such differentiation is the variations in customer preferences. Product differentiation with correspondingly adjusted prices can benefit both the consumer and the producer by matching service options more closely to customer preferences, and by allowing more efficient use of generation capacity. The availability of metering options using microprocessor technology has created new opportunities in which electric power product differentiation is both feasible and desirable [Rosenfield, 1985].

Efficiency gains from differentiation are becoming common place in the regulated industries. For technical and historical reasons, regulated firms initially offered undifferentiated services of essentially uniform quality. Regulatory policies encouraged firms to emphasize supply-side economies of scale. Pricing was based primarily on cost recovery, including normal profits but excluding monopoly profits.

This approach has contributed to efficiency for the products actually offered, but it ignored the additional efficiency gains obtainable from the differentiation that adapts product designs to various market segments. Imperfect regulation, such as rates of return allowed to exceed the cost of capital, encouraged firms to emphasize capital investments and indirectly to maximize quality attributes relying on capital-intensive technologies [R.WILSON,1990].

Some of these regulatory policies were changed during the past fifteen years, resulting in substantially altered incentives for firms in regulated industries. In part these changes are reflected in lower allowed rates of return on capital, and in part they are reflected in an overall relaxation of regulatory supervision and a general presumption in favor of deregulation wherever possible. In addition, in the power industry environmental concerns and measures to economize fuels have focused attention on the inefficiency of serving all customers in peak periods with fuel-intensive peaking generators and at prices below actual marginal costs.

These changes have produced a considerable emphasis on product differentiation. Marketing unbundled products or services has long been common practice in many industries. The unbundled products typically are offered as product lines defined by various combinations of product attributes. The distinguishing characteristics are usually those for which customers' preferences are diverse and suppliers' costs differ significantly. Products are priced to present customers with a variety of quality/price tradeoffs. Each customer can then choose the combination that maximizes his benefit minus cost. Examples of the effects of product differentiation in some other industries that have had related experiences are discussed in the following.

In the securities market, for instance, customers' diversity appears in their risk preference, tax brackets, and need for liquidity. Various securities and investment opportunities offer different levels of risk, liquidity, and tax benefits, which are the product quality attributes. The market price for each security is determined by its relative supply and the distribution of customer preferences. Consequently, each customer is faced with a menu of quality/price combinations, from which he can choose an investment portfolio that is optimal with respect to his preferences.

The photocopier market offers another example of product differentiation. Customer diversity in this case is manifested primarily through usage patterns (peak and base load), copy quality requirements, and total volume. Products are differentiated according to machine capacity and speed, operating cost, and copy quality. Hence, customers face a menu of capital versus operating cost tradeoffs. A customer's optimal choice will depend on his volume, load pattern, and valuation of other features bundled in different machines.

Facsimile transmission services offer different rates for night versus day service, and for fast versus slow delivery. Customers who differ in the urgency of their transmission can choose the quality/price combination that best serves their needs. Other ways of product differentiation in telecommunications are voice versus data lines, trunk versus single lines, night versus day rates, store and forward service, WATS lines, and private versus party lines.

In transportation systems, particularly in the airline industry, customers are offered fares that correspond to a variety of service quality attributes such as class (first, business, coach), advanced commitment and cancellation privileges, length of stay, direct versus indirect flights, standby, charters, overbooking, and bumping. This menu of service quality options and their corresponding fares respond to customers' preference diversity, allowing them to make their own price/quality tradeoffs. At the same time, it allows airlines to better utilize their capacity, reduce their operating costs, and increase their sales and competitive effectiveness.

There are strong analogies between the transportation and telecommunication industries and the electric power industry. In all three cases, the product is non-storable and capacity is dictated by peak loads. Furthermore, customers' preference of diversity for delivery conditions offers opportunities for unbundling and product differentiation. The analogous attributes for electric power product differentiation are time of use, interruptibility, cycling (of air conditioners), voltage, bulk power, temperature cutoffs, warning time before interruption, and interruption insurance. Most of these attributes are already used in rate structures offered to industrial customers and in experimental residential rates.

II.4 Implementation Forms

Essential to the implementation of product differentiation are control and metering capabilities needed for monitoring and enforcement. In the airline industry, these capabilities are provided by the reservation and ticketing system, while in the telecommunications industry they are provided by switching and recording circuitry.

Until recently, metering and control of electric power loads at a sufficiently detailed level to enable product differentiation were economical and technologically feasible only for large industrial customers. The advances in microelectronics over the last decade may now allow economical metering and control at the residential level. This provides opportunities for the electric power industry to follow the lead of other industries and differentiate its product further.

In the following, we discuss the implementation forms of product differentiation that are either traditionally practiced or recently developed by electric utilities.

(a) Traditional Forms

Numerous innovative rate and service options that unbundle the characteristics of electric power have been introduced into the industry. A recent review [EPRI, 1989] of innovative rate designs being employed by the electric power industry lists three programs that tie the price of electricity to the reliability of service: (1) interruptible/curtailable rates, (2) demand subscription rates, and (3) direct load control. In addition, Real-Time Pricing (RTP) and Priority Service programs have been introduced recently. RTP programs do not explicitly vary the level of reliability, but customers' responses to the real-time prices can be used to infer the value of service reliability. Each of these programs is briefly described below.

(1) Interruptible/Curtailable Service

Interruptible/curtailable (I/C) service is a form of service differentiation that has been in use in the electric power industry for over 35 years. Interest in I/C service has increased substantially in the last 15 years. Between 1972 and 1986 the number of utilities with I/C programs has grown fivefold, with 71 percent of the large investor- owned utilities reporting I/C programs by 1986 [EPRI, 1989]. Under an I/C program, the customer specifies a maximum level of demand known as the firm power level (FPL). The customer can utilize electricity service up to the FPL as if standard service applied. However, usage above the FPL is subject to interruptions. In exchange for allowing a portion of its usage to be interrupted, the firm receives a bill discount, typically in the form of demand charge credit. I/C service is typically available as an optional tariff or service rider in the commercial and industrial sectors.

Interruptible/curtailable rates represent a simple form of priority service in which customers divide their loads into two reliability segments. Usage up to the firm power level is serviced at the standard reliability level. Usage above FPL (interruptible power) is serviced at a reduced reliability level, determined, in part, by contractual limits on the frequency of interruptions and the duration of interruptions.

(2) Demand Subscription Rates

Demand Subscription Rates (DSR) are any rate through which a customer receives a rebate or credit for subscribing to a predetermined maximum demand level that cannot be exceeded. In their basic form, DSR programs bear little resemblance to product differentiation because the customer can never purchase any amount greater than the contracted level. More advanced forms of DSR exist for which the subscription level is not always in force. In these cases, the program becomes similar to interruptible/curtailable rates. According to EPRI Innovative Rate Surveys (1989), there have been few implementations of demand subscription rates to date.

(3) Direct Load Control

The objective of product differentiation pricing is to provide customers with a spectrum of service quality options to which they can assign segments of their total load. In practice, this segmentation may take the form of dividing total usage by end-use, with different end-uses assigned different reliability levels. In this case, it is important to know how the value of service reliability varies by end-use. Direct Load Control (DLC) programs offer varying reliability levels for a particular end-use and, as such, are a potential source of information on reliability needs associated with those end-uses.

Direct Load Control programs are similar in structure to I/C rates. In exchange for a bill reduction, the consumer agrees to have a portion of his load interrupted by the utility, with limits placed on the number and duration of interruptions. However, DLC programs differ from I/C rates in two key ways. First, as noted above, DLC programs are targeted at specific end-uses, typically air conditioning and water heating. Second, DLC programs usually involve a cycling strategy. For example, under a residential DLC program, a household's air conditioner might be interrupted every other half-hour for a six-hour period. The objective of the cycling strategy is to reduce the effect of DLC activations on the customer. Unfortunately, this will also reduce the analyst's ability to infer end-use-specific outage costs from DLC programs, since the customer may perceive little or no change in service. Over the past decade, direct load control programs have become increasingly popular in the electric power industry, particularly with improvements in the technology used in their implementation [EPRI, 1989].

(b) Advanced Forms

(1) Real Time Pricing

Real Time Pricing (RTP) refers to a class of rate structures in which the price of electricity changes frequently to reflect current information on system costs. The extreme form of RTP, "spot pricing," allows the price of electricity to change instantaneously with no limitations on its level or variability over time. RTP is a relatively new rate structure in the electric power industry. Most of the existing programs are voluntary and in the industrial and commercial sectors.

(2) Priority Service

Priority service is a special kind of product differentiation that increases the range of choices available to consumers. The basic idea of priority service is to unbundle service reliability into a spectrum of priority classes, each priced to reflect the cost to the utility of providing that quality of service. Due to the extreme relevance and importance of priority service to this research, the remainder of this section is centered around this issue.

Priority service can be viewed as a special form of product differentiation in which the market is segmented into a spectrum of priority classes. Those customers willing to pay higher prices are assigned higher priority in receiving the product or service. The importance of this scheme is underscored by Milton Harris and Arthur Raviv (1981) who find that among all incentive mechanisms, priority service allows a monopoly to extract the highest profits from selling a scarce supply. In the more specific context of electric power, Maurice Marchand (1974), and John Tschirhart and Frank Jen (1979) consider a similar pricing scheme in which interruptible service is priced according to service reliability.

Product differentiation and priority service can be interpreted as a form of market organization that supplements and, in some cases, substitutes for spot markets. Spot prices are revised continually, whereas priority service contracts cover a period of extended duration. In principle, the price charged for each priority class is the expectation of what the spot prices would be for the same quality of service purchased in the spot markets. Priority service can be a less costly form of market organization if supplies are nonstorable, customer's valuations are stable over time, and transaction costs are significant [Chao and Wilson, 1987].

Compared with spot pricing, priority service offers two major advantages. One is that it yields important information about the distribution of customer's valuations that can be used to guide capacity planning. This information is unavailable from the observed choice behavior of customers in a spot market. The process of self selection among priority service contracts enables the seller to infer the allocation rule for every contingency. A second major advantage of priority service is that it enables supplemental insurance provisions to be incorporated into the contracts. When the role of customer's risk aversion is recognized, efficient risk sharing requires that any form of market organization be accompanied by insurance provisions. When the producer or a third party underwriter is the most efficient bearer of risk, the efficient insurance contracts cover all or most of the customer's risk. It is then in the underwriter's interest to allocate supplies to minimize claims.

Although spot pricing is used in wholesale markets for bulk trades among power producers, proposals to use spot pricing in retail markets have not been successful. A common explanation for this is that customers want prior assurance about the size of their monthly bills.

More fundamental reasons exist, however. First, is the argument of infeasibility. Electricity is essentially nonstorable, whereas failures of generation equipment can occur within timeframes of milliseconds. Spot prices therefore can fluctuate quickly and greatly. This requires allocation rules that can be implemented relatively quickly. Technologies that enable such quick responses by customers are not feasible at present. Moreover, there are no known methods for establishing new equilibrium prices without time-consuming iterations or collection of bids and offers.

A second argument, appropriate to intermediate timeframes, introduces transaction costs. Continually monitoring spot prices and adjusting demands responsively imposes appreciable costs on customers. Even if predetermined response rules are implemented automatically, a considerable investment in equipment is required. Priority service, however, takes advantage of the fact that the optimal rationing rule is based on priority assignments determined by customers' relative valuations of service. This makes automation of rationing rules feasible, and the centralization of the rationing rules economizes on the costs of implementation.

The implementation of priority service in the electric power industry has been proposed in several organizational forms. Each form has considerably different implications for costs and risks. The literature in this subject has studied some market forms in detail. In the following, three notable forms of implementation that differ in terms of contract forms and market organizations are discussed.

In the first form, consumers purchase multiple units rather than a single unit (say, one kilowatt of power). A consumer's demand comprises several units with different valuations. These units can be ordered by their valuations, from a base unit with a high valuation down to a marginal or peak unit with a lower valuation. Actually most of the priority service models are based mainly on this assumption. This form of implementation allows each consumer to select different reliability levels and corresponding rates for different increments of demand. In such a scheme, the responsibility for estimating the chances of interruption and for interpreting contractual obligations rests primarily with the utility. However, this approach poses practical difficulties. First, the utility usually has imperfect knowledge of the distribution of consumers' valuations. A misspecification of the price menu could result in too few customers selecting a low priority to enable provision of the higher level of service required for high-priority customers. Further, ambiguities in the interpretation and enforcement of such a contract may arise unless the contracts are specified in terms of observable events.

In the second form of implementation, consumers purchase service insurance and can expect to be compensated for an interruption by an amount that depends on the insurance premium paid in advance. During supply shortages, the utility first will interrupt the service of those consumers who select the lowest coverage. In such a scheme, the utility is committed to the priority ranking determined by the risk premium or interruption compensation stipulated unambiguously in the service contract, but not to a probability or frequency of service (although these may have been used to design the menu and inform customers about the predicted consequences of their selections). Michael Manove (1983) shows that, in general, if the insurance is provided by the producer, it will be free from the moral hazard problem, inducing both the producer and consumer to reduce efficiency losses. Therefore, this form of implementation requires relatively little monitoring and control. In the third implementation form, each consumer buys priority points, which are then assigned to demand segments. A market will be created to allow consumers to exchange their priority point holdings. The utility is relieved of the task of designing a price menu. In an emergency, the utility curtails those demand units assigned the fewest priority points. In this approach the burden of assessing the likelihood of interruption is transferred to the market maker and participants. The market transactions of priority points will provide relevant information about the distribution of consumer valuations and a direct indication to the utility of whether capacity expansion is justified.

However, an efficient implementation of this scheme requires that customers' expectation be "rational", in the sense that their selections are based on reliability assessments that are eventually consistent and correct.

II.5 Summary

The discussion here has revealed two different treatments of the issue of reliability differentiation. Traditional treatment forms, which mainly use random rationing schemes, overlooked the issue of customers preferences. However, these forms correctly recognized the limitations of the power system to provide for limited reliability options of service. Advanced forms, in contrast, permit each customer to opt for a specific reliability level for each increment of his/her demand. These forms yield price structures generally very complex to implement. The assumption is that the advanced electronic technology could provide services in a timely and cost-feasible manner. The need for a structure that does not compromise customers' choices, and at the same time is easy to implement is evident.

Chapter III

Literature Review

Two major areas of the economic literature central to this research are (1) peak-load pricing in public utilities, and (2) product-differentiation and discriminatory pricing. A review of both subjects is provided in this chapter. The product-differentiation subject is discussed in more detail. Developments in the public utility pricing literature are important, as they provide the foundation for this work and its benchmark results. The early contributions of economists in this area have been toward solving the deterministic peak-load pricing problem. In the uniform demand case, where capacity can be expanded continuously and infinitesimally, their work indicates that price should be set and expanded in such a way that price equals marginal cost. Since, in practice, capacity can only be expanded in lumpy increments and since that means large fluctuations in prices, Boiteux and others have advocated that prices be set at long-run marginal costs (LRMC) at all times. The main result for the two period case (that is, containing a peak and an off-peak) is that capacity and prices should be set in such a way that only peak consumers pay for capacity costs, whereas both peak and off-peak consumers pay for the energy charges or variable costs. Wenders (1976) shows that with the use of different types of technologies (that is, capital intensive units for base load and gas turbines for peak load) even off-peak consumers will be responsible for a certain amount of capacity cost. This literature has been reviewed in detail by Munasinghe (1979) and Crew and Kleindorfer (1979).

None of these studies, however, takes into account the uncertainty in supply or demand. Starting with Brown and Johnson (1969), several papers introducing demand uncertainty appeared in the literature. Brown and Johnson make the assumption that any excess demand not met can be rationed costlessly. Because of the unrealistic nature of the assumption, their optimal solution exhibits frequently occurring excess demand. Their work, however, helped to focus attention on this problem. Much of the further work on this subject essentially was to make improvements on this model and to reach more plausible conclusions. Visscher (1973), Meyer (1975), Carlton (1977), Crew and Kleindorfer (1978), Sherman and Visscher (1978) are some of them. Many of these papers used the notion of rationing cost, which is the cost of making sure that only the marginal consumers are cut off in the event of a supply outage. Other rationing schemes, like random rationing schemes, were also considered. Additionally, the notions of a reliability constraint in the peak-load problem [Meyer, 1975], and of optimizing the reliability level (Crew and Kleindorfer, 1978) were introduced. Meyer introduces the approach of chance--constraint on demand to obtain prices that reflect specified standards of system reliability. Crew and Kleindorfer also use the chance-constraint approach and the rationing cost approach to establish safety (reliability) margins. They argue that the approach becomes very complex under a diverse technology system such as an electric power system. However, in a less diverse technology system such as a gas distribution system, the chance-constrained approach was successfully used to account for demand uncertainties and specified reliability standards [Guldmann, 1981, 1983]. Two recent papers bring into consideration both demand and supply uncertainties. Lioukas (1983) shows that off-peak consumers should be charged with capacity costs according to the loss of load probability in any period. Chao (1983) brings into consideration a multiple generating technology environment and also considers different types of demand uncertainties.

An important issue, not considered in the above papers, is that consumers face different levels of shortage costs from the non-supply of electricity, and their willingness to pay for a certain level of reliability varies according to their vulnerability to shortage. For example, let U(q,R) be the consumer willingness to pay for the q^{th} unit of electricity at reliability R. Implicitly, it is assumed here that a demand exists as a function of reliability. The question, however, is how the reliability enters the demand function. This can be seen only by modelling the consumer's choice problem. This suggests the possibility, however, that consumers could be given a choice with respect to reliability of service by providing different options. This is known as interruptible supply pricing. Essentially, it deals with the rationing of supply shortfall, but by using predetermined contracts to interrupt consumers according to the priority they have chosen. Several papers have discussed the problem of interruptible supply pricing: Marchand (1974), Harris and Raviv

(1981), Tschirhart and Jen (1978), and Dansby (1979). Marchand (1973) was the first to derive welfare-maximizing prices in an interruptible-service framework. In his model, consumers are free to choose their probability of service, and prices are levied on both the maximum demand and the mean demand of each consumer. Dansby (1979) examined peak-load pricing with "ripple" control whereby consumers are limited to prespecified levels of service during instances of excess demand. Harris and Raviv (1981) take the pricing method as endogenous and show how selecting a method is related to the capacity constraint. They find that among all incentive mechanisms, priority service allows a monopoly to extract the highest profits from selling a scarce supply. In the more specific context of electric power, Tschirhart and Jen (1979) develop a model of a profit maximizing (instead of welfare maximizing) monopolist. They also consider a similar pricing scheme in which interruptible service is priced according to service reliability. While these papers make interesting contributions to the formulation of the problem, the crucial missing aspect is the link between consumer shortage cost and how it explicitly affects consumer choice.

The second major literature stream, product differentiation and discriminatory pricing, is more important and relevant with respect to this research. The idea of reliability-differentiated supply is in the spirit of non-linear pricing and quality differentiation. Goldman, Leland and Sibley (1977), Faulhaber and Panzar (1977), Oren, Smith and Wilson (1982), Chiang and Spatt (1982) and Mussa and Rosen (1978) have written some of the early papers in this area. The idea common to many of them is consumer self-selection from a set of non-linear outlay schedules or from a menu of alternative quantities with different qualities. The latter aspect of quality differentiation is more relevant here. Continuous reliability and pricing menus are developed by Oren, Smith and Wilson (1986), Chao and Wilson (1987), Wilson (1988) and Viswanathan and Tze (1989) to analyze the efficiency of certain market organizations and supplier investment levels in electric power. In general, these models consider a priority index based on a single attribute, for example, the customer's service value, which leads to one-dimensional menu choices and price functions. In Oren, Smith and Wilson (1985) a pricing structure with additional additive capacity-based components is developed. In Chao, Oren, Smith and Wilson (1986), a two-dimensional additive price function is developed for the attributes of reliability and load shapes. In a related paper, Oren (1988) considers the case of triangular shortfalls in which interrupted service is gradually resumed based on a second set of priorities. This leads to a very complicated analysis in which pricing functions exist only under certain conditions.

Chao and Wilson (1987) develop a priority service model based on a simple form of customers' preferences and supply technology. Their approach is considered the basis of most of the above mentioned papers. In this model, each unit of demand is associated with a valuation privately known to the customer. The statistical distribution of customers' valuations is known to the utility. The major problem of this approach, although theoretically sound, is that it yields an optimal price schedule that is only implicitly defined and whose implementability and practicality are questionable. On the production side, supply is obtained from several technologies with constant marginal costs. The capacity of each technology is assumed to be comprised of infinitesimal increments with independently and identically distributed failure probabilities.

Several problems are also associated with this approach. First, it does not relate reliability to system costs. System costs and reliabilities depend on each other inherently; that is one reason the product is differentiated. The second problem is that the estimation of the system's and units' production (energy) in this approach is not accurately provided. To provide for an accurate estimation, a stochastic production cost simulation model should be used. Such a model considers any form of demand distribution and all scenarios of units failures.

Chao and Wilson's model leads to the conclusion that efficient priority service improves the welfare of every customer, compared to uniform quality service at a single price, and enables the utility to meet the same revenue requirement. This conclusion was challenged in several other studies [Baron, 1981 and Rau and Hegazy, 1990]. Baron demonstrated that in a regulated framework attaining social welfare goals is limited unless the regulator has the same information as the firm and has the authority to directly regulate quality. Rau and Hegazy develop a model in which consumers have to share the burden of fixed costs in proportion to their taste for quality (reliability). A customer with a taste for higher quality has to pay a higher share. They conclude that improvement in welfare gain is uncertain.

In the following some of the above mentioned papers are discussed in more detail.

The paper by Chao et. al.[2] provides a general overview of priority service and illustrates how it can reduce utilities' risks through introducing greater product differentiation in electric power service. It illustrates the basic features of menu-based service offerings and demonstrates that all customers' costs can be reduced through an appropriately chosen priority service plan, when both interruption costs and service costs are considered. It illustrates that the menu of service offerings must be matched both to the supply characteristics of the particular utility and to the reliability preferences of the customers it serves. Thus, the collection of customer preference information is essential to designing the most effective priority service plans. Since menus permit customers to select their individually preferred service plans, it is not necessary for the utility to know the service valuations of individual customers. Instead, the authors argue, the utility requires only the distribution of values throughout the customer population.

Chao and Wilson[3] developed the fundamental theoretical properties of priority service from a basic model of customer preferences and supply technology. In their model, each unit of demand is associated with a valuation known privately to the customer. This point has been widely adopted in most other product differentiation models. It assumes a continuous relationship between consumption, valuation, and quality. It implies, also, that utilities should be able to offer a

[2] Chao, Oren, Smith, and Wilson: Managing risk by unbundling electric power service, Electric Power Research Institute, EPRI P-5350, (1987).

[3] Chao and Wilson: Pricing, Investment, and Market organization, Electric Power Research Institute, EPRI-P5350, (1987).

continuous stream of different qualities of production. This assumption is questionable, because in practice electric power systems can only provide a limited number of discrete service quality options.

In their model, the authors assume that the statistical distribution of customers' valuations, and the probability distribution of supply are known to the utility. Supply is obtained from several technologies with constant marginal costs. The capacity of each technology is assumed to be comprised of infinitesimal increments with independently and identically distributed failure probabilities. Due to the importance of this model, we will discuss it in more detail.

(a) Consumer Choice

Each customer is characterized by a single unit of demand and the associated marginal willingness-to-pay v, which takes a value in the interval from 0 to V. Aggregate demand is uncertain. The aggregate demand function is represented by D(.,w), and the inverse demand function, or the willingness-to-pay function, is represented by p(.,w), both contingent on the state of the world w, which is a real-valued random variable. It represents a set of disturbances such as changes in consumer tastes, firm technology, weather or market conditions, etc. The objective of each consumer is to choose from the menu M a priority option that maximizes his/her expected surplus, where M is a pricing scheme menu of options = {(p,s,r)}. For each option (p,s,r), p is the priority charge (payable in advance), s is the service charge (payable as service is provided) and r is the service reliability, which is the probability of receiving the service. Selection of an option by a consumer is equivalent to accepting a contingent forward contract for delivery in an event having the specified reliability. Therefore, for each v, the consumer's problem is to find

$$S(v)=\max[r(v-s)-P\,|\,(P,s,r)\in M]$$

(b) <u>The Cost Model</u>

The authors assume that there are n technologies with marginal capacity costs K_i, and marginal operating costs C_i, for i = 1,...,n, respectively. These technologies are ranked in ascending order by operating cost. For each technology, the total capacity consists of a continuum of homogeneous generation units, each of which is subject to random failures. Let X_i denote the installed capacity of technology i, and the random function $Y_i(X_i, w)$ denote the available capacity, whose realization is known to the producer. The unit availability factor is denoted by a constant (a_i). The authors assume here that the capacity increments, that is, $Y_i(\Delta x, w)$, are independent of each other and of all other random variables.

The system operation is based on the prespecified loading order of technologies from 1 to n. For a given installed capacity configuration (X_1,...,X_n), the total available capacity of technologies 1,...,i is denoted by

$$Z_i(w)=\sum_{j=1}^{i} Y_j(X_j,w),$$

and the short-run marginal cost function is given by

$$C(z,w)=C_i, \qquad if \qquad Z_{i-1}(w) < z < Z_i(w) \ ,$$

where z is the actual production level.

(c) Optimal Price Menu

The authors proceed in two steps to design a price menu that maximizes social welfare. First, they present the conditions for optimal allocation, assuming that w is fully observable. Then, assuming that only the utility knows the distribution of w, they assume a price menu and demonstrate that it will induce consumer choices consistent with the optimal allocation obtained with perfect information. Then, the problem is to find the intersection of the marginal willingness-to-pay function and the marginal cost function, as illustrated in figure [III.1]. Denote the instantaneous equilibrium (spot) price by $\bar{p}(w)$, [spot price associated with a given random outcome w]. Then the reliability of a type v consumer can be expressed as

$$R(v)=Pr[\bar{P}(w)\leq v],$$

indicating that the consumer is served in the event that the spot price is less than his willingness to pay. Formally, the spot price is characterized as follows:

$$\bar{P}(w)=\min[\max[P(z,w),C(z,w)]\,|\,z\geq 0]$$

$$=\min[\max[P(Z_i(w),w),c_i]\,|\,i=1,\ldots,n].$$

And the authors construct the price menu M^{*} as follows:

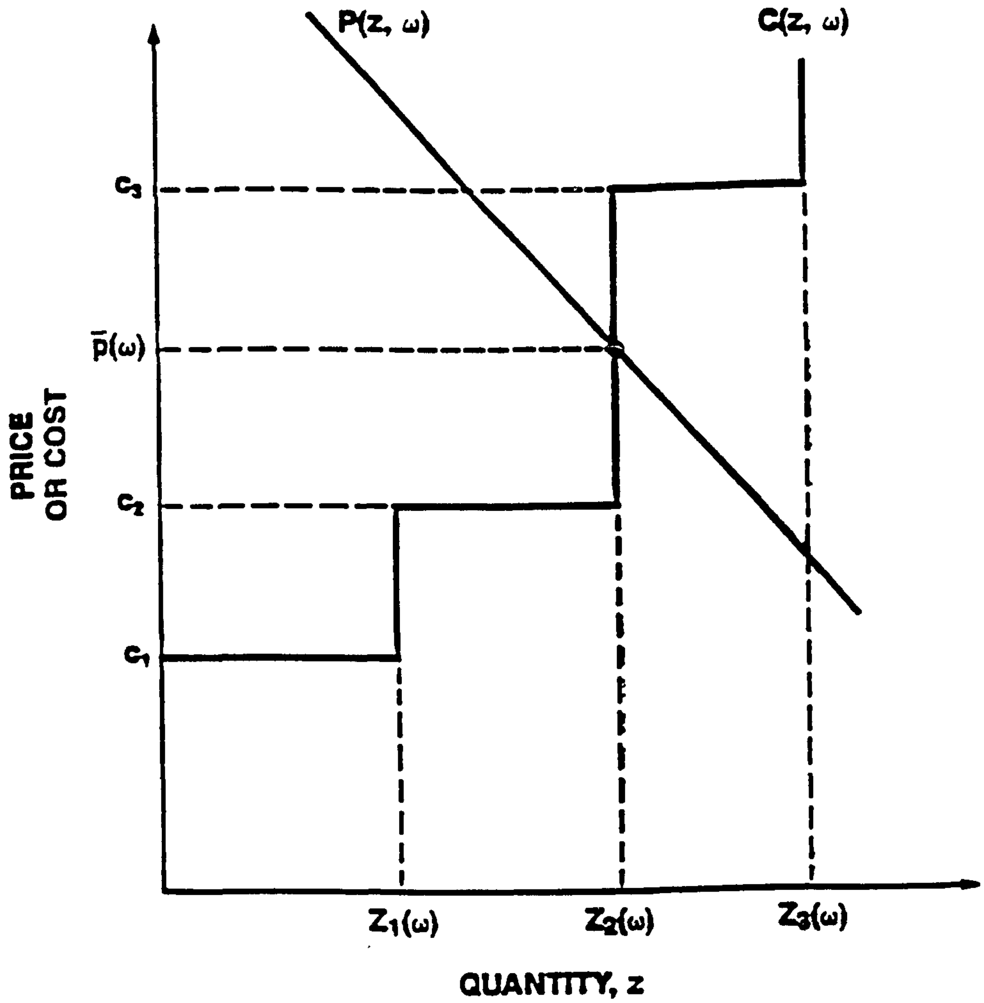


Figure III.1: Determination of spot price.

$$r^*(v) = \begin{cases} R(v), & \textit{if} \quad v \geq c_i, \\ 0, & \textit{if} \quad v < c_i. \end{cases}$$

$$s^*(v) = \begin{cases} \overline{c}_i, & \textit{if} \quad c_i \leq v < c_{i+1}, \\ 0, & \textit{if} \quad v < c_i \end{cases}$$

$$P^*(v) = \int_0^v [r^*(v) - r^*(u)]du - s^*(v)r^*(v).$$

$$M^* = [(P^*(v),\ s^*(v),\ r^*(v)) |\ 0 \leq v \leq V]$$

The authors argue that, since R(v) is nondecreasing in v, M^* will induce optimal consumer choices. This menu can be implemented in many forms. Several other papers have developed a menu based on the same concepts and assumptions, differing only in their implementation framework.

The above model concludes that the prices are to be the expectation of what would be paid for equivalent quality service if purchased in a spot market. Priority service and spot markets alike yield the same fully efficient allocation of scarce supplies. However, the utility's revenues are the same under the two forms of market organization only if demand is certain or the marginal component of demand is stochastically independent of total demand and supply.

Oren, Smith and Wilson[4] introduce a paper based on the doctoral research of Richard Pitbladdo, who studied the situation in which customers obtain further information about their valuations of service after they select their service priority assignments. His main finding is that an optimal rationing is a variant of priority service having the following features: each priority class has an assigned energy charge per unit of service actually delivered and higher priorities have higher energy charges. Due to this energy charge, a customer is predicted to demand service only if his subsequent valuation exceeds the energy charge assigned to his selected priority class. The utility's rationing rule is to serve customers in order of their selected priority classes, but only if they demand service.

Wilson's[5] paper studies the design of efficient rationing schemes and their implementation in both regulated and competitive environments. The general rationing problem is to elicit customers' privately known preferences via self-selection from a menu of options, and then to use this information to guide the allocation of scarce supplies. An efficient scheme is one that maximizes expected total surplus. Contrary to the previous papers in which priority service has been shown to be an efficient method for a restricted class of customers' preferences and supply technologies, this paper provides more general models of customers' preferences and

---

[4] Oren, Smith, and Wilson: Pricing Priority Service: Further Characterization of service menus, Electric Power Research Institute, EPRI-P5350, (1987).

[5] Wilson, R B : Efficient and Competitive rationing Via priority Service, Electric Power Research Institute, EPRI-P5350, (1987).

supply technology.

The main conclusions derived from this study can be summarized as follows. First, for a monopoly, priority service is an optimal rationing scheme, and for a regulated monopoly priority service is both efficient and Pareto superior to random rationing. A few priority classes suffice to realize most of the efficiency gains. Uncertain customer valuations can be addressed via priority-dependent energy charges, and sufficient information is revealed to guide capacity planning. Prices are the expected spot prices for equivalent service quality. In competitive markets, however, firms lack a profit incentive to differentiate their service classes, and generally the efficiency gains from better rationing of scarce supplies are obtained only from entry of additional firms. However, if firms' supplies are imperfectly correlated and not completely pooled, then the diseconomies of imperfect pooling can more than offset the potential gains from more efficient rationing. In this case, an optimal form of organization for the industry is a single regulated monopoly offering differentiated classes of priority service.

A paper by S. Smith[6] describes a series of linear programming models for designing customer menus of price/reliability tradeoffs in electric power service. Total welfare maximization achieves a socially efficient distribution of demand, subject to existing generation capacity, and determines the marginal economic value of additional capacity. Marginal-cost prices for each service option are obtained from

[6] Smith, S.: A Linear Programming Model for Pricing Alternative service Conditions for Electric Power, EPRI-P5350, (1987).

the dual variables of the linear program.

The model's inputs are the marginal costs, capacities, and reliability characteristics of the utility's current generating units, and the customers' values of service, aggregated by applications or end-uses. The outputs are the demands and prices for the various service reliability options, the maximal operating levels for the utility's generation units or blocks, and the shadow prices for additional generation capacity of the various types. The model is formulated so as that to select the values of the decision variables (U,x,Q,y):

Maximize $\Sigma_i\Sigma_a\ h_a x_{ai}\ V_{ai} - \Sigma_j\ f_j U_j - \Sigma_j\Sigma_s\ e_j Q_{js} t_s$

subject to:

$\Sigma_i\ x_{ai} \ \leq\ 1$ for all a ,

$\Sigma_s\ t_s y_{is} \ \geq\ r_i t,$ for all 1,

$\Sigma_a\ h_{as}\ \Sigma_{m<i}\ x_{am} - \Sigma_j\ Q_{js} \ \leq\ B(1-y_{is}),$ for all i,s,

$Q_{js} \ \leq\ L_{js}\ U_j,$ for all j,s,

$A\ U \ \leq\ M$ [or $U_j < M_j$]

$U_j,\ Q_{js},\ x_{ai} \ \geq\ 0,\ \ y_{is} = 0,1.$

Where:

h_a = total potential demand (kwh) of type (a) during the contract period

= $\Sigma_s\ t_s h_{as}$

h_{as} = number of demand units (kw) of type (a) under scenario (s).

x_{ai} = fraction of demand units of type (a) assigned to product i

V_{ai} = net value of serving 1 kwh of application (a) with product i

U_j = maximum operating level for block j for the contract period.

M_j = maximum available capacity.

$L_{js}\ U_j$ = the maximum available capacity from the j^{th} unit under scenario s, given that the capacity level U_j was selected.

Q_{js} = actual operating level of the j^{th} unit under scenario s (where: $0 \leq Q_{js} \leq L_{js}\ U_j$),

t_s = the number of hours that the system operates under scenario s during the contract period t. [Thus $\Sigma_s\ t_s = t$.]

f_j = marginal capacity cost (per KW) for unit j for the contract period.

e_j = marginal energy cost (per KWH) for unit j for the contract period.

r_i = assured reliability level of product i.

d_{is} = demand for reliability r_i or better under scenario s =

$= \Sigma_a \Sigma_{m<i}\ x_{am}\ h_{as}$,

y_{is} = 1 if product i is to be served under scenario s, and 0 otherwise.

B = a "very large" constant.

The linear program maximizes total welfare (value of service minus service cost), subject to the generation capacity constraints, assuming the optimal product/application assignments are made. The first constraint indicates that applications may be assigned to one product at most, but need not be assigned at all.

The second constraint guarantees that when averaged over all scenarios, each service option delivers at least its promised reliability level as measured by $r_i t$, the amount of time that service is provided. The third constraint states that if $y_{is} = 1$, that is, that product i is to be served under scenario s, then supply must be greater than or equal to demand under scenario s. If $y_{is} = 0$, the large constant B effectively removes that constraint. The fourth constraint enforces the upper bound on available capacity. Finally, $AU < M$ characterizes the feasible set for selected capacity levels U_j.

The feasibility of this approach is hindered by the assumption that a customer can assign numerous increments of his load to different service options. The utility and its power system under this scenario are assumed to be able to provide all these services. While the idea of allowing the customer to select different values for each increment of demand is the essence of all these studies, it yields price structures and menus that are extremely difficult to implement and for customers to understand.

Oren, Smith and Wilson[7] introduce another model where supply and demand are no longer linear functions of the system's parameters, but are modeled as general functions. The utility selects nominal operating levels for each of its generation sources at the beginning of the contract period. Each unit provides a randomly varying output that depends upon the chosen operating level. Customer demand also varies randomly, but is dependent upon the price and reliability of each offered service grade. The reliabilities of the service grades are not pre-determined, but are specified by the utility, based on its estimates of customer demand and the supply

[7] Oren, Smith, and Wilson: Discrete models for Pricing Reliability Options.

costs and capacities of its generation resources. Once the utility specifies the reliability of each service grade, it is committed to deliver service of that quality or better during the contract period.

The formulation of this approach was presented in a mathematical programming algorithm similar to the above formulation, although more general.

In an environment where the utility is assumed to be risk neutral and consumers risk averse, an allocation mechanism introduced by Oren and Doucet[8] offers consumers the opportunity to purchase some type of compensatory insurance in order to recoup losses incurred in the event of an interruption. While the reliability distribution for each consumer is assumed to be public information, each consumer's valuation of the uninterrupted power is private information (that is, known only to the consumer). Because of this asymmetry of information the utility is unable to identify the particular type of a given consumer and must rely on self-selection. In particular the tariff should offer incentives so that each consumer makes a compensation choice which truthfully reveals his/her valuation. This permits the utility to ration electricity efficiently by way of minimizing compensation payments, that is, interrupting consumers in the order of their selected compensation levels. Furthermore, assuming that the utility is risk-neutral whereas consumers are risk averse, a socially efficient tariff should induce transfer of all the risk to the utility through full insurance.

---

[8] Oren, S., and J. Doucet: Interruption Insurance for generation and Distribution of Electric Power.

The author's proposed tariff structure consists of a service charge in excess of the supply cost (paid only when service is delivered) and an insurance premium depending on both the probability distribution of consumer type and selected compensation.

Summary

In the reviewed literature, several assumptions have been made that we found to be critical to the validity and implementability of the reliability-based pricing concept. This research is prompted by the obvious need to reconsider all these assumptions in order to have a reliability-based price structure with conclusive results. These critical assumptions are listed as follows:

1. It was assumed in most of the previous work that each customer could select the appropriate service options for each increment of his load. The implicit assumption is that the reliability of a power system changes smoothly with any incremental change in production. Since in practice capacity can be added only in lumpy increments, reliability also changes in the same manner. Therefore, we assume that the customer should opt for different reliability level for each block of consumption, instead of opting for each increment.

2. It was assumed in all of the literature that each customer knows fully his valuation of service at the time he chooses among the menu of options offered and also that the menu of options is fixed once it is

offered. The proposed model constructs a more realistic case, where the customer makes his selection based on an initial estimate of his valuation of service, whereas at the time service is offered the customer knows exactly his valuation of service. Therefore, the proposed model will allow customer and utility alike to continue adjusting their valuation for quality via exchanging information until both valuations are matched at the time of actual service. An equilibrium-iterative process is proposed to handle that.

3. The relation between system reliability and system costs has not been developed. For example, the concept behind peak-load pricing was that peak consumption induces more costs for the system than off-peak consumption. If the same concept is applied to reliability-based pricing, as the literature correctly claims, then the link between cost and reliability of production should have been developed.

4. Previous product quality analyses have treated supply cost as predetermined and have not considered the impact of the supplier's decisions on the probability distribution of available supply. The proposed model uses stochastic production cost simulation to estimate expected production, cost, and reliability while taking into consideration the stochastic nature of the system.

5. One of the basic notions behind reliability differentiation is the fact that customers differ in the shortage costs they sustain during an

interruption and hence differ in their willingness to pay to avoid it. However, outage costs have not been considered in any of the above studies.

6. Rate structures should be implementable and easy to understand. The price menus that are suggested in the literature are not attractive in both regards.

Chapter IV

The Model

The proposed model is based upon the concepts employed in the nonlinear pricing literature. We assume that a single welfare-maximizing public utility owns and operates the generating plants and transmission network of the electric power system under consideration and sells to independent customers. The utility is assumed to be able to instantaneously set and communicate prices, with a different price for each customer class and reliability level. Supply outages are assumed to occur randomly, with the probability of an outage being known. Although only generation outages are considered in this study, the analysis can be extended to include other causes of outages. While both electric supply and demand vary throughout the day and throughout the year, this research focuses on one time cycle, the day, and divides it into 24 periods. We assume that electric power supply and demand in one period are independent of supply and demand in any other period.

The model employs the traditional measure of welfare used to evaluate public utility pricing policies [Crew and Kleindorfer, 1987], with:

$$W = CS + PS,$$

where W = net social welfare, CS = consumers' surplus and PS = producer's surplus. The consumers' surplus (CS) is the area under the marshallian demand curve, i.e

$$CS(x) = \int_0^x [p(y) - p(x)]\, dy.$$

where P(X) is the inverse demand function. It is also called the "willingness-to-pay function" or "marginal utility function". The consumer surplus represents the difference between what the consumer is willing to pay for a given level of consumption, and what he actually pays (p). The latter price p is determined by market forces or a regulatory body. The producer's surplus is measured by its profits, i.e., the difference between total revenue and total cost:

$$PS(X) = P(X)\, X - TC(X) - FC$$

where TC(X) is the total variable production cost, and FC is the total fixed cost. The net benefits are then measured by:

$$W = CS + PS = \int_0^x p(y)dy - TC(X) - FC.$$

In the model, customers are aggregated into four independent classes (residential, commercial, and small and large industrial). Customers in each class are assumed to have the same valuation of electricity and reliability. Residential customers are assumed to be price-taking, utility-maximizing households. Large and small industrial, and commercial customers consist of price-taking, profit maximizing firms. Further discussions of customers' consumption behavior and reliability choices are outlined in section IV.1

To allow for service quality (reliability) differentiation among customers, we rank customers into two groups: a high reliability group and a low reliability group. High reliability customer classes are assumed to opt for "firm" power service, that is, a 100 percent continuous service. The customer classes in the low reliability group are ranked in the order of the interruption level they would opt for. At any outage situation, the utility will interrupt low reliability customer classes in accordance with their chosen rank slot. A price-incentive mechanism to encourage customer classes to make rational choices is proposed in section IV.3.

Because supply uncertainty is an important element of this research, a stochastic production cost simulation is used to represent the supply component of the model. The use of stochastic production cost simulation allows accounting for all possible forms of randomness in both supply and demand. This simulation technique is designed to relate demands to production costs, marginal costs, and reliability levels. The technique also provides a new way to estimate the expected value of the marginal costs.

In the following section we discuss in detail the elements of the proposed model.

IV.1 Customer Choice Modeling

Traditional consumer theory ignores the notion of quality to focus on that of quantity. Both production and utility functions are defined as functions of the quantities of the various goods. At the other extreme, the Lancasterian approach ignores the notion of a good to focus on the notion of characteristics. Neither approach can help analyze consumer behavior with regard to both consumption and quality simultaneously. In this research, our interest is to express customer classes's choices for electricity service with respect to both the quantity and the quality (reliability) of service.

The main proposition is that customers can often adjust their activities (that are related to the quantity and quality of electricity consumption) faster than product characteristics can. This proposition is obvious in the case of electricity service. In the short run, customers have to constrain their preferences to comply with the limitation of the power system. In the long run, however, power systems should be able to adjust to satisfy all choices. Customers, therefore, cannot associate each increment of their consumption with a different level of quality in the short run (as was suggested in the literature), but they can divide their consumption into a few limited blocks and opt for a different level of quality for each. This research devotes its attention to the short-term analysis, in which reliability choices for customer classes are bounded by the limitation of system reliability.

In the following we discuss the proposed behavior of each customer class. The residential class is characterized by the behavior of a household, and the three other classes by the behavior of a firm.

The Household

Consumer demand theory tells us that the demand for a commodity such as electricity is determined by the maximization of the consumer's utility function subject to his income constraint. If income is assumed large enough, and the prices of all other commodities except that of electricity are presumed to be constant, then one obtains the price-demand function for electricity as the marginal utility with respect to the demand of electricity. Obtaining price-demand function is essential in order to measure consumer surplus. However, consumer surplus does not always yields a precise measure of welfare. As price changes while the consumer's money income remains constant, real income may itself affect the consumption of the good and shift the demand curve. Unless it is assumed that income effects are negligible and that the income elasticity of demand for the good is zero, the consumer surplus cannot be assumed to be an accurate measure of welfare. In most practical situations, a small fraction of income is spent of the good in question, and estimation of consumer surplus from observable demand curves will yield an adequate measure of consumer welfare.

In order to characterize customer behavior in this class, let us first formulate its utility maximization problem. The customer is supposed to maximize $U = f(x_1,..,x_n)$ subject to the linear income constraint $\Sigma_i\ p_i x_i = y$, where p_i represents the price of the i^{th} commodity, x_i the demand for the commodity, and y his total expenditures "income". The first order condition of this problem is:

$$\frac{\partial U(x^*)}{\partial x_i} = \lambda\ p_i, \qquad i=1,...,n$$

providing the inverse demand function:

$$p_i = \frac{1}{\lambda}\frac{\partial U}{\partial x_i},$$

where λ is the Lagrange multiplier, also known, in this context, as the marginal utility of income. As mentioned above, if the income level is sufficiently large, one can assume that the marginal utility of income λ is equal to one. Economic theory shows that the "no-income-effect" assumption results in two simplifications. First, it is possible to measure the value of good x to each individual by the area under the individual's ordinary demand curve for x (otherwise it might be necessary to use the appropriate individual compensated demand function). Second, all the information needed for the analysis can be obtained from aggregate or market demand functions.

Since consumer's behavior is traditionally summarized by means of a utility function, we should therefore display the desirable form and analytical properties of the utility function that fits the purpose of this research. It should show a diminishing marginal utility for any one input alone, and be capable of yielding negative price elasticities of any finite absolute value. Under the no-income-effect assumption (the demand for x is independent of budget level), the consumer's preference can be represented by a an additive utility function in the level of the composite (numeraire) good (all other goods). A quadratic separable utility function form can satisfy such requirements. This form has consistently a wide range of price elasticities, is concave, is twice differentiable, and yields downward slope demand functions that fit the purpose of this research.

One way to represent the impact of service reliability on customer satisfaction is to assume that unreliable service directly reduces utility. That is, if R is the probability that the amount X of electricity will be served, the utility is $U = U(X,R,..)$, with $\partial U/\partial R > 0$. The rationale that reliability is a desirable attribute of electricity is just the same as the rationale that efficient gasoline burning is a desirable attribute of a car. Consequently, greater reliability would affect the utility obtained from any given consumption of electricity.

To consider explicitly a reliability index as a parameter in the utility function is to assume that reliability is observable in advance. In the economic literature, there are three types of goods with regard to quality: search goods, experience goods, and credence goods. Search goods are goods where the quality can be ascertained by consumers before a purchase (for example, dresses). Experience goods are goods where quality is learned after the good is bought (for example, the taste of canned food or the quality of a restaurant). Credence goods are goods where aspects of quality are rarely learned, even after consumption (for example, the amount of fluoride in a toothpaste).

Electricity is an experience good with regard to reliability of service. However, it is not always the case that reliability is learned after electricity is bought. Learning in this case is a stochastic process because service may or may not be interrupted. For such an experience good, the main issue is information. The probability of interruption is the most important piece of information that a consumer needs in advance, and is the most developed index a power company can provide. The consumer's utility function, therefore, can be expressed in the following form:

$$U(X,R) = R\ U_1(X,..) + (1-R)\ U_2(X,..).$$

The first part, U_1, represents the satisfaction of the consumer in consuming the amount X of electricity with reliability R. The second part is the loss or dissatisfaction caused by an interruption of X with a probability (1-R). Both parts are functions of the consumption variable X and the parameter R, noticing that X is always positive and larger than zero.

Following the above discussion, and assuming a quadratic, additive (quasi-linear with respect to the composite good) utility function form with no income effect, we propose the following form for the residential class:

$$U_t = R_t(a_t X_t - \frac{b_t X_t^2}{2}) - (1-R_t) X_t OC_t + m, \qquad t=1,..,24$$

where:

U_t = the utility function of the residential class at period (hour) t,

X_t = residential class demand for electricity (MW) at period t,

R_t = the reliability level (probability that the load will be served) at period t,

OC_t = the customer class outage cost at period t,

a_t ,b_t = the parameters of the utility functions,

m = the composite (numeraire) good.

This utility function form leads to the following inverse demand function for the residential class:

$$P_t(X,R) = R_t(a_t - b_t x_t) - (1-R_t).OC_t, \qquad t=1,..,24$$

The Firm

We consider a short-term-rational competitive firm whose benefit from electricity usage at any time depends on consumption at that time only, which means that the firm maximizes its profits at each time t independently. Consider a firm producing, at time period t, an output q_t, sold at price r_t, with a set of non-electricity inputs denoted by the vector L_t and whose prices are w_t; and an electricity input X_t whose price is P_t. The firm's production function is written as $q_t = h_t(X,L)$. According to economic theory, the firm is supposed to maximize its economic profits, given the factor and output prices that it faces in the market, and given its production technology. The representative profit-maximizing firm will seek to produce its output by methods that minimize its cost of production. That is, the profit maximization problem is analytically separable into two problems. First, find the cost minimizing combination of non-electricity inputs L for producing any given output level using a fixed amount of electricity. This gives rise to the input demands $L_t^*(q,w,X)$ for the non-electricity inputs, and hence to the cost function $C_t^*(q,w,X)$, representing the cost of purchasing those non-electricity inputs. Second, produce the output level and select the amount of electricity that maximize profits:

$$Max.\Pi_t = r_t h_t(X,L^*) - C_t^*(q,w,X) - P_t X_t.$$

Let V_t be the value-added or benefit function for the firm use of electricity at time period t. Thus, V_t is the firm's revenue minus the cost of all non-electricity variable inputs:

$$V_t(X,w,r) = r_t h_t(X,L^*) - C_t^*(q,w,X).$$

Assuming that the firm is facing a competitive market for its output and all non-electricity inputs, then V_t can be expressed as a function of electricity inputs only[9]. Then maximizing profits is equivalent to:

$$Maximize\ \ \Pi(x_t) = V_t(X) - P_t X_t,$$

which gives the identity $\partial V_t/\partial X_t = P_t$, where P_t is the price-demand function of electricity at time period t.

In order to estimate the consumer surplus of a firm consuming electricity, the demand function for the good produced by the firm must be estimated. As a practical matter, it may be difficult to obtain the necessary data for estimating the demand for the output of industrial or commercial firms. Fortunately, Schmalensee [1976] has demonstrated that valuations of consumer surplus are identical from either input or output markets, provided the output market is competitive. Under the assumption that the framework of reliability impact on consumers holds for firms, and that $V_t(X_t)$

---

[9] The assumption that the consumer firm faces competitive markets in its output non-electricity inputs, means that the prices for its output and these inputs are fixed equal to their private (market) marginal costs.

has the same quadratic form as assumed for the residential class utility function, one obtains the following linear inverse-demand function for all the other three classes:

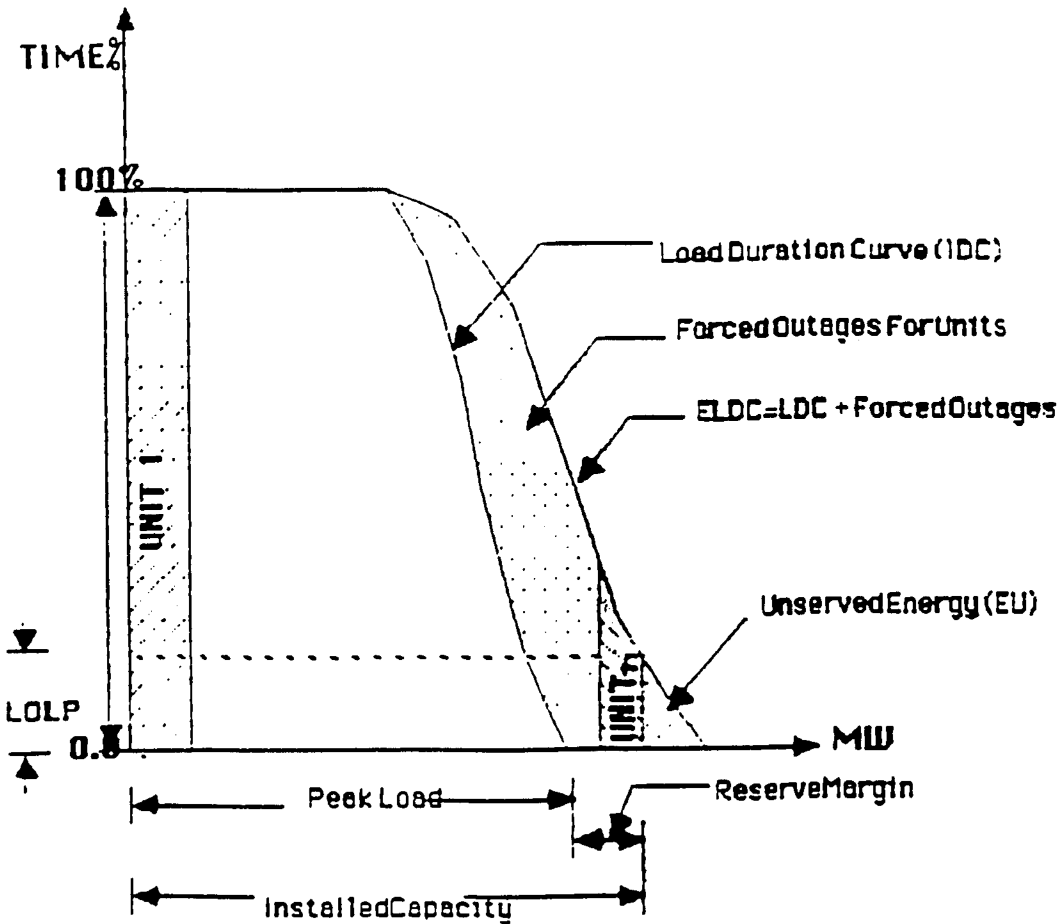
$$P_{it}(x) = R_{it}(a_{it} - b_{it}x_{it}) - (1 - R_{it}).OC_{it}, \qquad t = 1,...,24,$$

where i refers to the i^{th} customer class, and R is the reliability level.

IV.2 Supply Model

Production cost simulation takes as given a set of randomly available resources and a time-varying load. The simulation determines how much energy is served with each resource. A resource is any method that can serve load. Conventional electric generating units such as coal-fired plants are resources, but a less obvious example of resources are load management technologies such as devices for terminating electric service to air conditioners. These resources experience random outage periods when they are unavailable to serve load. The inability of a plant to serve load is called an outage. The load is the time series demand for electric power placed on the electric system by the customers.

Finding a computationally efficient method for explicitly representing the interaction between the randomly available resources and the time-varying load proved difficult until the concept of equivalent load was put forth by Baleriaux, Jamoulle, and Guertechin [1967]. Load variations over time are represented by the load duration curve (LDC) in figure IV.1. The graph shows the percent of time that the load is greater than or equal to the load shown on the horizontal axis. A complete review of the literature on production cost simulation is provided in



Unit 1= Cheapest unit (the least o&m cost)
Unit n= The highest o&m unit
LOLP= Loss of load probability

Fig IV.1: Reliability Indices (LOLP, EU)

appendix A.

The proposed stochastic production cost simulation technique is grounded heavily in the work of N.S. Rau [1979,1982]. His method is extended to include the estimation of the expected marginal cost (the probability distribution of system incremental cost) at each level of demand. The concept is illustrated in figure IV.2, where the probability density function (PDF) of load $f(L_o)$ is shown in the top right hand side. The probability of outage of a generation unit f(C) is represented in a binary fashion to the left of the vertical line. Note that q_i is the forced outage rate (FOR) of the i^{th} unit and p_i is its availability. A capacity of [-C] is shown to be available with probability p to reduce the demand. Hence the negative sign for C. To start with, if no units are available, the PDF of load is the unsatisfied demand (residual demand), and the zeroth moment of the PDF is the probability the demand is unserved (LOLP equals 1.00 in this case).

If only unit #1 were to be considered, the convolution of its binary representation $f(C_1)$ with load will result in the density shown in figure IV.2b. Observe that the density, compared to the PDF of load, is shifted to the left by C_1. This is due to the negative sign of capacity in its binary representation, signifying the fact that a maximum of C_1 can be reduced from the demand (or residual demand) by the operation of unit #1.

In figure IV.2b the density to the right of the vertical line illustrates the PDF of residual or unmet demand (RD). Similarly, the density shown in figure IV.2c is obtained by considering the second generation unit in the merit order of loading.

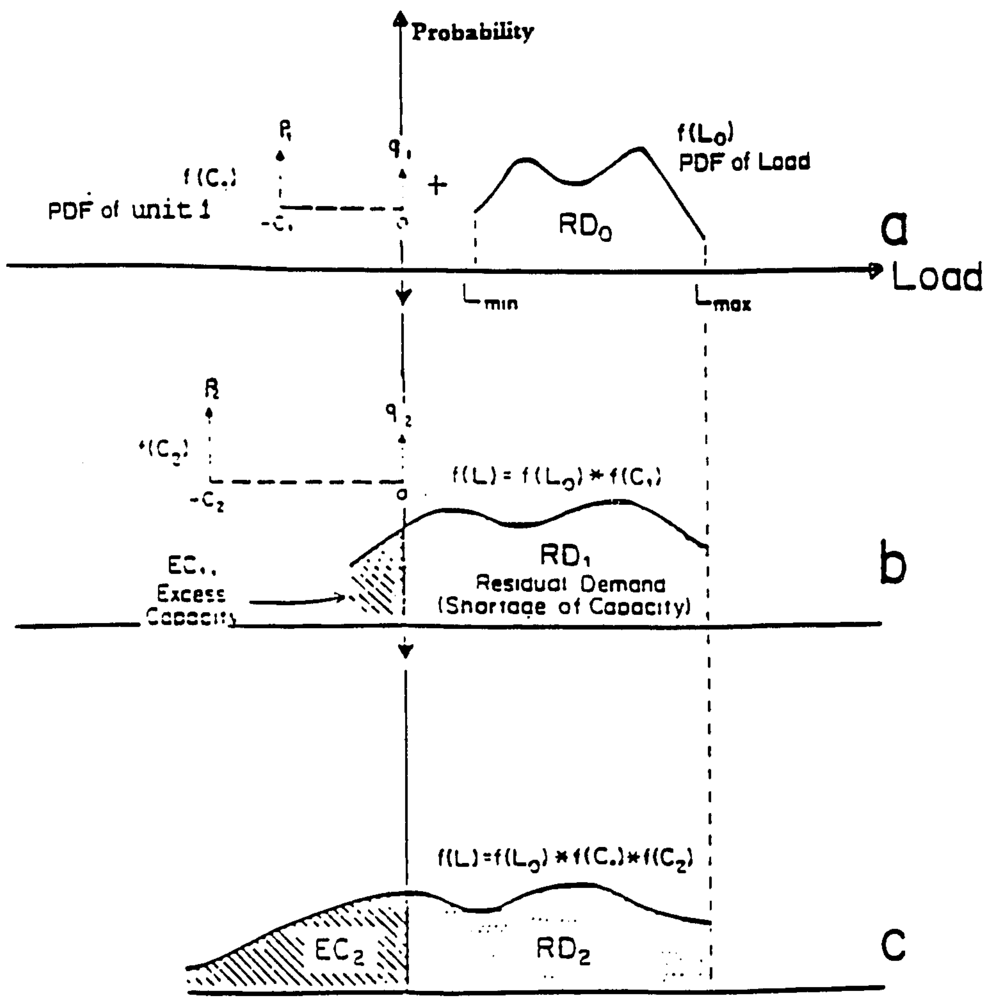


Figure IV.2: Convolution of residual demand and unit (-ve) availability.

It has been shown [Rau, 1979], that the difference between the means of any two residual demands is the mean or the expectation of energy production from the concerned unit. For example, the difference between the means of the densities to the right of the vertical line in b and a of figure IV.2 is the mean energy generated by unit #1, and the difference between c and b represents the mean energy generated by unit #2. A repetition of this procedure gives the expected production from each unit. The expected production cost can be obtained by multiplying the expected production from each unit by its cost of generation. Appendix A illustrate the steps of this procedure in detail.

We extend this analysis to obtain the expected marginal cost of the system by estimating the probability distribution of system lambda (incremental cost). Noting that the zeroth moment of the density to the right of the vertical line in (a) of figure IV.2 is the probability that the demand is unserved before committing any unit, and the zeroth moment of the density to the right of the vertical line in (b) of figure IV.2 is the probability that the demand is unserved after committing unit #1; the higher the value of this probability, the higher the probability that the system needs more units to be committed in order to serve the residual demand. Therefore, we argue that the difference between the zeroth moment of the densities to the right of the vertical line in (b) and (a) of figure IV.2 is the probability that unit #1 was needed to cover the total residual demand RD_0, or the probability that unit #1 is the marginal unit. If the residual demand is larger than the available capacity of the upcoming unit, then the unit is not the marginal one (and the probability is zero). On the other hand, if the unit available capacity is greater than or equal to the residual demand, the probability (the difference between two consecutive densities) is larger than zero and represents the marginality of the unit, and the probability that the incremental cost of this particular unit is the system incremental cost. The list of these probabilities represents the probability distribution of the system marginal cost. The summation of the product of each unit incremental cost by the probability that the unit is the marginal unit gives an accurate estimation of the expected value of the system marginal cost.

To further illustrate the above concept, let us consider the graphs in figure IV.3. Let f(L), shown in IV.3a, illustrate the PDF of the residual demand at any stage of the convolution process. Let the density shown in IV.3c represent the available capacity. This density may represent the total available capacity or the capacity available from any particular unit/units in a discrete form.

Consider an incremental load ΔL at a load level L. The probability that the load is between L and L+ΔL is

$$\int_{L}^{L+\Delta L} f(l)dl,$$

and is the hatched area in figure IV.3b. The probability that the energy produced is between L and L+ΔL is the product of the probability that the load is in this area and the probability that the capacity available is greater than L+ΔL. We shall

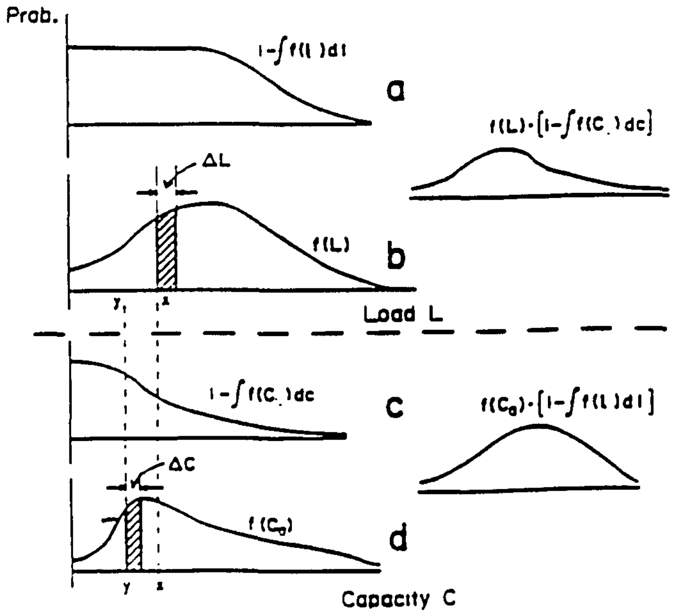


Figure IV.3: PDF of energy served by units consider the case of available capacity being less than L later. Expressed mathematically, the energy density function for capacity available between $C+\Delta C$ is:

$$\int_{L}^{L+dL} f(l)dl \cdot \int_{C}^{C+dC} f(c)dc, \qquad C>L$$

Similarly, the energy produced will be between L and $L+\Delta L$ for other values of C higher than $L+\Delta L$. By considering all values of $C > L$ and all values of the load, we obtain the total energy density function as:

$$f(L).\int_{L+dL}^{\infty} f(c)dc$$

Now, consider the case when available capacity is less than the load. The energy served will then be equal to the capacity available irrespective of the demand. It is then evident that the probability of the energy produced is equal to the PDF of capacity availability times the probability that the load exceeds the available capacity. Hence, one gets, for the second component of the PDF of energy production:

$$f(C).\int_{C+dC}^{\infty} f(l)dl$$

The total PDF of energy production is therefore:

$$= f(L) \int_{L+dL}^{\infty} f(c)dc + f(C) \int_{C+dC}^{\infty} f(l)dl$$

$$= f(L)\ [1-F(C)] + f(C)\ [1-F(L)]$$

where F(C) is defined as ∫f(c)dc, and F(L) is ∫f(l)dl. The above equation represents the PDF of the energy produced by a unit with available capacity f(C) meeting a residual demand f(L). Rearranging the equation, one gets:

$$f(L) - [f(L)F(C) + f(C) \int f(l)dl]$$

From the definition of the convolution process [Appendix A, equation A.18], the above formula is nothing but the difference between the PDF of the residual demand before committing unit C and the PDF of the residual demand after committing unit C.

In sum, this model provides estimations for the expected total production cost (TC), expected marginal cost (MC), and the system reliability level (1-LOLP) at each level of consumption and period of time. As shown in appendix A, there is no closed form function that can depict the relationship between total demand and these estimated parameters. However, in the following we express these parameters in the form used in the simulation:

$LOLP_i$ = the loss of load probability after committing unit i

= the zeroth moment of $[RD_{i-1} - RD_i]$, i is the last unit committed,

LOLP(X) = the system loss of load probability at consumption X.

= the zeroth moment of RD_n, n is the last unit available in the system,

TC(X) = total production cost at demand X,

= $\Sigma_i^n IC_i$.(the first moment of $[RD_{i-1} - RD_i]$),

MC(X) = the expected marginal cost,

= $\Sigma_i^n IC_i.[LOLP_{i-1} - LOLP_i]$,

where:

RD_i = the residual demand after committing i generating units (figure IV.2).

IC_i = the fuel cost of unit i ($/Mwh),

Appendix C defines some of the other acronyms used in this chapter. The next section introduces the proposed integrated welfare model, where the production cost simulation model is used to provide estimates for TC, MC, and LOLP as functions of total system demand.

IV.3 The Welfare Model

As discussed earlier, the electric utility is assumed to be welfare maximizing, and using as criterion the sum of consumers' and producer's surpluses, with:

Maximize Welfare (W) = Consumers' Surplus (CS) + Producer' surplus (PS),

or

$$Max\ W_{(x_{it})} = \sum_{i \in I} \int_0^{x_{it}} P_{it}(y_{it},R_t)\ dy\ -\ TC_t(x)\ -\ FC, \qquad t=1,..,24$$

where:

$$CS(X,R) = \sum_i \ [\int_0^{x_{it}} P_{it}(y_{it},R_t)dy\ -\ P_{it}(x_{it},R_t)x_{it}]$$

$$PS(X,R) = \sum_i P_{it}(x_{it},R_t)x_{it}\ -\ TC_t(X)\ -FC$$

$P_{it}(X,R)$ is the inverse demand function of customer class i at time t, $TC_t(X)$ is the utility total variable production cost, FC is the utility total fixed cost (assuming system capacity and generation mix are fixed), R_t is the system reliability level available at period t at total system demand, and x_{it} is the electricity consumption of class i at period t.

This problem statement implies that the welfare-maximizing utility allocates its resources so as to set the net marginal benefit for each customer at each time interval equal to the corresponding marginal costs:

$$\frac{\partial W}{\partial x_{it}} = P_{it}(x_{it},R_t) - \frac{\partial TC}{\partial x_{it}} = 0$$

We define the above marginal-cost pricing approach as SPOT2. A special case of the problem is when R=1 (perfect reliability), which is the traditional spot pricing scheme, defined here as SPOT1.

To comply with the regulatory mandate, the revenue collected by the utility has to be constrained to cover total costs. A Ramsey-type pricing formulation would do just that; however, it would lead to a welfare loss. The formulation of a Ramsey-type pricing model, in the case of independent demands, is as follows:

$$Max\ W_{(x_{it})} = \sum_{i\in I}\int_0^{x_{it}} P_{it}(y_{it},R_t)dy - TC_t(X) - FC$$

Subject to:

$$\sum_{i\in I} P_{it}X_{it} = TC_t(X) + FC$$

Solving the above model leads to the well-known inverse elasticity rule. We denote this formulation as RAMSEY2, and its special case of perfect reliability (R=1) as RAMSEY1. In reality, reliability is imperfect for a wide variety of reasons, representing a risk to customers' welfare. Therefore, the use of the traditional version of spot and Ramsey pricing (R=1) is not realistic. However, we will consider both reliability versions and use the four models (SPOT1, SPOT2, RAMSEY1, and RAMSEY2) in order to compare the impact of reliability on welfare. Note that the above four models (SPOT1, RAMSEY1, SPOT2, and RAMSEY2) do not take advantage of the potential benefits of differentiating service reliability, with different levels of willingness-to-pay for better reliability. In the following, such differentiation is considered.

We turn now to the problem of a regulated utility which is allowed to set an additional dimension of service, namely its reliability levels. The utility is allowed to serve its customer classes with different qualities (reliability) of service, according to each customer class's own choice. No regulatory oversight is assumed to take place over the reliability of service except that there is such oversight with regard to cost allocation. The utility is, then, in charge of designing the menu of choices which fits its system-specific limitations.

To allow for service reliability differentiation among customers, we rank customer classes into two groups: a high-reliability customer group and a low-reliability customer group. High-reliability customer classes are assumed to opt for "firm" power service, that is, a 100 percent continuous service. Customer classes in the low-reliability group are ranked in the order of the interruption level they would opt for. At any outage situation, the utility will interrupt low reliability customers in accordance with their chosen ranks. The mechanism to encourage customers to make rational choices is designed so that high reliability customers pay for the high quality "firm" services they opt for. Their incentive to do so is that the payment is less than the expected outage costs they may suffer if they are not protected from outages. Also, low-reliability customer classes are compensated with a payment higher than their expected outage costs. The payment/compensation mechanism is designed to ensure that the electric utility is profiting within the regulated net revenue constraint, and that customers are not cross-subsidizing each other. The value of the payment/compensation for any kilowatt-hour outage protection/curtailment that fits the above criteria is then the average value of all customers' outage costs. We call that value the social outage cost "SOC", and it is equal to the sum of all customers' outage costs divided by the total number of customers.

Because the high reliability customer classes are protected from outages, the total expected shortages must be allocated among the low reliability customers. The total expected shortage at any hour is estimated by using the production cost simulation model, and taken as the total expected demand at the hour multiplied by that hour's loss of load probability. Assuming that shortages happen mainly as a result of excessive demand[10] (excessive demand would lead to excessive use of generating units, which leads to higher probabilities of unit failures), we divide the total expected shortages into: (a) the expected shortages due to the high-reliability demands and (b) the expected shortages due to the low-reliability demands. Then, at any hour, each customer class in the low-reliability group will expect to suffer an "obligatory curtailment" equal to its expected demand multiplied by the loss of load

[10] Other reasons for generation shortages (capacity and/or energy shortages) include fuel shortages, droughts, labor strikes, malfunctions, etc..

probability, plus a "voluntary curtailment" equal to a portion of the expected shortages caused by the high-reliability demands. The voluntary portion is, of course, the customer's choice. The payment that the high-reliability customers have to pay is then the total expected shortages multiplied by the social outage cost. Each low-reliability customer, on the other hand, is compensated by a payment equal to his/her total expected curtailments multiplied by the social outage cost. This allocation mechanism minimize the total expected outage damages for society.

The regulated utility is entrusted to select the optimal set of prices[11] (quantities) which maximize economic welfare subject to the revenue constraint, and to allow for reliability differentiation within the incentive mechanism discussed above. Consider the general case where there are N customer classes, partitioned into two groups: classes 1,2,...,S, which opt for low-reliability service, and classes S+1, S+2,...,N, which opt for high-reliability service. The problem is then to maximize the sum of the aggregate consumers' and producer's surplus, under the constraint that the total revenue must not exceed total costs. The total revenue constraint is then divided into two set of constraints: (a) the first set has S constraints, each allocating part of the revenue to a low-reliability customer class. It is designed so that the low-reliability class covers part of the revenue with payment equal to the class's demand-proportional share of the total cost of service (TC+FC) less a compensation payment. The compensation payment includes two parts: (i) one related to the

[11] In the formulation of the model, we considered demands as the decision variables, knowing that the outcome would be the same.

expected average cost of interruptions caused (directly or indirectly) by the customer's own consumption, and set equal to [SOC*LOLP*customer class consumption], and (ii) one related to the expected average cost of interruptions caused by the total consumption of the high-reliability group and set equal to [SOC*LOLP*total consumption of S+1, S+2,...,N classes] multiplied by α, which is the share of that part of "voluntary" interruptions that the customer is willing to take. It is a customer choice, and the summation of all α's must equal to 1.0. (b) There is only one constraint, that allocates revenue to the high reliability group in amount equal to their demand-proportional share of the total cost of service, plus a penalty payment equal to the total expected average cost of interruptions [SOC*LOLP*total demand]. Note, that this compensation/payment mechanism is purely redistributive (the sum of the compensatory terms over all customers classes is zero) and ensures that the utility exactly covers its fixed plus variable costs. The mathematical formulation of this model is then as follows:

$$Max\ W_{(X_{it})} = \sum_i [\int_0^{X_{it}} P_{it}(y_{it},R_{it})dy] - TC(X) - FC, \qquad i=1,...,N, \qquad t=1,..$$

Subject to:

(a) Constraints related to the revenue collected from low-reliability customers:

$$P_{1t}(X_{1t},R_t)\ X_{1t} = \frac{X_{1t}}{X_{Tt}}(TC_t(X)+FC) - LOLP_t*SOC_t*[X_{1t} + \alpha_{1t}*(X_{s+1,t}+..+X_{Nt})]\ ,$$

$$\vdots$$

$$P_{st}(X_{st},R_t)X_{st} = \frac{X_{st}}{X_{Tt}}(TC_t(X)+FC) - LOLP_t*SOC_t*[X_{st} + \alpha_{st}*(X_{s+1,t}+...+X_{Nt})]\ ,$$

(b) Constraint related to the revenue collected from high reliability customers:

$$\sum_{s+1}^{N} P_{it}(X_{it},1)X_{it} = \frac{X_{s+1,t}+..+X_{Nt}}{X_{Tt}}(TC_t(X)+FC) + LOLP_t*SOC_t*X_{Tt}\ ,$$

where:

X_{it} = demand of class (i) in Mwh at time period t, (i=1,...,N, t=1,..,24),

X_{Tt} = total demand at period t.

N = total number of customer classes.

P_{it} = inverse-demand function of customer class (i). Based on the discussion in section IV.1, it is assumed to have the following form:

$$P_{it}(X,R) = R_{it}(a_{it} - b_{it}X_{it}) - (1-R_{it})OC_{it},$$

R_{it} = level of reliability chosen by (or made available to) customer class (i)[12]. For the high reliability classes R = 1, and for the low reliability classes R = 1 -

[12] Note that R (= 1 - LOLP) is a function of both total demand and system configurations. However, there is no closed form function that can express that relation, which is why we estimate it through simulations. Therefore, R is not considered as a decision variables in our model, although it varies with the variation of the other decision variables (X).

LOLP. However, some low-reliability classes will enjoy higher reliability level than (1-LOLP), and some will enjoy lower level, based on their choices for α's.

OC_{it} = interruption losses to class (i).

SOC_t = average value, among all customers, of social outage costs

= $\Sigma_i OC_{it}/N \quad (i=1,..,N)$

$LOLP_t$ = Loss of load probability, that is, the probability (percentage of time) that the generation system is unable to serve the demand. Note that R = 1 - LOLP.

UE = Unserved energy, that is, the expected amount of energy that the system will not be able to serve due to outages.

FC = Fixed costs of the electric power system ($/Mw).

TC_t = total variable cost ($/Mwh) estimated at each level of consumption by the production cost simulation model.

α_{it} = an allocation parameter related to the level of interruptions selected in advance by customer class i. It represents the share of the "voluntary" interruptions that customer i is willing to accept. Note that, $\Sigma^{s}_{i=1} \alpha_i = 1$

s = total number of classes that opted for low reliability service.

For the numerical and analytical purposes of this research, we simulate two special cases of the above model: (a) three-reliability-options, four-classes [RDP-3], and (b) two-reliability-options, four-classes [RDP-2]. In the analysis we assume that class 1 and class 4 are opting for high "firm" reliability, and classes 2 and 3 are opting for low reliability levels. This selection is based on the fact that in the special case study presented in chapter VI we assume that class 1 and class 4 have more inelastic demands and higher outage costs than in the case of the other two classes.. The formulation of the three-options reliability model is as follows:

RDP-3:

$$Max\ W_{(X_{it})} = \sum_i^4 [\int_0^{X_{it}} P_{it}(y_{it},R_t)dy] - TC_t(X) - FC, \qquad (t=1..,24)$$

Subject to:

$$P_{2t}(X_{2t},R_t)X_{2t} = \frac{X_{2t}}{X_{Tt}}(TC_t(X)+FC) - LOLP_t*SOC_t*[X_{2t} + \alpha*(X_{1t}+X_{4t})],$$

$$P_{3t}(X_{3t},R_t)X_{3t} = \frac{X_{3t}}{X_{Tt}}(TC_t(X)+FC) - LOLP_t*SOC_t*[X_{3t} + (1-\alpha)*(X_{1t}+X_{4t})]$$

$$P_{1t}(X_{1t},1)X_{1t} + P_{4t}(X_{4t},1)X_{4t} = \frac{X_{1t}+X_{4t}}{X_{Tt}}(TC_t(X)+FC) + LOLP_t*SOC_t*X_{Tt}$$

The two-options reliability model is formulated as follows:

RDP-2:

$$Max\ W_{(X_{it})} = \sum_i^4 [\int_0^{X_{it}} P_{it}(y_{it},R_t)dy] - TC_t(X) - FC, \qquad (t=1,..,24)$$

Subject to:

$$P_{2t}(X_{2t},R_t)X_{2t} + P_{3t}(X_{3t},R_t)X_{3t} = \frac{X_{2t}+X_{3t}}{X_{Tt}}(TC_t(X)+FC) - LOLP_t*SOC_t*X_{Tt},$$

$$P_{1t}(X_{1t},1)X_{1t} + P_{4t}(X_{4t},1)X_{4t} = \frac{X_{1t}+X_{4t}}{X_{Tt}}(TC_t(X)+FC) + LOLP_t*SOC_t*X_{Tt}$$

The conceptual framework to implement any of these models can be summarized as follows. The electric utility has an initial estimation (or forecast) of the hourly demands of each customer class. The utility also owns information about the generation mix, unit availabilities, dispatching, and fuel costs. The utility uses its dispatching center (production cost simulation) to estimate the minimum combinations of input factors (minimize fuel and maintenance costs) in order to maximize profit. Then the regulated utility is able to reveal information about total expected production costs (TC), marginal costs (MC), and level of reliability R, which are the components needed to design a price menu. Customers reactions to these price signals lead to changes in their consumption and reliability decisions, enforcing changes in production, costs and system reliability. A partial market equilibrium is reached when prices are reflecting the customers valuations to the service. The recent success of implementing some real-time pricing experiment, is a sign that the implementation of this mechanism should succeed since its a proxy of real-time pricing.

Chapter V

Data, Results, and Analysis

In this chapter, we present applications of the pricing models developed in chapter IV. The nonlinear nature of these models requires the use of the Iterative Newton-Raphson technique in order to obtain numerical solutions. To use this technique, the first- and second-order conditions of the optimization model must be obtained. The mathematical detail of these conditions is presented in appendix B. The flow chart in figure V.1 depicts the integration framework of the model. The following steps illustrate the procedure we use to conduct the analysis.

1- Prepare data inputs: power system generation mix and characteristics, customer class hourly initial load profiles, customer class hourly outage cost, and customer class hourly elasticity of demand.

2- Estimate the average cost prices using the production cost simulation model.

3- Use initial load profiles, estimated average prices, and assumed elasticities of demands, to estimate the parameters of the demand functions for each class, using the linear demand functions suggested

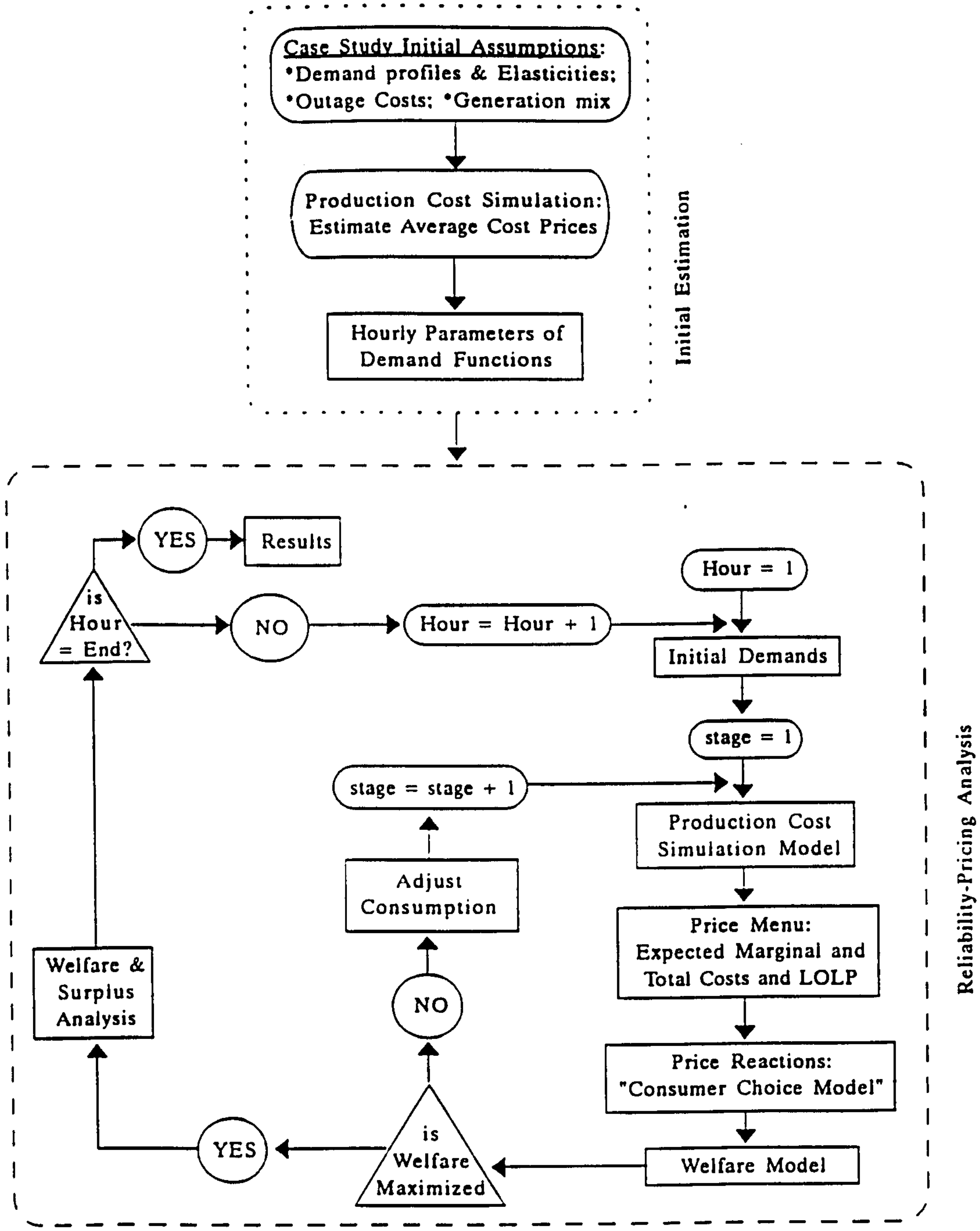


Figure [V• 1] Flow Chart of the Pricing Model in section IV.1.

5- The following iterative procedures are conducted at each hour to estimate the optimal prices/consumptions for any of the proposed pricing models:

(a) Set stage =1

(b) Estimate the expected marginal and total costs, and the reliability level (LOLP) for total consumption at this stage.

(c) Estimate the related prices.

(d) Use the Newton-Raphson technique to solve for the first and second order equations. The model converges at the optimal prices and consumptions, which gives the optimal related welfare and surpluses. If convergence is obtained, go to (6), otherwise go to (e).

(e) Change consumptions incrementally and iterate one more stage (go back to stage a).

(6) Repeat steps (a) to (e) for every time period of the study.

V.1 Data

This section describes the assumed generation system, the initial load profiles and elasticities, and customers' outage costs. The objective is to define a power system broadly enough to provide a basis for verifying the validity of our method and analysis.

V.1.1 Power System

Table V.1 shows the generation mix of the assumed power system. The system is identical to the IEEE reliability test system [IEEE-PAS,1972]. This generation system is well known to accurately represent a typical electric utility generation system and contains different sizes, types, costs, and availability rates of power plants. The table includes the type, size, incremental cost, forced outage rate (FOR), mean time failure (MF), and mean time repair (MR) of each unit in the system. The forced outage rate (FOR) is the probability (percentage of time) that the unit will be down or out of service. The mean time failure (MF) is the mean time between failures and the mean time repair (MR) is the mean duration of failures. Information about both parameters (MF and MR) can be easily found in the unit statistics reports. The forced outage rate (FOR) is related to these two parameters, by the general Markovian relation:

$$FOR = \frac{1}{MR+MF} * [\ MR + MF * \exp^{-t*\frac{MR}{MR+MF}}\ ]$$

where t measures the number of hours the unit has operated. This formula converges to the simple form:

$$FOR = \frac{MR}{MR+MF}$$

Table V.1 Details of Generation System

Unit size (MW)	No. units	Fuel type	Fuel cost ($/Mwh)	Forced outage rate	Mean time failure (hrs.)	Mean repair (hrs.)
400	2	Nuclear	5.59	0.12	1100	150
350	1	Coal	11.40	0.08	1150	100
197	3	#6 Oil	19.87	0.05	950	50
155	4	Coal	11.16	0.04	960	40
100	3	#6 Oil	22.08	0.04	1200	50
76	4	Coal	14.88	0.02	1960	40
50	6	Hydro	16.30	0.01	1980	20
20	4	#2 Oil	37.50	0.10	450	50
12	5	#6 Oil	28.56	0.02	2940	60

Total Installed Capacity = 3405 MW

when the number of hours the unit has operated becomes large (t = 5000 hours). If the unit has just started (t is small), the value of FOR becomes very close to unity (FOR = 1.0). In this research we consider only the simple formula. The problem with this assumption is that it yields the same reliability results for the same system with the same demand even if the analysis is five years earlier or later. The Markovian formula would relate the results to the history of unit operations. Since we are conducting our research for one day only, the simple formula is more appropriate for our analysis.

The information given in table V.1 about incremental costs of generation is in terms of dollars per megawatt-hour. This information has not been updated, since we are only interested in demonstrating the validity of the proposed model through a comparative analysis of different pricing schemes.

V.1.2 Demand Pattern

Four classes of consumers are considered: large industrial users, small industrial users, commercial users, and residential users. The consumption patterns of these classes are obtained from a typical utility (Ohio Edison) and the loads are scaled to match the IEEE generation model discussed above. Therefore, the total system peak load is made identical to the system peak suggested in the IEEE reliability test system [peak=2850Mw]. The hourly demands for a typical day are shown in table V.2. The analysis over a day is, of course, for illustrative purposes. More extensive studies would consider the demand over a year, or a season at least.

Table V.2 Initial Hourly Demand (MW)

Hour	Class#1	Class#2	Class#3	Class#4	Total
1	91.5	734.8	426.9	411.4	1664.6
2	90.4	726.1	363.1	387.9	1567.6
3	89.2	729.2	314.7	371.6	1504.9
4	88.9	716.8	291.1	360.7	1457.6
5	89.6	711.2	285.9	359.1	1445.8
6	92.2	733.6	305.5	367.4	1498.8
7	99.9	787.6	388.3	406.4	1682.3
8	110.9	851.6	475.7	529.9	1968.3
9	115.9	873.4	558.1	685.4	2232.8
10	118.1	880.8	628.1	806.5	2433.4
11	120.2	888.3	703.1	867.2	2578.9
12	120.2	878.3	750.9	915.1	2664.5
13	120.2	887.1	722.1	901.6	2630.9
14	122.5	890.7	743.2	943.5	2700.0
15	120.5	870.3	760.7	940.6	2692.1
16	115.1	831.7	862.5	905.4	2714.8
17	109.2	808.8	1058.5	873.5	2850.0
18	103.6	787.6	1089.9	734.1	2715.1
19	101.4	785.8	1032.8	680.8	2600.8
20	100.1	774.6	874.3	645.2	2394.3
21	99.1	769.6	858.9	634.3	2361.9
22	99.6	784.5	708.2	564.7	2157.2
23	97.2	778.3	634.7	502.3	2012.6
24	94.8	754.1	511.7	445.3	1806.1

Note: Class #1 = Large industrial customers,
Class #2 = Small industrial customers,
Class #3 = Residential customers, and
Class #4 = Commercial customers.

V.1.3 Average Price of Electricity

The generation and demand data described above are initially used for the purpose of estimating the parameters of the proposed demand functions. We base this study on the demand functions' parameters that are estimated at the average prices. The average prices estimated in this section are related to the average cost of production, to meet the demands in table V.2. This cost is obtained using the proposed stochastic production cost simulation model. The expected production cost of supplying the demand is found to be \$35,999.7/hour, or an average of \$12.6315/Mwh.

To account for capital and other fixed costs, we note that the fuel component in the price of electricity ranges from 30 percent to 70 percent. This component depends on the generation mix and fuel prices. Consequently, we assume in this research that the fixed costs represent one-third of the production costs (that is, \$11,999.90). The fixed costs include the undepreciated capital component of equipment.

To estimate the actual average price tariff, we allocate the fixed costs among customer classes according to the ratios of their demand at the system peak hour to the total system peak demand. Therefore, the following formula is used to calculate the average price of electricity for each class:

Average price for class (i) = [total production costs/total demand] +

fixed costs * [class (i) demand at system peak/system peak demand]

The initial average cost prices for each class are then:

P_1 = 12.7925 \$/Mwh (Large Industrial)

P_2 = 13.8260 \$/Mwh (Small Industrial)

P_3 = 14.4940 \$/Mwh (Residential)

P_4 = 13.9217 \$/Mwh (Commercial)

These prices are assumed constant over the period of study (24 hours), and reliability level is assumed to have no impact on their calculations. They are used with the hourly demand elasticities data for each customer class to estimate the required demand functions' parameters.

V.1.4 Demand Elasticities and Parameters

In this research we assume that the hourly elasticities of demand for each class are known. A typical electric utility (Pacific Gas & Electric, PG&E) has reported such information in its annual real time pricing report. Table V.3 is a slightly adjusted version of the data we borrowed from the report. The reason we adjusted the data was that we found that some reported hourly elasticities were positive, possibly due to a sudden shift of demand from the peak to the intermediate period (where price is also, but slightly, increasing).

The information in table V.3 is crucial to this analysis. We use it, in combination with the average-cost prices, to estimate the demand functions coefficients a's and b's. Using the price-demand function forms (assuming perfect reliability) discussed earlier, we get:

Table V.3 Demand Elasticities

Hour	Class#1	Class #2	Class#3	Class#4
1	-0.05	-0.7	-0.81	-0.2
2	-0.081	-0.7	-0.81	-0.2
3	-0.09	-0.7	-0.81	-0.2
4	-0.101	-0.7	-0.8	-0.01
5	-0.1	-1.0	-0.81	-0.05
6	-0.087	-1.25	-0.8	-0.011
7	-0.084	-1.25	-0.8	-0.011
8	-0.069	-1.25	-0.18	-0.02
9	-0.047	-1.5	-0.15	-0.011
10	-0.054	-1.5	-0.15	-0.011
11	-0.085	-1.45	-0.15	-0.02
12	-0.085	-1.4	-0.18	-0.02
13	-0.085	-1.4	-0.2	-0.015
14	-0.064	-1.0	-0.2	-0.01
15	-0.05	-1.0	-0.19	-0.011
16	-0.036	-1.0	-0.2	-0.01
17	-0.036	-1.0	-0.11	-0.015
18	-0.036	-1.2	-0.11	-0.015
19	-0.04	-1.4	-0.08	-0.1
20	-0.04	-1.2	-0.13	-0.1
21	-0.036	-1.0	-0.14	-0.19
22	-0.042	-0.7	-0.14	-0.17
23	-0.05	-0.7	-0.14	-0.15
24	-0.05	-0.7	-0.14	-0.19

Note: Class #1= Large industrial customers,
Class #2= Small industrial customers,
Class #3= Residential customers, and
Class #4= Commercial customers.

$$P_{it}(X_{it},1) = a_{it} - b_{it}X_{it} \qquad (1)$$

$$\eta_{it} = \frac{P_{it}}{X_{it}} \partial X_{it} / \partial P_{it} = \frac{P_{it}}{X_{it}}\left(-\frac{1}{b_{it}}\right) \qquad (2)$$

where i denotes the customer class, t is the time period (hour), the X_{it}'s are the initial demands depicted in table V.2, and the P_{it}'s are the average cost prices estimated in section V.1.3.

The two equations (1) and (2) are solved for the variables a_{it} and b_{it} for each hour and for each class of customers. Table V.4 presents the results.

V.1.5 Outage Costs

Outage costs reflect the price a customer is willing to pay to avoid an outage. These costs may be incurred by customers when service is lost in the form of lost wages or profits, or in the form of spoilage or interruption of a production process, or in other forms. Outage costs depend upon the timing, frequency, duration, notification, and extent of the outage. Outages tend to be less costly as customers have more warning time to prepare for the outage. The outage cost thus provides an accurate measure of the value of reliability to the customer.

In this research, it is assumed that outages occur without prior notification and are of "short and equal" durations. Outage costs are then defined as the costs incurred by the customers due to a sudden shortage of power supply, over and above any loss in consumers' benefit due to shortage. This is because a sudden shortage of supply results in losses due to interruption of production processes which could have been avoided by rescheduling the process if prior notice of the outage was available. The outage costs data for the four classes under study are drawn from studies and reports prepared by the Electric Power Research Institute (EPRI) over the last 10 years. The data for the residential class is drawn from EPRI-EA-2462 (June 1982, page 8-2 and table 8-1), conveniently provided for each hour of the day. The information for the three other classes is taken from EPRI-P6510 (Sep. 1989, page 5-56). Since these data were not provided per hour, we use the hourly distribution (percentage) curves to estimate the hourly outage costs for each class. These curves are also provided in EPRI-P6510 (page 2-32). Table V.5 combines all these informations.

Table V.4 Parameters of Demand Functions

Hour	a_1	a_2	a_3	a_4	b_1	b_2	b_3	b_4
1	268.6	33.6	31.7	83.5	2.79	0.026	0.041	0.169
2	170.7	33.6	31.7	83.5	1.74	0.027	0.048	0.179
3	154.9	33.6	31.7	83.5	1.59	0.027	0.055	0.187
4	139.4	33.6	31.9	1406.	1.42	0.027	0.061	3.859
5	140.7	27.6	31.7	292.3	1.42	0.019	0.061	0.775
6	159.8	24.9	31.9	1279.	1.59	0.015	0.058	3.444
7	165.1	24.9	31.9	1279.	1.52	0.014	0.045	3.114
8	198.2	24.9	93.1	710.1	1.67	0.013	0.165	1.313
9	284.9	23.0	108.8	1279.	2.34	0.010	0.169	1.846
10	249.7	23.0	108.8	1279.	2.01	0.010	0.150	1.569
11	163.2	23.3	108.8	710.1	1.25	0.010	0.134	0.802
12	163.2	23.7	93.1	710.1	1.25	0.011	0.105	0.760
13	163.2	23.7	85.2	942.1	1.25	0.011	0.098	1.029
14	212.7	27.6	85.2	1406.	1.63	0.015	0.095	1.475
15	268.4	27.6	88.9	1279.	2.12	0.015	0.098	1.345
16	368.1	27.6	85.2	1406.	3.08	0.016	0.082	1.537
17	368.1	27.6	143.2	942.1	3.25	0.017	0.121	1.062
18	368.1	25.3	143.2	942.1	3.43	0.014	0.118	1.264
19	332.6	23.7	191.6	153.1	3.15	0.012	0.171	0.204
20	332.6	25.3	123.4	153.1	3.19	0.014	0.124	0.215
21	368.1	27.6	115.6	87.2	3.58	0.018	0.118	0.115
22	317.4	33.6	115.6	95.8	3.05	0.025	0.143	0.145
23	268.6	33.6	115.6	106.7	2.63	0.025	0.159	0.184
24	249.7	33.6	115.6	87.2	2.49	0.026	0.198	0.164

Table V.5 Hourly Outage Costs

Class1 ¢/Kwh	Class2 ¢/Kwh	Class3 ¢/Kwh	Class4 ¢/Kwh
14.91	14.44	9.51	18
14.91	13.98	9.58	18
14.76	14.14	9.56	18
14.49	14.16	9.92	18
14.76	14.16	9.54	18
14.76	14.16	9.81	18
14.76	14.68	9.90	18
14.49	14.38	10.03	18
14.49	14.06	19.26	21.6
14.49	14.38	10.43	21.4
14.76	14.74	10.61	21.2
14.76	14.74	10.7	20.2
14.76	14.84	10.97	20.2
14.49	14.68	10.41	20.2
14.49	14.54	11.59	20.2
14.49	14.52	10.70	20.2
14.49	14.42	10.08	28.6
14.49	13.5	9.36	28.4
14.49	13.5	8.47	38.8
14.49	13.5	7.85	23.6
14.49	13.5	7.58	22.4
14.49	13.5	7.49	23.0
14.49	13.5	7.40	22.0
14.49	13.5	7.31	22.6

Note:
Class1 = Large industrial customers,
Class2 = Small industrial customers,
Class3 = Residential customers, and
Class4 = Commercial customers.

V.2 Results and Analysis

We examine the welfare gains and energy and reserve savings possibilities due to different pricing schemes. The base case for comparison is that of the traditional spot pricing. The analysis simulates the effect of different pricing schemes on the four customer classes. Benefits due to these pricing schemes are compared to the spot pricing, which reflects the traditional optimal "first best" pricing.

After conducting the initial calculations of average pricing and demand functions parameters, an estimation of the spot pricing models is undertaken. Under the assumed fixed costs, this estimation indicates that the utility's revenue exceeds its total cost. Therefore, there is overrecovery of fixed costs under the two forms of spot pricing [SPOT1 and SPOT2].

Next, the RAMSEY1 and RAMSEY2 models are solved to constrain utility revenues to not exceed total costs. RAMSEY2 differs from RAMSEY1 by accounting for the impact of reliability on customer satisfaction. In both models, the fixed cost to be recovered is fixed at all hours at the value of one-third of the production cost of the initial average-cost prices (that is, $11,999.90).

Finally, the effects of reliability-based pricing are estimated at two different levels: three reliability options [RDP-3], and two reliability options [RDP-2].

The results are presented in tables V.6 through V.16, and in figures V.1 through V.4, leading to several conclusions that must, of course, be qualified by the assumptions outlined earlier. Table V.6 indicates the economic welfare gains arising from the different pricing schemes, with spot pricing as the base case. The results show that under perfect reliability conditions, spot pricing (SPOT1) yields the highest economic efficiency and Ramsey pricing (RAMSEY1) yields the second best. However, the model of reliability-based pricing with two options (RDP-2) yields very close welfare gains and energy and reserve savings. In table V.7, the values of the loss of load probability (LOLP), estimated in relation to hourly total demands for each pricing schemes, demonstrates that the perfect reliability world does not exist. Therefore, we focus on comparing the four pricing models that consider imperfect reliability, namely SPOT2, RAMSEY2, RDP-2, and RDP-3.

Table V.6 indicates that reliability-based pricing yields the highest economic efficiency (7,061,631), consumer's surplus, and energy savings. Spot pricing yields the second best and Ramsey pricing yields the third best results. The significance of these results is very important, clearly showing that reliability-based pricing, in the proposed forms, is economically feasible. Reliability-based pricing is shown to improve efficiency and load factors, and save energy and reserve power.

The results provide many other important observations. Table V.7 presents the system hourly and total reliability level, under all pricing schemes. The reliability under Ramsey pricing is shown to be higher (smaller LOLP) than under spot pricing and the other models. The reason is, obviously, that prices under the Ramsey framework are generally higher [see tables V.9 through V.11 and figures V.2 through V.5], and consequently consumption is lower [see tables V.12 through V.16 and figure

V.5]. Since the proposed form of the production cost simulation model yields higher reliability for lower demand[13], then the lower demands under Ramsey cause the estimated reliability level to be higher. Next, the overall and hourly reliability profile is much worse (smaller) under reliability-based pricing models [RDP-2,RDP-3] than under spot and Ramsey prices. The reason is that under the later schemes [spot and Ramsey], the reported reliability levels are served to all customer classes without discrimination. Under reliability pricing, however, the reported levels are served only to those who opt for the lower reliability level. The remainder of customer classes, under this scheme, are served with firm power. Finally, the reliability levels under the three reliability options model (RDP-3) are much smaller than the levels under the two reliability options model (RDP-2), because, under the two-options model, the allocation of outages and compensations among low-reliability customer classes is based on the inverse elasticity rule. Under this rule, higher prices are assigned to the less elastic class #3, which also happens to have the highest demand among all classes. This results in lower consumptions and higher reliability levels under the two-options model.

Tables V.9, V.10, and V.11, and figures V.2 through V.5 depict the price-menus for each class under the different pricing schemes. Several points can also be noted: (a) The reliability-based models yield higher prices for the high-reliability customers

---

[13] The reason is that the forced outage rate (FOR) of generation units are assumed fixed and not function of the unit operating time. Also assumed is that all units are available at the initial stage (t=0). Using Markovian analysis to account for the status of each unit in relation to its operation history, would improve the accuracy of the estimated relation between reliability and demand levels.

during peak and shoulder hours than under the Ramsey model. This result is consistent with the basic economic premises that motivated this research, namely, that higher prices should be paid for higher quality commodities or services; (b) Reliability-pricing models yield lower prices for the high-reliability customer classes during off-peak hours than under the Ramsey model. Table V.7 shows that the LOLP values during off-peak hours are zeros, that is, perfect reliability. Therefore, during off-peak hours, reliability plays no role in the determination of prices. When reliability is 100 %, the role of cost allocation in proportion to class demands becomes more important in price design. Since high-reliability classes have the lower demands, their prices during off-peak hours become lower than they are under Ramsey, where only the inverse elasticity rule affects prices.

The above analysis could also be applied to the prices for the low-reliability customer classes. The reliability-based prices are higher during off-peak hours than under Ramsey and spot pricing schemes. The reason is that reliability plays no role during off-peak hours. Then, the allocation rule allocates higher revenues to be collected from those customer classes with higher demands, which happen to be the low-reliability classes. However, during peak and shoulder hours, the results show that the low-reliability classes are paying lower prices under reliability-based pricing that what they are under Ramsey and spot pricing. It has to be mentioned that this logical result is only true for the three-options reliability model. Under the two-options model, the prices for class 2 (low-reliability class) during peak and shoulder hours are closer and, at some hours slightly higher than the corresponding prices under Ramsey pricing. The explanation is that, under the two-options model, class 2 and class 3 prices are determined by both reliability and the inverse elasticity rule. Since class 2 has a much higher price-elasticity of demand during peak hours (see table V.3), its prices become much higher than those of class 3 under the same two-option pricing, and slightly higher than its own prices under Ramsey pricing.

In the three-options reliability model [RDP-3], we arbitrary allocate the part of outages caused by the demand of the high reliability customer classes LOLP*$(X_1 + X_4)$ between the low-reliability customer classes. We assumed that each class opts to share 50 percent of these interruptions, that is, $\alpha = 0.5$. In real cases, of course, customers can chose any other value for their share of interruption. We did run the model with several values for α, and the results are shown in table V.8. When α = 0.0, class 3 carries all the interruption caused by the demand of the high reliability customer classes, and when α = 1.00, class 2 carries all these interruptions. The results reveal that the more any of the low reliability customer classes is interrupted, and consequently compensated, the higher its surplus becomes. However, they also show that the total economic welfare increases only when customer class 3 opts for a higher share of interruptions. This result is due to the fact that class 3 has a much higher demand (less elastic) than class 2, which means that the increase in class 3 surplus due to interruption compensation is much higher than any decreases in class 2 surplus.

Table V.6 Changes in Welfare and Energy Under Different Pricing Schemes

	Spot1	Ramsey1	Spot2	Ramsey2	RDP-2	RDP-3
Welfare	7077795	7077032	7028047	7027265	7061632	7042272
% change	0.0	-0.01	-0.7	-0.71	-0.23	-0.5
Surplus:						
Class#1	282517	281767	280855	280126	282956	282699
Class#2	84282	82064	82449	80718	69893	82721
Class#3	636837	644863	629968	638466	658371	635315
Class#4	5753585	5780335	5714371	5739956	5762500	5753527
Total	6757229	6789030	6707645	6739268	6773722	6754263
% change	0.0	0.47	-0.73	-0.26	0.24	-0.04
Peak Demand MW	2501.8	2468.6	2440.1	2428.8	2457.3	2651.2
% change	0.0	-1.32	-2.46	-2.91	-1.77	5.97
Energy Mwh:						
Class1	2464.9	2451.3	2463.9	2450.2	2464.2	2463.8
Class2	14321.1	14276.6	14344.5	14108.6	13282.7	14972.7
Class3	14777	14821	14753	14795	14708	14575
Class4	15065	15049	15066	15046	15066	15052
Total	46643.1	46611.4	46641.9	46409.7	45539.0	47077.5
% change	0.0	-0.07	-0.002	-0.5	-2.3	0.9

Table V.7 Reliability Levels
Loss Of Load Probability*

	Pricing Schemes					
Hour	SPOT1	RAMSEY1	Spot2	Ramsey2	RDP-2	RDP-3
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0
7	0	0	0	0	0	0
8	0.114	0.075	0.155	0.0741	0.0664	0.046
9	0.262	0.63	0.779	0.563	0.578	0.84
10	1.14	1.73	2.5	1.68	2.01	4.2
11	2.42	3.98	4.9	3.49	4.07	10
12	7.62	6.09	6.7	5.2	6.27	17
13	3.3	4.73	2.8	4.38	5.48	14
14	13.1	10.7	11.1	9.48	13	29
15	13.1	10.8	6.95	9.59	13.1	29
16	14.1	12.2	8.06	10.8	15	34
17	32.6	26.8	22.6	21	25	74
18	13.3	11.2	11.3	10	11.3	32
19	6.34	5.69	6.39	5.25	5.23	14.5
20	1.52	2.57	2.75	2.48	2.49	4.7
21	1.64	2.63	1.45	2.54	2.65	4.1
22	0.992	0.977	1.00	0.959	0.976	1.18
23	0.157	0.257	0.154	0.256	0.261	0.266
24	0	0	0	0	0	0
Total Loss Of Load Probability (LOLP/day):						
	111.7	101.0	89.5	87.7	107.4	268.8

* Unit = LOLP x 10^3

Table V.8 RDP-3 with Different Allocations

Allocation Factors*	α=0.0	α=0.25	α=0.33	α=1.00
Welfare	7046372	7044491	7043821	7036199
Surplus:				
Class2	80970	81805	82088	84898
Class3	639443	63736	636863	629551
Total Consumers' surplus:				
	6758388	6756530	675871	6748179

* α is the share of interruption chosen by class 2, and $(1-\alpha)$ is the share chosen by class 3.

Table V.9 Spot and Ramsey Prices ($/Mwh)

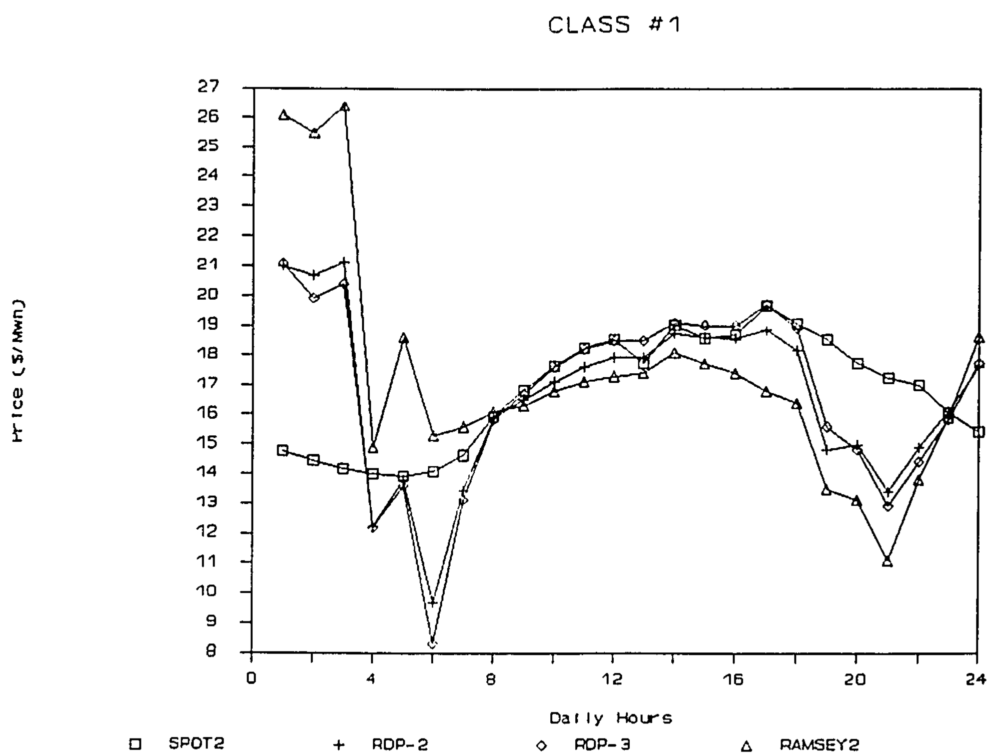
		RAMSEY1				RAMSEY2			
SPOT1	SPOT2	p1	p2	p3	p4	p1	p2	p3	p4
14.78	14.78	26.1	15.4	15.4	17.7	26.1	15.4	15.4	17.7
14.43	14.43	25.5	15.4	15.3	19.1	25.5	15.3	15.3	19.1
14.18	14.18	26.4	15.3	15.2	19.9	26.4	15.2	15.2	19.9
13.99	13.99	14.9	14.1	14.1	25.2	14.9	14.1	14.1	25.1
13.93	13.93	18.6	14.2	14.3	24.4	18.6	14.3	14.3	24.4
14.07	14.07	15.3	14.2	14.2	24.5	15.3	14.2	14.2	24.5
14.63	14.63	15.6	14.9	14.9	21.1	15.6	14.9	14.9	21.1
15.96	15.96	16.1	15.8	15.9	17.0	16.1	15.9	15.9	17.0
16.28	16.82	16.2	16.4	16.4	15.7	16.3	16.4	16.4	15.5
17.06	17.64	16.8	17.3	17.1	14.5	16.8	17.1	17.1	14.5
17.61	18.24	17.1	17.8	17.4	14.2	17.1	17.4	17.4	14.1
18.66	18.54	17.3	18.2	17.8	13.9	17.3	17.8	17.8	13.8
17.87	17.73	17.4	18.1	17.9	13.8	17.4	17.8	17.8	13.8
19.18	19.02	18.1	18.9	18.7	12.9	18.1	18.6	18.6	12.9
19.18	18.57	17.7	18.9	18.6	12.9	17.7	18.5	18.5	13.0
19.24	18.71	17.4	19.0	18.8	12.5	17.4	18.7	18.7	12.6
19.91	19.65	16.8	19.7	18.7	12.1	16.8	18.6	18.6	12.3
19.19	19.05	16.4	19.0	18.1	12.2	16.4	18.1	18.1	12.4
18.48	18.51	13.4	18.3	15.7	16.3	13.5	15.7	15.7	16.3
17.26	17.72	13.0	17.5	16.1	15.7	13.1	16.1	16.1	15.7
17.31	17.22	11.0	17.5	15.8	16.3	11.1	15.8	15.8	16.3
16.97	16.99	13.8	16.8	15.9	16.1	13.8	15.9	15.9	16.1
16.07	16.07	15.9	16.2	16.1	16.1	15.9	16.1	16.1	16.1
15.41	15.41	18.6	15.6	16.8	16.3	18.6	16.8	16.8	16.3

Table V.10 Price Menu
The 2-options Reliability Model [RDP-2]

Hour	SPOT1	P1	P2 ($/Mwh)	P3	P4	LOLP X 10^{-3})	Outage Mw 10^{-1}
1	14.78	21.01	16.89	16.64	15.86	0	0
2	14.43	20.67	17.28	16.94	16.36	0	0
3	14.18	21.12	17.59	17.19	16.61	0	0
4	13.99	12.21	17.71	17.34	18.93	0	0
5	13.93	13.81	17.81	19.11	19.28	0	0
6	14.07	9.66	17.38	19.84	20.29	0	0
7	14.63	13.42	16.49	18.08	17.91	0	0
8	15.96	15.91	15.88	16.73	16.29	0.0664	0.118915
9	16.82	16.51	16.54	15.67	16.12	0.578	1.133978
10	17.64	17.12	17.22	15.12	16.12	2.01	4.203714
11	18.24	17.61	17.83	14.48	16.21	4.07	8.870972
12	18.54	17.94	18.21	14.72	16.31	6.27	13.41842
13	17.73	17.93	18.04	14.73	16.26	5.48	12.16943
14	19.02	18.73	18.61	13.99	16.62	13.1	30.82168
15	18.57	18.61	18.63	14.05	16.65	13.01	30.58911
16	18.71	18.59	18.76	14.29	16.77	15	35.6175
17	19.65	18.84	19.45	14.21	17.48	25	61.4325
18	19.05	18.15	18.89	14.94	16.74	11.3	26.32787
19	18.51	14.82	18.24	15.21	16.82	5.23	11.57974
20	17.72	14.96	17.55	15.27	16.52	2.49	5.270334
21	17.22	13.39	17.38	15.29	16.77	2.65	5.62913
22	16.99	14.91	16.74	15.54	16.43	0.976	1.9642
23	16.07	16.11	16.19	16.11	16.18	0.261	0.494151
24	15.41	17.59	15.69	17.17	16.04	0	0

Table V.11 Price Menu
The 3-options Reliability Model [RDP-3]

p1	p2	p3	P4	Expected power interruption to class2 (Mwh)	Class3 (Mwh)
	$/Mwh				
21.1	16.8	16.8	15.8	0	0
19.9	17.2	17.2	16.5	0	0
20.4	17.5	17.5	16.7	0	0
12.2	17.6	17.6	18.9	0	0
13.6	18.2	18.2	19.4	0	0
8.3	18.3	18.3	20.9	0	0
13.1	17.1	17.1	18.1	0	0
15.8	16.2	16.2	16.3	0.0044	0.0035
16.7	16.1	16.1	16.1	0.0885	0.0795
17.6	16.1	16.1	16.2	0.4663	0.4513
18.2	16.0	16.0	16.5	1.1395	1.1755
18.5	15.9	15.9	16.9	1.9625	2.1105
18.5	15.9	15.9	16.7	1.624	1.688
19.1	15.6	15.6	17.5	3.5925	3.6145
19.0	15.6	15.6	17.6	3.547	3.659
19.0	15.5	15.6	17.8	3.982	4.547
19.7	14.3	15.6	20.4	8.38	11.235
18.9	15.6	14.7	18.1	3.2585	4.7415
15.6	15.9	15.7	17.5	1.4015	2.0315
14.8	16.1	16.0	16.7	0.459	0.575
12.9	16.1	16.1	17.0	0.4045	0.4895
14.4	16.1	16.1	16.5	0.1197	0.1201
15.8	16.1	16.1	16.2	0.026	0.0243
17.7	16.3	16.3	16.0	0.000	0.000



<u>Figure V.2</u>

Hourly Prices Under Different Pricing Schemes.
(Large Industries)

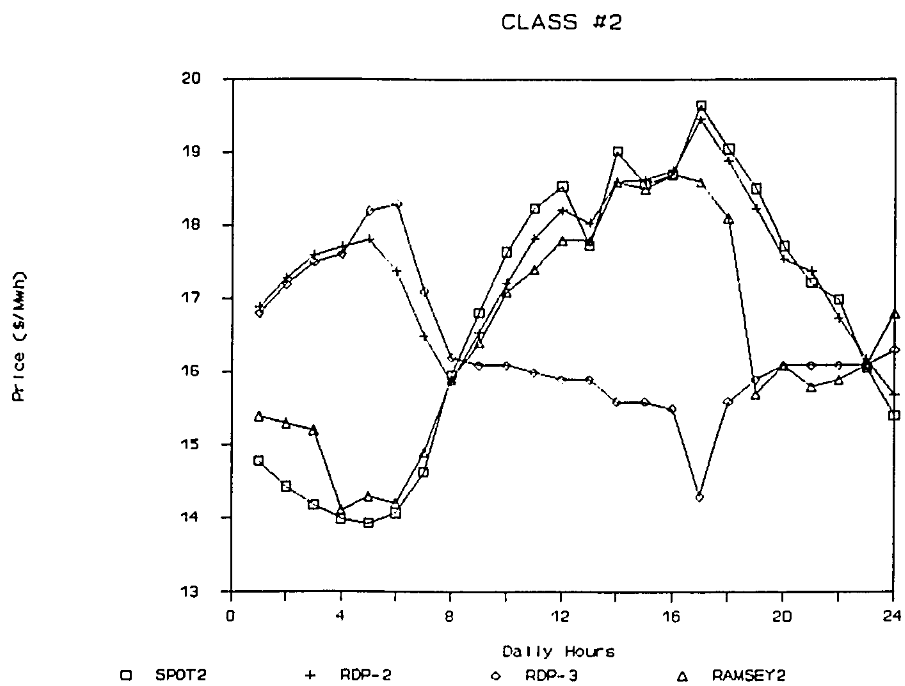
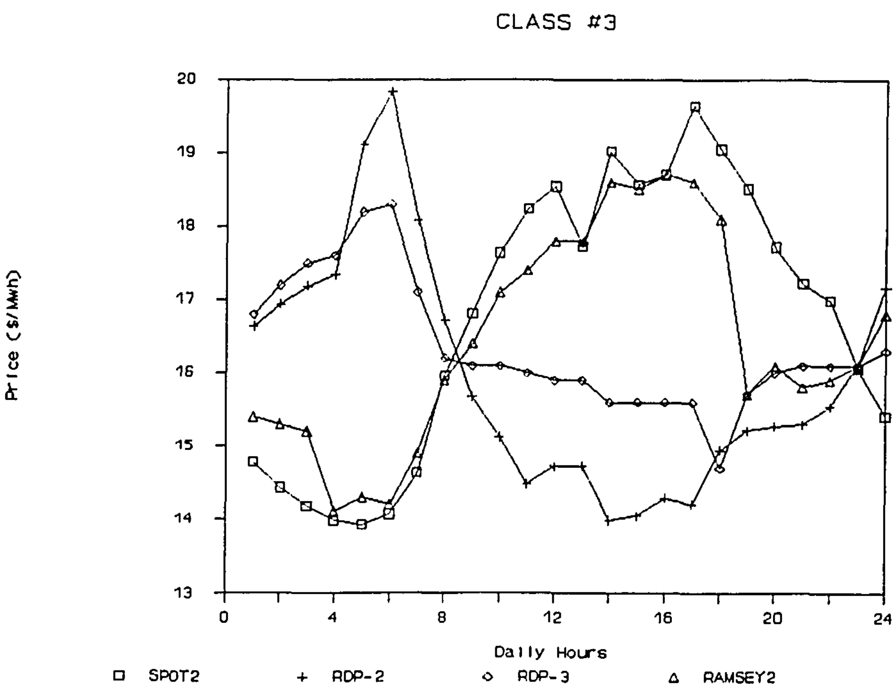


Figure V.3

Hourly Prices Under Different Pricing Schemes.
(Small Industries)



<u>Figure V.4</u>

Hourly Prices Under Different Pricing Schemes.
(Residential Class)

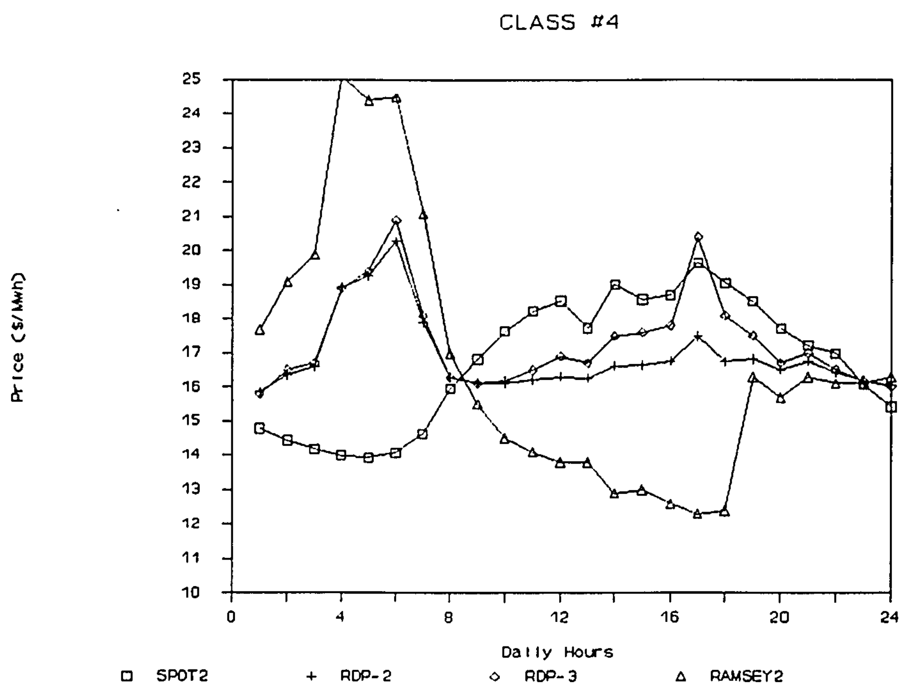


Figure V.5

Hourly Prices Under Different Pricing Schemes.
(Commercial Class)

Table V.12 Total Hourly Demands (Mwh)

Hour	Spot1	Spot2	Ramsey1	Ramsey2	RDP-2	RDP-3
1	1580.2	1580.2	1547.4	1547.4	1476.7	1475.2
2	1518.4	1518.4	1450.1	1450.1	1365.4	1363.5
3	1479.7	1479.7	1389.1	1389.1	1292.4	1290.6
4	1451.7	1451.7	1444.7	1444.7	1264.2	1263.8
5	1442.3	1442.3	1405.6	1405.6	1153.1	1146.1
6	1463.9	1463.9	1468.6	1468.6	1165.4	1131.2
7	1552.3	1552.3	1587.7	1587.7	1405.8	1383.7
8	1829.6	1830.1	1799.4	1799.2	1790.9	1766.2
9	1893.4	1990.1	1970.1	1959.5	1961.9	1997.7
10	2028.8	2116.9	2074.1	2070.7	2091.4	2184.6
11	2113.1	2205.9	2176.5	2159.1	2179.6	2317.2
12	2268.7	2250.5	2235.9	2213.3	2240.1	2397.6
13	2151.9	2131.1	2200.1	2189.5	2220.7	2365.9
14	2352.5	2326.1	2320.6	2301.7	2352.8	2485.5
15	2352.3	2255.2	2322.3	2303.5	2351.2	2485.2
16	2363.4	2277.1	2341.7	2322.1	2374.5	2508.6
17	2501.8	2440.1	2468.6	2428.8	2457.3	2651.2
18	2355.1	2329.9	2328.5	2310.6	2329.9	2501.7
19	2241.5	2242.7	2225.9	2214.7	2214.1	2368.3
20	2059.4	2128.9	2120.5	2116.1	2116.7	2200.1
21	2067.9	2054.1	2123.5	2119.1	2124.2	2182.6
22	2014.1	2015.1	2012.2	2010.7	2012.5	2032.1
23	1853.5	1852.1	1891.9	1891.5	1893.3	1894.8
24	1707.6	1707.5	1706.4	1706.4	1704.9	1684.1
Total	46643.1	46641.9	46611.4	46409.7	45539	47077.5

Table V.13 Large Industrial Class Demands (Mwh)

Hours	Spot1	Spot2	Ramsey1	Ramsey2	RDP-2	RDP-3
1	90.6	90.6	86.7	86.7	88.5	88.5
2	89.3	89.3	83.1	83.1	85.9	86.3
3	88.2	88.2	80.7	80.7	83.9	84.5
4	88.1	88.1	87.3	87.3	89.3	89.3
5	88.7	88.7	85.5	85.5	88.8	88.9
6	91.3	91.3	90.6	90.6	94.2	95.1
7	98.3	98.3	98.1	98.1	99.5	99.8
8	109.3	109.3	108.9	108.9	109.1	109.1
9	114.1	114.5	114.5	114.4	114.4	114.3
10	115.6	116.1	116.1	116.1	115.9	115.7
11	115.8	116.5	116.8	116.7	116.4	115.9
12	116.2	116.1	116.5	116.4	116.1	115.6
13	116.6	115.4	116.4	116.3	116.1	115.6
14	119.1	118.8	119.2	119.1	118.8	118.6
15	117.8	117.3	118.2	118.1	117.8	117.6
16	113.2	112.8	113.6	113.5	113.2	113.1
17	107.3	107.1	107.9	107.7	107.3	107.1
18	101.9	101.8	102.5	102.4	102.1	101.8
19	99.8	99.8	101.2	101.1	100.8	100.5
20	98.5	98.8	100.1	100.1	99.5	99.5
21	97.6	97.5	99.5	99.5	98.9	99.0
22	98.4	98.4	99.3	99.3	98.9	99.1
23	95.6	95.6	96.1	96.1	95.9	96.1
24	93.6	93.6	92.5	92.5	92.9	92.8
Total	2464.9	2463.9	2451.3	2450.2	2464.2	2463.8

Table V.14 Small Industrial Class Demands (Mwh)

Hours	Spot1	Spot2	Ramsey1	Ramsey2	RDP-2	RDP-3
1	683.5	683.5	673.9	673.9	620.8	623.9
2	693.3	693.3	667.4	667.4	598.9	603.1
3	710.1	710.1	672.1	672.1	590.2	594.5
4	709.3	709.3	706.7	706.7	576.0	579.8
5	704.7	704.7	691.5	691.5	506.4	484.9
6	701.7	701.7	709.1	709.1	497.3	436.1
7	683.6	683.6	712.9	712.9	598.0	554.2
8	723.5	723.9	697.5	697.3	693.2	665.5
9	555.2	645.4	625.9	615.9	614.1	652.7
10	504.8	586.3	543.5	540.5	549.0	649.6
11	463.6	547.9	513.5	497.6	501.6	650.8
12	525.8	509.9	487.5	467.5	468.4	642.6
13	461.6	443.4	497.2	488.2	490.7	652.9
14	594.1	571.9	560.9	545.4	551.8	709.2
15	579.7	498.2	547.5	532.2	537.7	695.5
16	539.9	469.9	515.4	500.1	503.1	663.3
17	506.6	454.3	464.3	431.6	428.6	645.3
18	469.1	447.3	433.9	419.1	415.7	602.2
19	468.2	469.6	431.2	422.5	420.7	585.1
20	491.2	549.5	524.4	521.1	518.9	610.8
21	542.9	532.6	565.6	562.6	566.8	634.5
22	677.1	677.7	667.1	666.1	667.5	691.8
23	654.7	653.6	682.9	682.6	684.5	686.5
24	676.9	676.9	684.7	684.7	682.8	657.9
Total	14321.1	14344.5	14276.6	14108.6	13282.7	14972.7

Table V.15 Residential Class Demands (Mwh)

Hours	Spot1	Spot2	Ramsey1	Ramsey2	RDP-2	RDP-3
1	402	402	398	398	367	362
2	352	352	340	340	306	300
3	312	312	296	296	260	255
4	293	293	292	292	239	235
5	289	289	283	283	205	220
6	303	303	304	304	208	234
7	364	364	372	372	303	324
8	467	467	465	465	460	463
9	540	546	545	544	549	546
10	604	609	608	608	621	614
11	671	678	679	678	697	686
12	716	714	716	714	743	728
13	677	675	684	683	714	698
14	698	695	696	693	740	716
15	717	703	715	712	757	734
16	808	792	806	802	854	829
17	1019	1011	1021	1015	1051	1031
18	1053	1050	1056	1054	1080	1067
19	1011	1011	1024	1023	1025	1019
20	843	850	859	858	865	857
21	827	825	845	844	848	841
22	691	692	696	696	698	694
23	617	617	622	622	622	622
24	503	503	499	499	496	500
Total	14777	14753	14821	14795	14708	14575

Table V.16 Commercial Class Demands (Mwh)

hours	Spot1	Spot2	Ramsey1	Ramsey2	RDP-2	RDP-3
1	404	404	389	389	400	400
2	383	383	359	359	374	373
3	369	369	339	339	357	356
4	361	361	358	358	359	359
5	359	359	345	345	352	352
6	367	367	364	364	365	365
7	406	406	404	404	405	405
8	529	529	528	528	528	528
9	684	684	684	684	684	684
10	804	805	806	806	805	805
11	862	863	867	867	864	864
12	909	909	915	915	912	911
13	897	897	902	902	899	899
14	940	940	944	944	942	941
15	937	936	941	941	938	938
16	902	901	906	906	903	903
17	869	868	875	874	870	867
18	730	730	735	735	732	731
19	662	661	669	668	666	663
20	626	630	637	637	633	632
21	599	598	613	612	609	608
22	547	547	549	549	547	547
23	486	486	490	490	490	489
24	433	433	430	430	432	432
Total	15065	15066	15049	15046	15066	15052

Chapter VI

Summary, Conclusions, and Recommendations

VI.1 Summary and Conclusions

The objectives of this research were (a) to develop a price structure that unbundles electricity service into different levels of reliability, and (b) to analyze the implications of this structure for economic welfare, system operation, and load and energy management, in comparison with other traditional price structures.

The first objective was accomplished through the development of a reliability-based pricing scheme that integrates the following models:

(a) A customer choice model based on the notion that demands are derived from benefit maximization behavior, expressing customers' benefits as contingent upon the level of reliability with which they are served.

(b) A stochastic production cost simulation model, which replaces the traditional pre-estimated cost function as a supply model. It directly relates demands to system production, marginal costs, and levels of reliability. The integration of production cost simulation within the pricing model is an important contribution of this research. We also extend the production cost model to estimate the probability density function of system incremental (marginal) cost. This research has made clear that the use of the traditional deterministic total and marginal costs in most of the existing pricing models, is hindering the accuracy and the validity of these models. The need to estimate the probability density of system incremental cost and then the expected value of total and marginal costs, turns out to be critical.

(c) A welfare model that integrates the two previous models within the framework of Ramsey, real-time, and priority service, resulting in an integrated model that is an accurate, fast, and reliable tool for studying any power system and demand profiles.

The second research objective was accomplished by conducting a comprehensive comparison between the proposed reliability-pricing model and other traditional models. This comparison has produced the following conclusions:

(1) Reliability-based pricing, in the proposed form, is superior to spot, Ramsey, and average-cost pricing, unless the reliability level of the system is extremely high.

(2) When system reliability level is very high (near 100 percent), the need to unbundle electricity service vanishes. Traditional spot and Ramsey pricing then maintain their positions as the most and second most efficient pricing mechanisms.

(3) While reliability-based pricing proves to be superior, increasing the number of reliability options proves to decrease this advantage. In our formulation, the increased number of reliability options means more disaggregation of the revenue constraint, thus neutralizing the role of the inverse elasticity in allocating revenues.

VI.2 Recommendations for Further Research

The research presented in this dissertation could be extended along the following lines:

(1) Considering and incorporating transmission outages, in addition to generation outages, in reliability-based pricing. Outages due to transmission failures are very significant. Measuring transmission reliability is usually conducted using stochastic load flow and transient stability models. The integration of one of these models with a reliability-pricing model would be very useful.

(2) Using the history of a generating unit operations and failures when estimating its forced outage rate is extremely important. The assumption that FOR is fixed at all time, is inaccurate over the long term. Therefore, the use of a Markovian process to consider the time transition impact on any unit performance could enhance the model if applied to long term studies.

Appendix A

Production Cost Simulation

The evaluation of power system reliability and cost is of fundamental importance to any electric utility. Historically, production cost evaluation has usually been considered as an exercise separate from the evaluation of reliability, since the models and evaluation techniques were generally different. Using some form of probability evaluation and simulation techniques, electric utilities were able, more recently, to estimate both production cost and reliability simultaneously. This development is of importance to this research on reliability-based pricing. Considering both economic and reliability parameters simultaneously leads to an improved and more consistent decision-making process.

The general purpose of a modern production cost simulation model is to simulate the dispatch, operation, and availability of power systems. The typical function of a model is to estimate system production costs, requirements for energy imports, availability of energy for sales to other systems, fuel consumption, and system adequacy and reliability. Production cost simulation models are widely used as an aid in long-range system planning, fuel budgeting, and system operation. The production cost simulation model provides the link between the physical description of a utility (unit sizes, operating costs, maintenance requirements, forced outage rates) and its expected load. Therefore, one of the most important requirement for electric utilities is to have a production simulation model that is fast, accurate, and with low computational cost.

This section reviews current and new methods of reliability analysis and production cost simulation, and then explains in detail the method and the proposed extension used in this research. First, the most well-known reliability indices and measures are reviewed, including: reserve margin, loss of load probability, loss of energy probability, frequency and duration, and Markov chain analysis. Second, the most widely used production cost simulation techniques are described. The probabilistic simulation methods are presented in more depth, since they uniquely allow for the simultaneous estimation of both production cost and reliability. A proposed extension of the stochastic production cost simulation method employs the cumulant technique to allow for the estimation of the probability density function of system lambda (system incremental cost). The probability density function of system lambda is useful in estimating the expected value of the system marginal cost. The expected value of system marginal cost is used in this research instead of the traditional deterministic marginal cost. The proposed extension is discussed at the end of this appendix.

A.1. Commonly Used Reliability Measures

A.1.1 Reserve Margin

The reserve margin of a utility system is perhaps the most common measure of system reliability. By definition, it represents the amount of extra capacity a utility has installed over the peak demand it faces within a time frame.

Reserve Margin (%) = (total installed capacity - peak demand)/ peak demand

The typical time frame for reserve margin is annual or seasonal. The total installed capacity then represents all units available for service within that period.

Reserve margin is a quick and easy measure of system reliability. Its calculation requires a minimum of information. In the past, maintaining a selected reserve margin insured a desirable level of reliability.

Reserve margin does not consider shortages at all. It is an index which, without knowledge of the size and number of units in the system, past history, and the distribution of demand, does not provide much information about system reliability. With such knowledge, one still could not say much about the occurrence of shortages in the future. It is not even clear that past shortages can be related to reserve margins. Since it considers only one point in time (peak demand), it is not possible to logically conclude that a shortage at any other time might have been prevented by a higher reserve margin.

A.1.2. Loss of Load Probability and Loss of Energy Probability

Loss of load probability (LOLP) is a common index of system reliability. It estimates the amount of time, within a time frame, during which the system capacity would be unable to meet a given demand. It is defined as follows:

$$LOLP = \sum_{i=1}^{n} P(systemcapacity < d_i),$$

where n is the number of discrete time periods (usually hours) in the time frame being studied (usually one year) and d_i is the forecast demand in time period i. If each "hourly" probability is interpreted as the fraction of the hour the system cannot meet demand, then LOLP represents the total amount of time during the time frame in which capacity is less than demand. In practical applications, a reduced set of hours is sometimes used to shorten computation time. For example, only peak weekday hours might be looked at, assuming the rest of the week has a negligible contribution to LOLP.

The probability distribution of capacity out of service is computed using unit forced outage rates and maintenance scheduling. For each time period (or hour), units scheduled for maintenance are removed from the set of all units in the system, leaving n units available for service. A discrete probability mass function of capacity out of service is computed as follows. Each unit is assumed to be out of service with a probability equal to its forced outage rate, and in service with a probability equal to (1-forced outage rate). There are 2^n combinations of the n units being in or out of service. Each combination represents an outage of "x" MW where "x" is the sum of capacities of units in outage in that state. The probability of being in that state is equal to the product of each unit's probability of being either in or out of service. A cumulative distribution function is then computed from this probability mass function.

There are refinements to this basic methodology which improves its accuracy. The definition of generating units can be broadened to include firm purchased power, emergency purchased power, or other emergency actions. For each source of power, a capacity and reliability are needed. Partial outages can also be included. In this case, the number of combinations of outage states will increase to $\prod_{i=1}^{n} K_i$, where K_i is the number of capacity states for unit i ($K=2$ for no partial outages). Demand also can be treated as a stochastic variable. The demand d_i can represent a distribution of demands. LOLP for each hour is computed for each possible demand in the distribution and weighted by the probability of occurrence of each demand state.

Loss of load probability is a useful reliability measure for utilities. Its use as a probability, however, is incorrect. LOLP is the sum of expected values and, as such, estimates only the total outage time which can be expected within a time frame. This information is important to planners, who are concerned with minimizing total outage time. However, as a shortage measure, LOLP tells only one side of the story. Knowing the expected total outage time says nothing about the characteristics of the shortages which are expected to occur. The method for calculating LOLP can be modified to calculate loss of energy probability, which provides more information on shortages.

The loss of energy probability measure (LOEP) has not gained wide acceptance until recently. Utilities that are able to calculate their LOLP should be able to compute LOEP. In the calculation of the LOLP, the probability of an outage for hour (j) is equal to the probability that system capacity y will be less than load d_j. This can be expressed as:

$$P\ (capacity < load\ d_j) = \sum_{i=1}^{M} P(capacity = y_i)$$

where M represents a subset of all possible capacity states such that $y_i < d_j$. The loss of energy probability weighs each of these states by the amount of energy shortage. LOEP is calculated for a single hour j using this formula:

$$LOEP_j = [\sum_{i=1}^{M} P(capacity = y_i) * (d_j - y_i)]$$

The LOEP for all hours in the time frame being studied are then summed. The result is the expected amount energy not served. Like LOLP, the LOEP is a sum of expected values, rather than a probability.

The use of LOEP as a reliability measure or planning tool is not common since minimizing unserved energy is only secondary to the problem of minimizing outages of any size. However, the index is gaining more recognition due to the increase need for energy exchange. The index also shares many of the problems of LOLP, in that it does not provide any notion of the frequency or duration of a single outage. Rather, it gives a total figure over a long period of time. It is an improvement over LOLP since it does give an indication of the magnitude of the outage which LOLP forecasts.

A.1.3. Frequency and Duration Method

The frequency and duration algorithm is a comprehensive method used to measure system reliability. It provides estimates of the expected time between outages and the expected duration of outages. More system data are required for computation, specifically mean repair times for each plant and duration of peak load. The algorithm was developed by Ringlee and Wood (1968,1969).

The results of the frequency and duration method are average (or expected) values. Used as a reliability measure, this information is useful in judging the exposure of a utility to an outage. Knowing the expected duration of outage is especially useful, enabling utilities to prepare in advance procedures for coping with such a problem. The main problem in using this method is the availability of data needed for the calculation. Data on mean repair time and duration of peak load are needed. Mean repair time data are available, but are not very good statistically. Realistically, it is not correct to apply a mean repair time based on data from many different units to a specific unit. A specific unit's mean repair time can be significantly different from the mean repair time of similar units. However, there may not be enough valid repair time information for a specific unit to calculate the mean.

As shortage measures, the frequency and duration indices have the same limitation as LOLP and LOEP; namely, they provide average value information.

Taken together, these three measures give a good picture of the average annual outage a utility might expect.

A.1.4. Markov Chain Method

As discussed above, the frequency and duration method does not result in a probability distribution of shortages. Ideally, this distribution is the most useful information for describing a system's exposure to shortages. The Markov Chain method lays a foundation for calculating a probability distribution of shortages and illustrates the type of information which can be extracted from this type of analysis.

For illustration, consider the probability distribution of capacity outages within some time frame. This coincides with the first step of the frequency and duration method, which is also applied independently of load. The distribution is discrete, owing to the discrete nature of unit capacity and to the convenience of representing time in increments of any desired period. The utility system is represented by n units, each with a capacity of c_i, forced outage rate f_i, and repair rate r_i. These parameters are assumed invariant over the time frame T being studied.

At any step in time, the system can be in state s(t), where s(t) represents the total capacity out of service at time t. Let a(t) be the set of units out of service and b(t) be the units in service.

$$s(t) = \sum_{i=1}^{n} c_i, \qquad i \in a(t)$$

The maximum number of states the system can be in is m, which represents the number of unique totals of unit capacities out of service.

In going from time t to time t+1, the following four transitions are possible:

i) Units in service can remain in service with probability $1-f_i$,

ii) Units in service can fail with probability f_i,

iii) Units out of service can be repaired with probability r_i,

iv) Units out of service can remain so with probability $1-r_i$.

This representation assumes that failure and repair times are binomially distributed, as is assumed in LOLP analysis. LOLP assumes that capacities between time steps are independent, and that assumption does not carry over here.

Given this information, we can now derive some pertinent probability distributions, namely: 1) $P(s,t \mid s = s^{*}, t=0)$, the probability of being in outage state s at time t; 2) $P(s > s^{*}, t \mid s=0, t=o)$, the probability of an outage greater than s^{*} after t periods, given all units were initially in service; and 3) $P(s=0, t \mid s=s^{*}, t=0)$, the probability that the system will have no outages at time t, given it was initially in outage state s^{*}.

The first of these distributions gives basic information about the system. The second and third distributions are used to compute expected values of time for failure and repair, which can be used to compute frequency and duration of outages.

The stochastic process described above has the necessary properties to be classified as a Markov process, since the state at time t+1 depends only on the state at time t, not on how the system arrived at that state. The process can be described by the mxm matrix Q. Each element of Q is the probability P_{ij} of going from s(t) = i to s(t+1) = j.

$$P_{ij} = P\{s(t+1) = j \mid s(t) = i\}$$

$$= \prod_{t=1}^{n} (1-f_i \mid i \in b(t) \text{ and } i \in b(t+1)) \quad \text{or}$$

$$= (f_i \mid i \in b(t) \text{ and } i \in a(t+1)) \quad \text{or}$$

$$= (r_i \mid i \in a(t) \text{ and } i \in b(t+1)) \quad \text{or}$$

$$= (1 - r_i \mid i \in a(t) \text{ and } i \in a(t+1)),$$

where:

a(t) = the set of units out of service at time t

b(t) = the set of units in service at time t

Using this transition matrix, we can find P_{ij}^{t}, the probability of going from state i to state j in t time periods. This probability is the element ij of matrix Q^t, where

$$Q^t = \prod_{r=1}^{t} Q$$

Given some initial state, such as $s(t=0)=0$, we can find the probability of being in any state at any point in time, or $P(s,t/s = s^{*}, t=0)$.

In sufficient time, it can be shown that the probability of being in any state s' will be independent of the initial conditions (the Markov Chain has the property that all states can be reached from any other state, and that no state is a trapping state [P_{ii} =1], and that there are no closed groups of states). These stationary or equilibrium probabilities P_{ij} are easily calculated by solving a system of m+1 linear equations. These probabilities have an important interpretation. They represent the average fraction of time the system will be in state i.

Up to this point, the above analysis had left aside all considerations of load. Theoretically, it would be possible to expand the scope of the transition matrix where each state would represent reserve, rather than capacity. Transition probabilities would reflect changes in load as well as capacity. Probability distributions of reserve can be calculated in a similar fashion. The transition matrices can be made a function of time, where maintenance, partial outages, and seasonal load distributions can be incorporated.

Practically speaking, these methods might be computationaly infeasible. In a small problem, with only three units and eight outage states, it was estimated [EPRI, 1982] that over 200,000 iterations would be needed to calculate the distribution to a cumulative probability of 0.9995.

A.2. Commonly Used Production Cost Simulation Methods

A.2.1. Monte Carlo Simulation

In this method, the simulation of power dispatch proceeds in chronological order. Suppose an hour is the unit time length of simulation. In order to determine which generation units are on forced outage, a random number is generated for each generating unit for an hour period. If the random number drawn is less than the forced outage probability of the unit under consideration, then that unit is decided to be on forced outage during that hour. If the probability distribution of forced outage duration is also known, the time of occurrence and duration of forced outage may be both decided by random numbers [S. Nakamura,1984].

Once it is determined which generating units are on forced outage in the hour, the dispatch simulation for the hour becomes deterministic; thus the simulation becomes simple and able to deal with more detailed aspects than the probabilistic simulation method. The chronological order of simulation makes it possible to consider: (a) start-up of cycling units; (b) an accurate simulation of the pump storage unit(s); (c) unit commitment rules; (d) spinning reserve; and (e) effect of inter-tie flow. The disadvantages of this approach are that the results are subject to variance (the result from different runs are always slightly different even with the same input data) and that there is an error due to the pseudo-randomness of the random numbers (the random numbers generated by a computer are not truly random).

A.2.2.Derating Method:

In the derating method [S. Nakamura, 1984], the equivalent capacity is defined as the capacity of a unit times (1-P), where P is the forced outage rate, thus incorporating the effects of forced outage. The generating units with equivalent capacity are dispatched deterministically, so the computational time required for this operation is very short. The disadvantage is, however, that the energy generated by peaking units tends to be severely underestimated whereas the energy generated by cycling (intermediate) units tends to be overestimated. Furthermore, the loss of load probability (LOLP) cannot be calculated by this approach. The derating method is often applied in combination with the probabilistic simulation method for the purpose of reducing the computational cost of probabilistic simulation without causing a significant reduction in accuracy. In this case, small peaking units are treated by the derating method while all other units are convoluted (summed) probabilistically.

A.2.3.Probabilistic Methods

An important advance in production cost simulation was the introduction by Baleriaux of a technique to account for the random nature of load and of generating units outages. The method was further refined by Booth (1972). The method rests heavily on obtaining a load duration curve (LDC) and the corresponding load distribution function. By considering the outage of generating units as part of the demand (as they are a burden on the system as much as the demand is), the notion of equivalent demand is defined. This equivalent load may be viewed as an augmented load caused by the random outages of generating units. Appropriate areas under the probability distributions of demand are used to obtain expected unit energy generation. Units are loaded according to a merit order decided upon their average incremental cost. The equivalent load is obtained by a convolution formula given in terms of a recursive algorithm. The basic concept of this method is illustrated as follows.

(i) Capacity and Demand as Random Variables

The load demand is usually specified by the "Load Probability Function", L(x). This is also known as the "Inverted Load Duration Curve" and "Complementary Distribution Function". L(x) is defined as the probability that a random load x will equal or exceed a demand level x (MW), with:

$$L(x)=P(\bar{x}\geq x) \tag{A.1}$$

The load frequency (density) function, l(x), is defined by

$$l(x)\Delta x=P(x<\bar{x}\leq x+\Delta x) \tag{A.2}$$

where $l(x)\Delta x$ is the probability that a random load of x MW takes a value between x MW and $(x + \Delta x)$ MW. From equations (A.1) and (A.2), it is evident that:

$$L(x)=\int_x^{\infty} l(u)du \tag{A.3}$$

Typical load probability and load frequency curves are shown in figure A.1 and A.2

The forced outages of a generating unit are specified with the outage capacity frequency function, q(C), as

$$q(C)dC=P(C<\bar{c}\leq C+dC) \tag{A.4}$$

Where the q(C) dC is the probability that a random outage of c MW will be between the unit's generation levels C and (C +dC) MW. If C is the unit's rated capacity then

$$\int_0^C q(u)du=1 \tag{A.5}$$

The unit's availability probability P is

$$P=\int_0^C q(u)\delta(u-0)du \tag{A.6}$$

Where δ is the Dirac delta function ($\delta(x)=1$ for $x=0$, $\delta(x)=0$ for $x\neq0$).

The probability, q, that the unit is unable to provide any power is

$$q=\int_0^C q(u)\delta(u-C)du \tag{A.7}$$

In many instances generating units are represented with two states: (a) available and capable of full power generation and (b) on total forced outage. In this representation q and P are the unit's forced outage and availability probability and equation (A.5) reduces to

$$P + q = 1 \tag{A.8}$$

A typical two state discrete outage capacity frequency function is shown in figure A.4 where P and q are represented with two spikes, at zero and full capacity respectively.

(ii) Equivalent Load Demand

Assume that the generating system consists of N generating units with loading order 1,2,3,....,n,...,N, with capacities $C_1,C_2,....,C_n,...,C_N$, and with availability probabilities $P_1,P_2,...,P_n,...,P_N$. Let the load probability function be L(x), as depicted in figure A.3. If the simulation periods is T hours, the expected energy generation E_1, of the first unit is (as depicted by the shaded area in figure A.3:

$$E_1 = T.P_1 \int_0^C L(x)dx \tag{A.9}$$

In order to calculate the expected energy generation of the second unit, we must realize that the second unit will move in the loading order to replace the whole or part of the first unit, when the latter is totally or partially on forced outage. This means that the load demand for the second unit will be higher when the first unit is on forced outage than otherwise. Therefore, before calculating the energy of the second unit, we should increase the load demand by the amount of outage of the first unit. The new load demand is called Equivalent Load Demand and its load probability function, $EL_1(x)$, is defined as the probability that a random equivalent load of x_1 MW equal or exceed a demand level of x MW

$$EL_1(x)=P(\bar{x}_1>x), \tag{A.10}$$

where x_1 is the sum of random load demand x and outage capacity C_1 of the first unit:

$$\bar{x}_1=\bar{x}+\bar{C}_1 \tag{A.11}$$

Let $l_1(x,c)$ be the combined frequency function of x and C_1, i.e.

$$l_1(x,C)=P(x<\bar{x}\leq x+dx\ ,\ C<\bar{C}_1\leq C+dC) \tag{A.12}$$

Assume that x and C_1 are independent random variables. Thus,

$$EL_1(x)=P(\bar{x}+\bar{C}_1>X)\int_{-\infty}^{\infty}\int_{X-C}^{\infty}l_1(x,c)dxdc \tag{A.13}$$

However, the necessary and sufficient condition that random variables x and C_1 are independent is [Cramer, 1974]:

$$l_1(x,C)=l_1(x)q_1(C) \qquad \text{(A.14)}$$

Where

l(x,C) is the combined frequency function of x and C_1

l(x) is the frequency function of x

$q_1(C)$ is the frequency function of C_1

x is the load (MW), and

C is the outage capacity of unit one in MW

Substituting equations (A.15) and (A.3) in (A.14) yields

$$EL_1(x)=\int_{-\infty}^{\infty}q_1(C)dC\int_{X-C}^{\infty}l(x)dx=\int_{-\infty}^{\infty}q_1(C)L(x-C)dC \qquad \text{(A.15)}$$

Since $q_1(C) = 0$ for $C<0$ and $C>C_1$, where C_1 is the maximum capacity of the first unit,

$$EL_1(x)=\int_0^C L(x-C)q_1(C)dC \qquad \text{(A.16)}$$

which is the familiar convolution equation of random variables x and C_1 [Cramer, 1974].

When $q_1(c)$ is discrete, equation (16) becomes:

$$EL_1(x)=P_1(x)+\sum_{j=1}^{J} L(x-C_j)q_{1j}, \tag{A.17}$$

where q_{1j} is the discretized outage capacity frequency function for the first unit defined by:

$$q_{1j} = q_1(C_j), \qquad j=1,2,...,J$$

In the two state representation, we have

$$EL_1(x)=P_1L(x)+q_1L(x-C_1) \tag{A.18}$$

This form of the convolution has been suggested by Baleriaux [1959] and later used by Booth (1972).

The equivalent load probability function of the n-th unit $EL_n(x)$, that results from the convolution of the first n generating units with the system load is defined as:

$$EL_n(x)=P(x_n>x), \tag{A.19}$$

where:

$$x_n=x+\sum_{i=1}^{n} C_i \tag{A.20}$$

$EL_n(x)$ is found by the recursive application of equation (A.16) i.e.

$$EL_n(x)=\int_0^C EL_{n-1}(x-C)q_n(C)dC \tag{A.21}$$

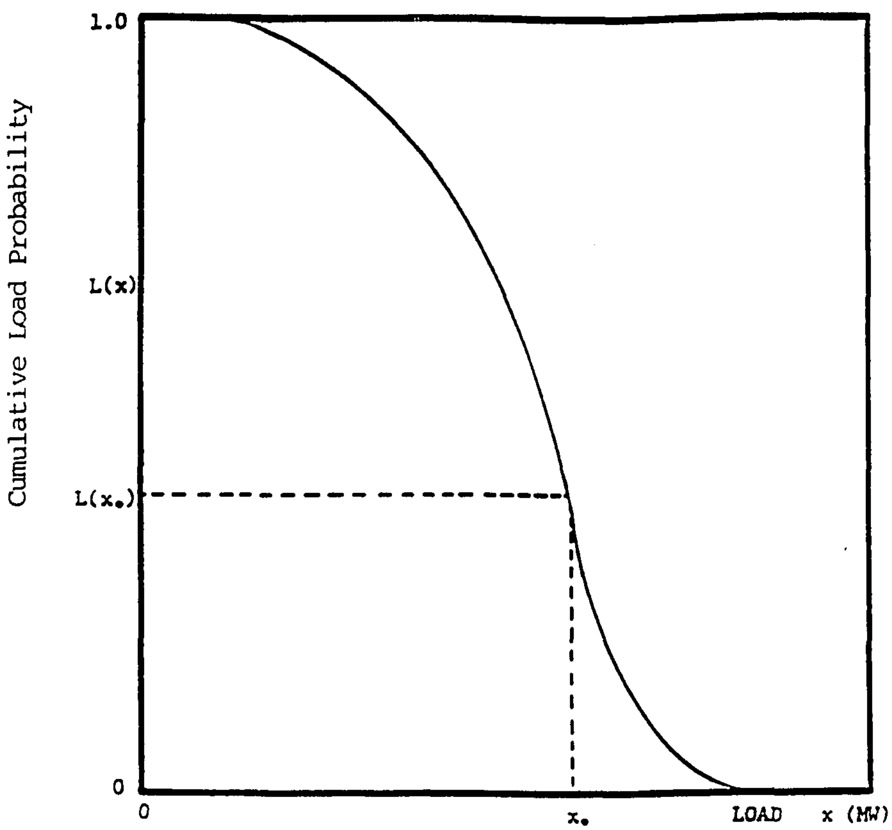


Figure A.1: Typical Load Probability Curve

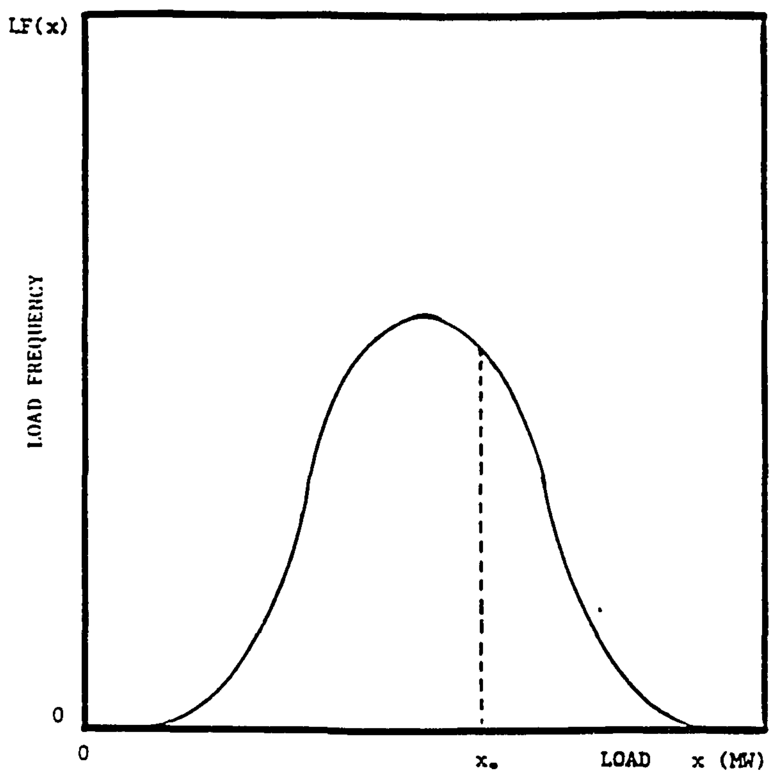


Figure A.2: Typical Load Frequency Curve

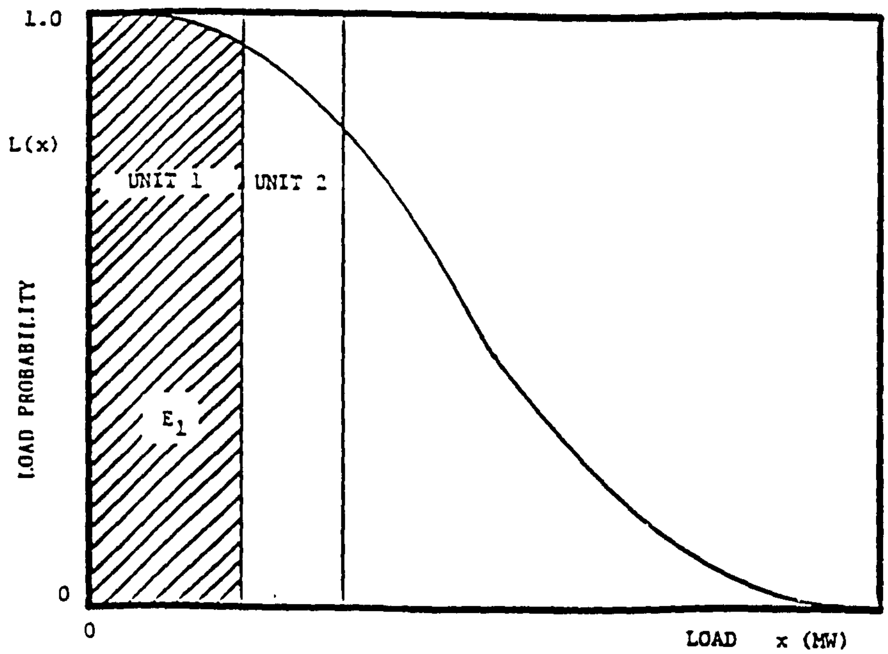


Figure A.3: Load Probability Curve. The Shaded Area Depicts the Expected Energy Generation by the First Generating Unit

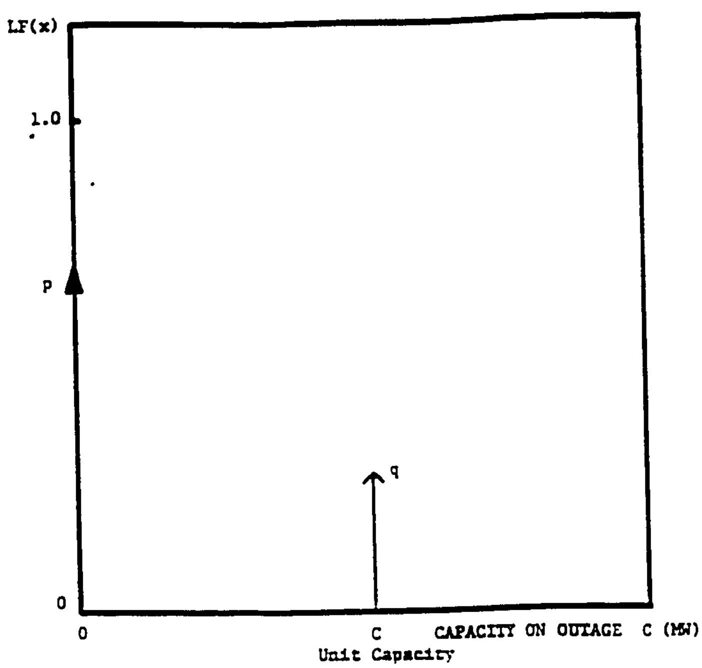


Figure A.4: Two-State Discrete Outage Frequency Function.

Since the sum of random variables is cumulative, the order with which units 1 through n are convolved is unimportant. Therefore, if we know the equivalent load probability function $EL_n(x)$ of the first n units and we want to find the equivalent load probability function, $EL_k(x)$, of n-1 units that does not incorporate the effect of outages of any unit k, (k<n), we must solve the equation:

$$EL_n(x) = \int_0^{C_k} EL_{n-1}(x-C) q_k(C) dC \tag{A.22}$$

Where:

$El_{n-1}(x)$ is the equivalent load probability function of units 1,2,...,k-1,k+1,...,n.

$q_k(C)$ is the k^{th} units outage capacity frequency function, and

C_k is the k^{th} units capacity, in MW

The process is called deconvolution and is used when generating units are represented with more than one capacity blocks.

(iii) Evaluation of the LOLP and the Unserved Energy

The power system's reliability and security of supply levels provided by a given configuration are usually represented in terms of two quantities: the Loss-Of-Load Probability (LOLP) and the amount of Energy Not Served (UE). The LOLP can be defined as the percentage of time (referred to as the total period of time considered) during which the system load exceeds the available generating capacity of the system.

The Energy Not Served is the amount of energy required by the system and which cannot be supplied by the generating equipment existing in the system.

The LOLP is evaluated as the average value of the function $ELDC_n$ (the final equivalent load duration curve) at an ordinate equal to the total system generating capacity on the x MW axis, as shown in figure A.5.

$$LOLP=\frac{1}{\Delta x}\int_{\sum_i {}^n c_i-\Delta x}^{\sum_i {}^n c_i+\Delta x} ELDC_n(X)dX \tag{A.23}$$

The energy not served (UE) is evaluated as:

$$UE=T.\int_{\sum_i {}^n c_i}^{\infty} ELDC_n(X)dX, \tag{A.24}$$

where:

T = total length of the period covered by the LDC (hours).

The load probability function or load duration curve (LDC), and the equivalent load probability function (equivalent load duration curve ELDC) are shown in figure A.5.

The advantage of the probabilistic simulation arises from the fact that the load for a long period, such as month, quarter or even a year, may be handled in a single load probability function regardless of the length of the period. Besides, this approach does not suffer from variance, as does the Monte Carlo method, which needs repetitive use of dispatch simulation.

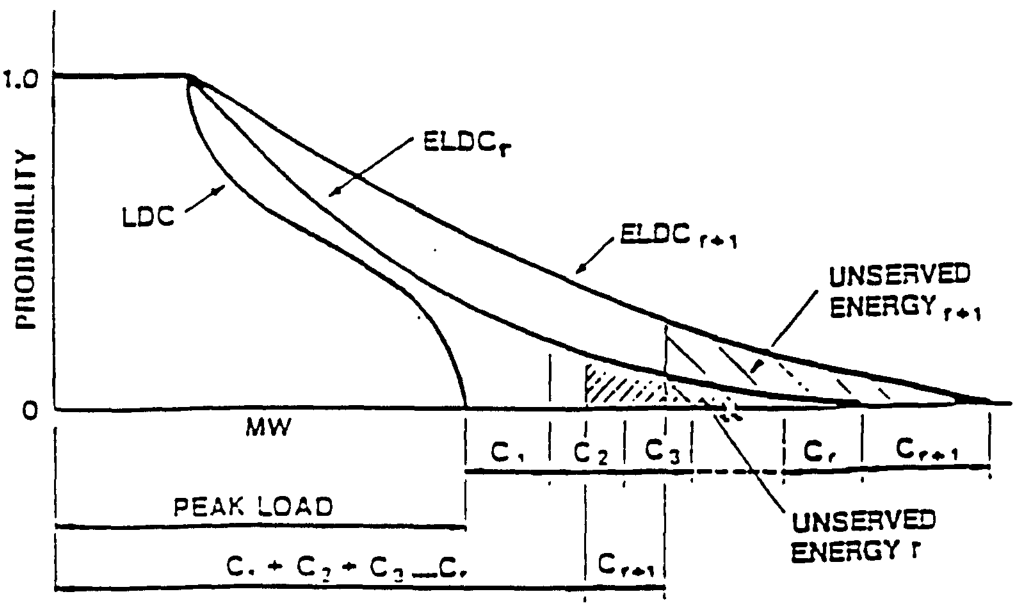


Figure A.5: Final Equivalent Load Duration Curve.

The computation of costs and the accuracy of the probabilistic simulation are affected significantly by the choice of the numerical algorithm adopted. The following methods are most frequently used:

-Booth-Baleriaux method.

-Fourier Series Expansion.

-Piecewise Linear Approximation.

-Cumulant Approximation.

Booth Method

Booth [1972] and others have used a numerical representation in which several hundred rectangular areas represent the equivalent load curve. These are represented in the computer as a one dimension array with the base of each rectangle representing a constant megawatt value step size. Convolving the effect of forced outages of a generation unit into the load curve requires the recursive recalculation of every value as does the deconvolution of a unit. If a unit sizes are not integral multiples of step size, interpolation is required. Booth's interpolation is simple and ingenious, but it is an approximation. As units are convolved, the number of steps to be calculated increases. This leads to an exponential effort as the power system grows. Each calculation introduces inherent digital truncation errors, so that, for very large systems, recursive calculations eventually lead to intolerable errors. Double precision reduces this problem but does not eliminate it. Reducing the step size, which improves representation of the inverted LDC and reduces interpolation errors, actually causes digital truncation errors to increase because of the increased number of recursive calculations.

Fourier Transform Method

J. P. Stremel and R. T. Jenkines [1980] use another method they borrowed from the notion of electric wave form of analysis. They fit inverted LDC by a quarter Cycle of Fourier series. Figure A.6 depicts the fitting process that converts the load duration function into a Fourier expansion series.

By assuming that the inverted LDC function in figure A.6.d is a quarter cycle of a wave function, one can express the LDC in the following Fourier form:

$$F(X) = \frac{a_o}{2} + a_1\cos X + b_1\sin X + \ldots + a_n\cos X + b_n\sin X + \ldots$$

where a's and b's are the Fourier coefficients expressed as:

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} F(X)\cos nX dX,$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} F(X)\sin nX dX,$$

$$n = 1,2,\ldots$$

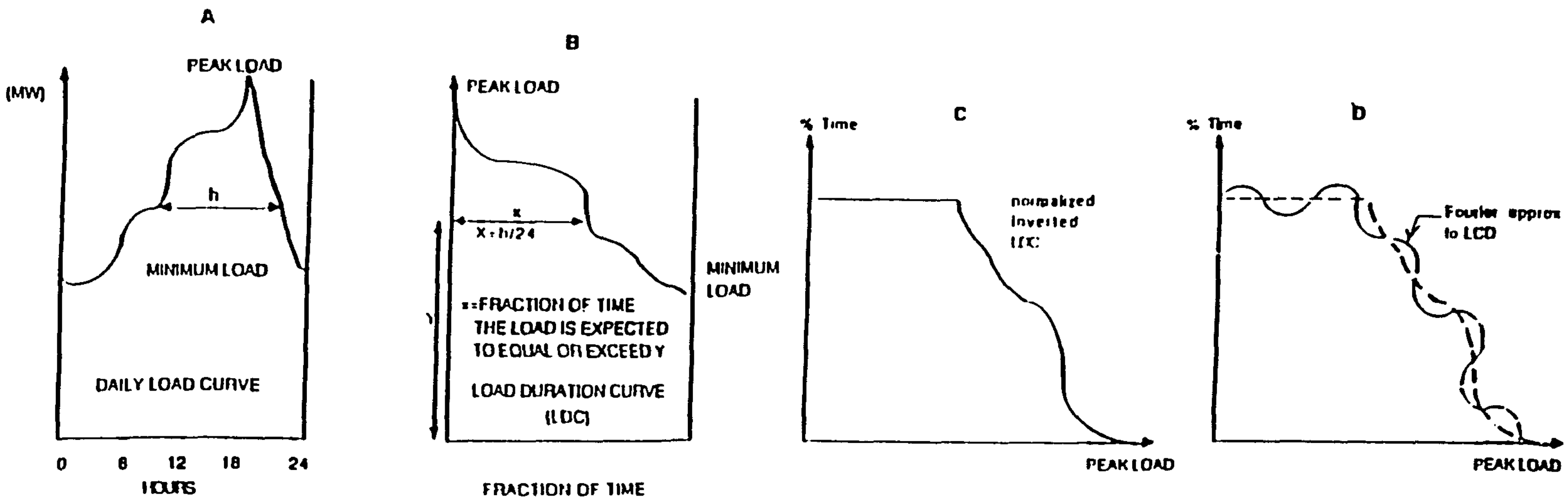


Figure A.6: Conversion of Hourly Load Curve to Normalized Inverted LDC Using Fourier Series Approximation

With the Fourier Series Expansion, the equivalent load probability functions are developed through convolutions and deconvolutions. The convolution and deconvolution are performed by simple relations among the Fourier coefficients of only the same Fourier mode. This process has the advantage of being constant in computational effort regardless of system size and it alleviates the inherent truncation errors of the numerical method. The accuracy of computing the energy generated by each unit and the LOLP is good except when the LOLP becomes extremely small (less the 10E-04). This method has been used in the WASP-package and other TVA software packages [Nakamura and Brown, 1977].

Piecewise Linear Approximation Method

This method was proposed by S. Nakamura and S. Tzemos [1981] to calculate LOLP more accurately than by the Fourier Series Expansion method. In this method, the load probability function, L(x), is represented with a piecewise linear polynomial. The load axis x is divided in grid points x_i,

$$x_i = i.\Delta x, \qquad i=1,2,3...,I,$$

where Δx is the interval between two consecutive grid points, and i is the grid number. Given the values $L_i = L(x_i)$ of the load probability function for all the grid points, the approximate representation of L(x) is given by $L_a(x)$:

$$L_a(x)=\frac{x_{i+1}-x}{\Delta x}L_i+\frac{x-x_i}{\Delta x}L_{i+1} \qquad \text{(A.25)}$$

The value of L(x) at the interval between grid points is found by linear interpolation between the grid values. The representation of the equivalent load probability function EL(x) is similar to that of L(x).

The discrete forms of the convolution and deconvolution equations are:

$$EL_n(x_i)=\sum_{j=0}^{J} EL_{n-i}(X_i-C_j)q_j \tag{A.26}$$

$$EL_{n-1}=\frac{1}{p}[EL_n(x_i)-\sum_{j=1}^{J} EL_{n-1}(x_i-C_j)q_j] \tag{A.27}$$

The values of $EL(x_i - c_j)$ between the grid points EL_i - $EL(x_i)$ are approximated by $L_a(x)$ from equation (A.25). Both equations (A.26) and (A.27) are applied recursively for i=1 to i=I, where I is the total number of grid points. In equation (A.28) we set:

$$EL_{n-1}(x_i-C_j)=1, \tag{A.28}$$

for any

$$x_i-C_j\leq 0$$

The computational cost of the piecewise linear method is proportional to the number of convolution and deconvolution operations in each application. Calculating only LOLP, without calculating the energy generated, does not require a large number of convolutions and deconvolutions, and, accordingly, the computing cost is much smaller than when energy calculation is required. Therefore, the linear approximation method may be used only for LOLP calculations, while the energy calculations are performed with the Fourier expansion method or the cumulant method.

Cumulant Method

Cramer [1974] shows that there are several methods of developing an orthogonal expansion derived from normal distributions in order to represent an arbitrary frequency function. One of these expansions in orthogonal polynomials became known as the Gram-Charlier series of type A, while another was developed by Edgeworth from an entirely different approach through the theory of elementary errors. Both expansions have been used in production costing models.

Rau et al. [1980] have introduced the Cumulant method using the Gram-Charlier expansion to represent the probability distributions of load and capacity outages. With the cumulant method, load probability functions and forced outage probability distributions are all expressed in the form of low order cumulants.

Cumulants are linear combinations of statistical moments. They exhibit two highly desirable characteristics [Stremel, 1980]: first, they can be used to describe the probability of total and partial outages of each dispatchable operating level of each unit (there is no requirement to assume discrete step sizes or identical units); second, a random variable (such as system outage) which is the sum of independent random variables, is characterized by cumulants. The second property translates to simply summing the cumulants determined for each generating unit in order to develop the corresponding cumulants for the system. Thus, convolutions and deconvolutions are performed by addition and subtraction of cumulants. The number of cumulants that are subject to addition or subtraction at each convolution or deconvolution is far smaller than the number of the Fourier coefficients.

Moments are expected values of probabilistic variables (or functions of these variables). Cumulants are functions of moments. For example, consider a random variable X, having a probability density function (PDF) f(X). The i-th moments about a constant, C, is defined as :

$$E[(X-C)^i] = \sum (X_i-C)^i f(X_i) \qquad \textit{(discrete)}$$

$$= \int (X-C)^i f(X)dX \qquad \textit{(continous)}$$

The cumulants of a distribution are defined in terms of the first moment about the origin (the mean) and the higher moments about the mean. For the random variable X having the PDF f(X), the expected value of X is :

$$M_1 = E(X) = \sum X_i f(X_i) = \bar{X}, \qquad \textit{(discrete)}$$

$$= \int X f(X)dX, \qquad \textit{(continous)}$$

Higher moments about the mean, x ,are calculated as:

$$M_i = E[(X-x)^i] = \sum (X-x)^i f(X_i), \qquad (discrete)$$

$$= \int (X-x)^i f(X) dX, \qquad (continous)$$

where M_i is the i-th moments about the mean.

The first six cumulants are defined as:

$K_1 = M_1$

$K_2 = M_2$ (variance = sigma^2)

$K_3 = M_3$

$K_4 = M_4 - 3(M_2^2)$

$K_5 = M_5 - 10(M_2{*}M_3)$

$K_6 = M_6 - 15(M_2{*}M_4) - 10(M_3^2)$

These relations are derived in detail in Kendal and Stuart [1977].

In essence, the higher cumulants, K_3 through K_6, of a distribution with PDF f(X), measure the departure of f(X) from a normal (Gaussian) distribution having mean K_1 and variance K_2.

There are two basic assumptions required for a probabilistic simulation: (a) outage occurrences are independent of the load; (b) outage occurrence are independent of other outages. With this assumptions, it has been shown [Kendall and Stuart, 1977] that the cumulants of the sum of independent random variables are equal to the sum of the cumulants of these random variables. Since load and outages are random variables, the equivalent load is the sum of the load and all the outages up to the point of interest on the partial or complete equivalent load curve. The convolution of forced outages into the equivalent load curve consists of simply adding the appropriate outage cumulants, order for order, to the cumulants of the "load" or "partial equivalent load" curve. The important property of this operation is that it is equivalent to the application of the laborious convolution equation:

$$\acute{F} = P\ F(X) + q\ F(X-C), \qquad where \qquad P+q=1$$

In the Fourier transformation approximation, the above equation must be applied to every Fourier coefficients. Thus, a series with 100 coefficients would require several hundreds of multiplications. With cumulants, convolution becomes a trivial operation.

To describe the cumulant relationship, it is necessary to follow the next three steps to calculate the equivalent load curve ordinate at a chosen MW value X.

Step 1 Calculate the deviation of the chosen X from the mean: (X-x) and calculate the standard deviation "sigma".

Step 2 Standardize the deviation from the mean in terms of standard deviation, i.e., calculate Z = (X-x)/sigma (Z is commonly referred to as standard variate).

Step 3 Standardize the cumulants K_3 to K_6 of the equivalent load curve ELC(X):

$G_1 = K_3 / \text{sigma}^3$

$G_2 = K_4 / \text{sigma}^4$

$G_3 = K_5 / \text{sigma}^5$

$G_4 = K_6 / \text{sigma}^6$

Then, the probability density function f(Z) can be expressed in Gram-Charlier series type A as follows:

$$f(Z)=N(Z)-\frac{G_1}{6}N^3(Z)+\frac{G_2}{24}N^4(Z)+\frac{G_1^{\,2}}{72}N^6(Z)+...,$$

where Z is a standardized variable represent load plus unit outage with mean zero and variance one; N(Z) is the normal distribution function and $N^i(Z)$ is the i-th derivative of N(Z).

Since most calculations for probabilistic simulation use ELC rather than PDF f(Z), we can derive ELC from f(Z) as follows:

$$ELC(Z)=1-\int_{-\infty}^{Z}f(X)dX$$

$$=\int_{Z}^{\infty}N(X)dX+\frac{G_1}{6}N^2(Z)-\frac{G_2}{24}N^3(Z)-\frac{G_1^{\,2}}{72}N^5(Z)+..$$

$$=\int_{Z}^{\infty}N(X)dX+\frac{G_1}{6}N^2(Z)-\frac{G_2}{24}N^3(Z)-\frac{G_1}{72}N^5(Z)+..$$

This equation is used repeatedly in probabilistic simulation to calculate unit expected generation, as follows {figure A.7}:

This implies

$$E_i = P_i \int_A^{A+C_i} ELC(X)dX$$

$$E_i = P_i \int_A^{\infty} ELC_{i-1}(X)dX - P_i \int_{A+C_i}^{\infty} ELC_{i-1}(X)dX$$

Note that the first term in the right hand side of the equation represents the unserved energy after convolving (adding) i-1 units; and the second term represents the unserved energy after convolving i units. Also, from figure {A.7}, note that S represent the loss of load probability LOLP after convolving i units. The same analogy is used in this dissertation to compute the probability that unit (i) is the marginal unit (the unit that cover the whole residual demand of ELC(X)) as the difference between S_i and S_{i-1}, as discussed in section 4.2.

The computational cost with the cumulant method is approximately one order of magnitude smaller than the cost with the Fourier expansion method. However, the method may lead to poor accuracy in LOLP and energy calculations for peak generating units when the LOLP is small or when the number of units in the generating system is small. Multimodel load shapes will also cause difficulties in approximating the load distribution. As a result, negative values for LOLP as well as the expected energy generation for the peaking units may occur.

Stremel et al [1980] have considered a variation of the cumulant method which begins with the chronological demand curve and obtains a probability density function PDF of demand and the corresponding cumulants. The moments and cumulants of the load PDF are obtained directly by sampling the chronological demand every hour and assigning to each sample equal probability. However, both variations of the cumulant method are as computationally efficient.

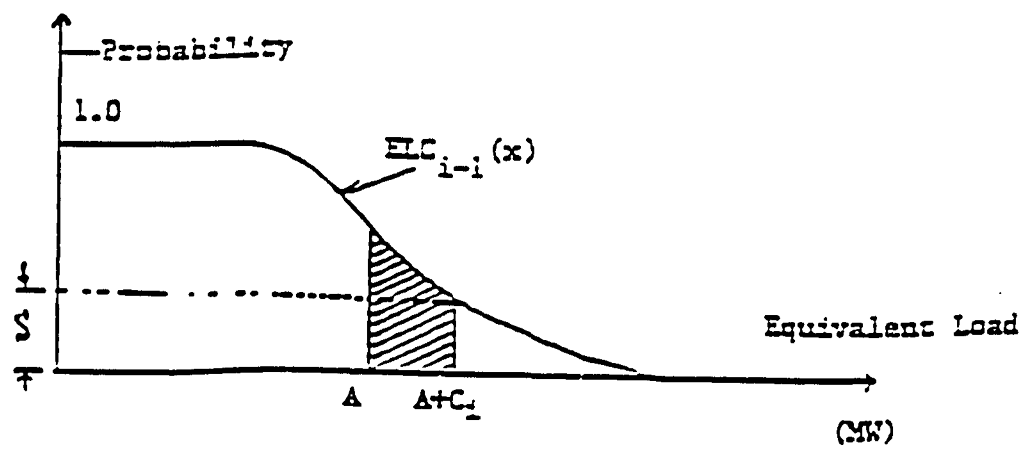


Figure A.7: Unit (i) Expected Generation.

Appendix B

Pricing Models Mathematical Solution

The welfare maximization model used in this research has the following form:

Maximize $f(x_1,\ldots,x_n)$,

Subject to:

$g_1(x_1,..,x_n) = b_1$

$\vdots$

$\vdots$

$g_m(x_1,\ldots,x_n) = b_m$, for $m < n$,

which is a classical equality-constrained optimization problem. The method for dealing with this problem is that of Lagrange multipliers. The procedure begins with formulating the Lagrange function:

$$\acute{L}(x_1,..,x_n,\lambda_1,..,\lambda_m) = f(x_1,\ldots,x_n) + \Sigma^{m}{}_{i=1}\ \lambda_i(g_i(x_1,..x_m)\text{-}b_i),$$

where $\lambda_1,\ldots,\lambda_m$ are called Lagrange multipliers. The method is reduced, then, to analyzing the unconstrained function $L(x_i,\lambda_j)$, where $i=1,..,n$, and $j=1,..,m$. Thus, the $(n+m)$ partial derivatives of function L with respect to x_i's and λ_j's are set equal to zero and then the critical points are obtained by solving the equations:

$$F_1 = dL/dx_1 = 0,$$

$$\vdots$$

$$F_n = dL/dx_n = 0,$$

$$F_{n+1} = dL/d\lambda_1 = 0,$$

$$\vdots$$

$$F_{n+m} = \partial L/\partial\lambda_m = 0,$$

To solve this set of nonlinear simultaneous equations, we use the Newton-Raphson algorithm. This method requires the evaluation of both the function $F(x,\lambda)$ and the derivatives $F'(x,\lambda)$ at arbitrary points. The method is derived from the familiar Taylor series expansion of a function in the neighborhood of a point X. If we let X denote the entire vector of values x_i and λ_j, each of the functions F_i can be expanded in Taylor series:

$$F_i(X+dX) \approx F_i(X) + \Sigma_{j=1}^{n+m}(dF_i(X)/dX_j)\, dX_j + (d^2F_i(X)/dX^2_j)\, dX^2_j/2 + ...$$

By neglecting terms of order dX^2 and higher, and setting $F_i(X+dX) = 0$, we obtain a set of linear equations for the corrections dX that moves each function closer to zero simultaneously, namely:

$$dX_j = - F_i(X)/\Sigma_{j=1}^{n+m} F'_{ij}(X), \qquad \text{or}$$

$[dX] = -[F/F']$, where F' is the Jacobian matrix.

The corrections are then added to the solution vector,

$$x^{new}_i = x^{old}_i + dx_i, \quad i=1,...,n,$$

$$\lambda^{new}_j = \lambda^{old}_j + d\lambda_j, \quad j=1,...,m,$$

and the process is iterated to convergence.

To solve the proposed pricing models, following the above procedure, we then need to obtain the first and second order conditions. In the following, the mathematical details of these conditions are provided for each model. Note that, although all variables and parameters in these models are time specific, we drop the notation t (for time period) to simplify the presentation.

Ramsey Pricing [RAMSEY2]

Maximize welfare:

$$Maximize\ W = \sum_{i=1}^{4} \int_0^{X_i} P_i(y,R)dy - TC(X) - FC$$

Subject to:

$$\sum_{i=1}^{4} P_i(X_i,R)\ X_i = TC(X) + FC$$

where $P_i = R_i(a_i - b_iX_i) - (1-R_i).OC_i$.

Note that in this model, all customers are served with the same reliability level, i.e. $R = R_i = (1\text{-LOLP})$, where LOLP is the system loss of load reliability at the total level of consumption $X_T = \Sigma_{i=1}^{4} X_i$.

The Lagrange equation is:

$$L = \sum_{i=1}^{4} [\int_0^{X_i} P_i(y,R)dy] - TC(X) - FC + \lambda(\sum_{i=1}^{4} P_i X_i - TC(X) - FC)$$

The first order conditions are:

$$F_1 = dL/dX_1 = P_1 - MC + \lambda(P_1 - X_1 R b_1 - MC) = 0,$$

$$F_2 = dL/dX_2 = P_2 - MC + \lambda(P_2 - X_2 R b_2 - MC) = 0,$$

$$F_3 = dL/dX_3 = P_3 - MC + \lambda(P_3 - X_3 R b_3 - MC) = 0,$$

$$F_4 = dL/dX_4 = P_4 - MC + \lambda(P_4 - X_4 R b_4 - MC) = 0,$$

$$F_5 = dL/d\lambda = \Sigma_{i=1}^{4} P_i X_i - TC(X) - FC = 0.0$$

where, MC and TC are the system marginal cost and total production costs at X_T.

Both are outputs of the production cost simulation model.

The second derivatives are as follows:

$$(H_{ij} = dF_i/dX_j,\ i=1 \rightarrow 5 \text{ and } j=1 \rightarrow 4;$$

$$= dF_i/d\lambda,\ i=1 \rightarrow 5 \text{ and } j=5)$$

$$H_{11} = -R_1 b_1 + \lambda(-2R_1 b_1)$$

$$H_{12} = 0.0$$

$$H_{13} = 0.0$$

$$H_{14} = 0.0$$

$$H_{15} = P_1 - X_1 b_1 R_1 - MC$$

$H_{21} = 0.0$

$H_{22} = -R_2b_2 + \lambda(-2R_2b_2)$

$H_{23} = 0.0$

$H_{24} = 0.0$

$H_{25} = P_2 - X_2b_2R_2 - MC$

$H_{31} = 0.0$

$H_{32} = 0.0$

$H_{33} = -R_3b_3 + \lambda(-2R_3b_3)$

$H_{34} = 0.0$

$H_{35} = P_3 - X_3b_3R_3 - MC$

$H_{41} = 0.0$

$H_{42} = 0.0$

$H_{43} = 0.0$

$H_{44} = -R_4b_4 + \lambda(-2R_4b_4)$

$H_{45} = P_4 -X_4b_4R_4 - MC$

$H_{51} = P_1 - X_1b_1R - Mc$

$H_{52} = P_2 - X_2b_2R - MC$

$H_{53} = P_3 - X_3b_3R - MC$

$H_{54} = P_4 - X_4b_4R - MC$

$H_{55} = 0.0$

Reliability-based Pricing [RDP-3] (three reliability options)

$$Maximize\ W = \sum_{i=1}^{4} \int_0^{X_i} P_i(y,R)dy - TC(X) - FC$$

subject to:

$$P_2X_2 = T1.Z - LOLP.SOC.[X_2 + \alpha.(X_1+X_4)]$$

$$P_3X_3 = T2.Z - LOLP.SOC.[X_3 + (1-\alpha).(X_1+X_4)]$$

$$P_1X_1 + P_4X_4 = T3.Z + LOLP.X_T.SOC,$$

where

$T1 = X_2/X_T$

$T2 = X_3/X_T$

$T3 = (X_1+X_4)/X_T$

X_T = total consumption

SOC = average value of social outage costs

Z = TC(X) + FC

LOLP = the system loss of load probability

Lagrange equation is:

$$L = W + \lambda_1[P_2X_2 - T1.Z + LOLP.SOC.[X_2 + \alpha.(X_1+X_4)]]$$

$$+ \lambda_2[P_3X_3 - T2.Z + LOLP.SOC.[X_3 + (1-\alpha).(X_1+X_4)]]$$

$$+ \lambda_3[P_1X_1 + P_4X_4 - T3.Z - LOLP.X_T.SOC]$$

The first order conditions are:

$$F_1 = dL/dX_1 = P_1(1+\lambda_3) - MC[1+\lambda_1 T1+\lambda_2 T2+\lambda_3 T3]$$
$$- Z[\lambda_1 T1^1+\lambda_2 T2^1+\lambda_3 T3^1]$$
$$+ LOLP.SOC[\alpha.\lambda_1+(1-\alpha).\lambda_2-\lambda_3] - \lambda_3.R_1.b_1.X_1=0.0$$

$$F_2 = dL/dX_2 = P_2(1+\lambda_1) - MC[1+\lambda_1 T1+\lambda_2 T2+\lambda_3 T3] - Z[\lambda_1 T1^2+\lambda_2 T2^2+\lambda_3 T3^2]$$
$$+ LOLP.SOC[\lambda_1 - \lambda_3] - \lambda_1.R_2.b_2.X_2 = 0.0$$

$$F_3 = dL/dX_3 = P_3(1+\lambda_2) - MC[1+\lambda_1 T1+\lambda_2 T2+\lambda_3 T3]$$
$$- Z[\lambda_1 T1^3+\lambda_2 T2^3+\lambda_3 T3^3]$$
$$+ LOLP.SOC[\lambda_2 - \lambda_3] - \lambda_2.R_3.b_3.X_3 = 0.0$$

$$F_4 = dL/dX_4 = P_4(1+\lambda_3) - MC[1+\lambda_1 T1+\lambda_2 T2+\lambda_3 T3]$$
$$- Z[\lambda_1 T1^4+\lambda_2 T2^4+\lambda_3 T3^4]$$
$$+ LOLP.SOC[\alpha.\lambda_1+(1-\alpha).\lambda_2-\lambda_3] - \lambda_3.R_4.b_4.X_4=0.0$$

$$F_5 = dL/d\lambda_1 = P_2X_2 - T1.Z + LOLP[X_2 + \alpha.(X_1+X_4).SOC] = 0.0$$

$$F_6 = dL/d\lambda_2 = P_3X_3 - T2.Z + LOLP[X_3 +(1-\alpha).(X_1+X_4).SOC] =0.0$$

$$F_7 = dL/d\lambda_3 = P_1X_1 + P_4X_4 - T3.Z - LOLP.X_T.SOC = 0.0,$$

where

$T1^i = dT1/dX_i$

$T2^i = dT2/dX_i$

$T3^i = dT3/dX_i, \quad i=1,2,3,4$

Note that, the higher order derivatives of T1, T2, and T3 are very small and can be neglected. Then, the elements of the second order (Hessian) matrix are obtained as follows:

$H_{11} = - R_1 b_1 (1+2\lambda_3)$

$H_{12} = 0.0$

$H_{13} = 0.0$

$H_{14} = 0.0$

$H_{15} = - MC.T1 - Z.T1^1 + \alpha.\ LOLP.SOC$

$H_{16} = - MC.T2 - Z.T2^1 + (1-\alpha).LOLP.SOC$

$H_{17} = P_1 - X_1.R_1.b_1 - MC.T3 - Z.T3^1 - LOLP.SOC$

$H_{21} = 0.0$

$H_{22} = - R_2 b_2 (1+2\lambda_1)$

$H_{23} = 0.0$

$H_{24} = 0.0$

$H_{25} = P_2 - X_2 R_2 b_2 - MC.T1 - Z.T1^2 + LOLP.SOC$

$H_{26} = - MC.T2 - Z.T2^2$

H_{27} = - MC.T3 - $Z.T3^2$ - LOLP.SOC

H_{31} = 0.0

H_{32} = 0.0

H_{33} = - $R_3b_3(1+2\lambda_2)$

H_{34} = 0.0

H_{35} = - MC.T1 - $Z.T1^3$

H_{36} = P_3 - $X_3R_3b_3$ - MC.T2 - $Z.T2^3$ + LOLP.SOC

H_{37} = - MC.T3 - $Z.T3^3$ - LOLP.SOC

H_{41} = 0.0

H_{42} = 0.0

H_{43} = 0.0

H_{44} = - $R_4b_4(1+\lambda_3)$

H_{45} = - MC.T1 - $Z.T1^4$ + α.LOLP.SOC

H_{46} = - MC.T2 - $Z.T2^4$ + $(1-\alpha)$.LOLP.SOC

H_{47} = P_4 - $X_4R_4b_4$ - MC.T3 - $Z.T3^4$ - LOLP.SOC

H_{51} = - MC.T1 - $Z.T1^1$ + α.LOLP.SOC

H_{52} = P_2 - $X_2R_2b_2$ - MC.T1 - $Z.T1^2$ + LOLP.SOC

H_{53} = - MC.T1 - $Z.T1^3$

H_{54} = - MC.T1 - $Z.T1^4$ + α.LOLP.SOC

$H_{55} = 0.0$

$H_{56} = 0.0$

$H_{57} = 0.0$

$H_{61} = - MC.T2 - Z.T2^1 + (1-\alpha).LOLP.SOC$

$H_{62} = - MC.T2 - Z.T2^2$

$H_{63} = P_3 - X_3R_3b_3 - MC.T2 - Z.T2^3 + LOLP.SOC$

$H_{64} = - MC.T2 - Z.T2^4 + (1-\alpha).LOLP.SOC$

$H_{65} = 0.0$

$H_{66} = 0.0$

$H_{67} = 0.0$

$H_{71} = P_1 - X_1.R_1.b_1 - MC.T3 - Z.T3^1 - LOLP.SOC$

$H_{72} = - MC.T3 - Z.T3^2 - LOLP.SOC$

$H_{73} = - MC.T3 - Z.T3^3 - LOLP.SOC$

$H_{74} = P_4 - X_4R_4b_4 - MC.T3 - Z.T3^4 - LOLP.SOC$

$H_{75} = 0.0$

$H_{76} = 0.0$

$H_{77} = 0.0$

Reliability-based Pricing [RDP-2] (two reliability options)

$$Maximize\ W = \sum_{i=1}^{4} \int_0^{X_i} P_i(y,R)dy\ -TC(X)-FC$$

subject to:

$P_1X_1 + P_4X_4 = T1.Z + LOLP.X_T.SOC$

$P_2X_2 + P_3X_3 = T2.Z - LOLP.X_T.SOC$

where

$T1 = (X_1+X_4)/X_T$

$T2 = (X_2+X_3)/X_T$

X_T = total consumption

SOC = average value of social outage costs

Z = TC(X) + FC

LOLP = loss of load probability

The Lagrange equation is:

$$L = W + \lambda_1[P_1X_{1+P}4X_4 - T1.Z - LOLP.X_T.SOC$$

$$+ \lambda_2[P_2X_2+P_3X_3 - T2.Z + LOLP.X_T.SOC$$

The first order conditions are:

$$F_1 = dL/dX_1 = P_1(1+\lambda_1) - MC[1+\lambda_1T1+\lambda_2T2] - Z[\lambda_1T1^1+\lambda_2T21]$$

$$+ LOLP.SOC[\lambda_2-\lambda_1] - \lambda_1.R_1.b_1.X_1 = 0.0$$

$$F_2 = dL/dX_2 = P_2(1+\lambda_2) - MC[1+\lambda_1 T1+\lambda_2 T2] - Z[\lambda_1 T1^2+\lambda_2 T2^2]$$
$$+ LOLP.SOC[\lambda_2 - \lambda_1] - \lambda_2.R_2.b_2.X_2 = 0.0$$

$$F_3 = dL/dX_3 = P_3(1+\lambda_2) - MC[1+\lambda_1 T1+\lambda_2 T2] - Z[\lambda_1 T1^3+\lambda_2 T2^3]$$
$$+ LOLP.SOC[\lambda_2 - \lambda_1] - \lambda_2.R_3.b_3.X_3 = 0.0$$

$$F_4 = dL/dX_4 = P_4(1+\lambda_1) - MC[1+\lambda_1 T1+\lambda_2 T2] - Z[\lambda_1 T1^4+\lambda_2 T2^4]$$
$$+ LOLP.SOC[\lambda_2-\lambda_1] - \lambda_1.R_4.b_4.X4 = 0.0$$

$$F_5 = dL/d\lambda_1 = P_1X_1 + P_4X_4 - T1.Z - LOLP.X_T.SOC = 0.0$$

$$F_6 = dL/d\lambda_2 = P_2X_2 + P_3X_3 - T2.Z + LOLP.X_T.SOC = 0.0,$$

where

$$T1^i = dT1/dX_i$$
$$T2^i = dT2/dX_i, \quad i=1,2,3,4$$

By neglecting the higher order derivatives of T1 and T2, the formulation of the second order conditions (Jacobian) matrix is as follows.

Second order conditions:

$H_{11} = - R_1 b_1 (1+2\lambda_1)$

$H_{12} = 0.0$

$H_{13} = 0.0$

$H_{14} = 0.0$

$H_{15} = P_1 - R_1 b_1 X_1 - MC.T1 - Z.T1^1 - LOLP.SOC$

$H_{16} = - MC.T2 - Z.T2^1 + LOLP.SOC$

$H_{21} = 0.0$

$H_{22} = - R_2 b_2 (1+2\lambda_2)$

$H_{23} = 0.0$

$H_{24} = 0.0$

$H_{25} = - MC.T1 - Z.T1^2 - LOLP.SOC$

$H_{26} = P_2 - R_2 b_2 X_2 - MC.T2 - Z.T2^2 + LOLP.SOC$

$H_{31} = 0.0$

$H_{32} = 0.0$

$H_{33} = - R_3 b_3 (1+2\lambda_2)$

$H_{34} = 0.0$

$H_{35} = - MC.T1 - Z.T1^3 - LOLP.SOC$

$H_{36} = P_3 - X_3 R_3 b_3 - MC.T2 - Z.T2^3 + LOLP.SOC$

$H_{41} = 0.0$

$H_{42} = 0.0$

$H_{43} = 0.0$

$H_{44} = - R_4 b_4 (1 + \lambda_1)$

$H_{45} = P_4 - R_4 b_4 X_4 - MC.T1 - Z.T1^4 - LOLP.SOC$

$H_{46} = - MC.T2 - Z.T2^4 + LOLP.SOC$

$H_{51} = P_1 - R_1 b_1 X_1 - MC.T1 - Z.T1^1 - LOLP.SOC$

$H_{52} = - MC.T1 - Z.T1^2 - LOLP.SOC$

$H_{53} = - MC.T1 - Z.T1^3 - LOLP.SOC$

$H_{54} = P_4 - R_4 b_4 X_4 - MC.T1 - Z.T1^4 - LOLP.SOC$

$H_{55} = 0.0$

$H_{56} = 0.0$

$H_{61} = - MC.T2 - Z.T2^1 + LOLP.SOC$

$H_{62} = P_2 - R_2 b_2 X_2 - MC.T2 - Z.T2^2 + LOLP.SOC$

$H_{63} = P_3 - X_3 R_3 b_3 - MC.T2 - Z.T2^3 + LOLP.SOC$

$H_{64} = - MC.T2 - Z.T2^4 + LOLP.SOC$

$H_{65} = 0.0$

$H_{66} = 0.$

Appendix C

Acronyms

Convolution The convolution of two functions h(t) and g(t), denoted by g*h, is

$$g*h = \int_{-\infty}^{\infty} g(\tau)h(t-\tau)d\tau$$

defined by

Note that g*h is a function and that g*h = h*g.

Cumulants Linear combinations of statistical moments. The cumulants of distribution are defined in terms of the first moment about the origin (the mean) and the higher moments about the mean.

ELDC Equivalent Load Duration Curve. It is a LDC augmented (convoluted) with the load caused by the random outages of generating units.

FOR Forced Outage Rate. Probability of unpalnned outage of a generation unit.

LDC Load Duration Curve. The probability that a random load X will be equal or exceed a demand level X (MW).

LOEP Loss of Energy Probability. The expected fraction of total energy sales the utility would be unable to serve due to generation shortages.

LOLP Loss of Load Probability. It estimates the amount of time, within a time frame, during which the system capacity would be unable to meet a given demand.

MF Mean Time Failure. The mean time between failures.

Moments Expected values of probabilistic variables.

MR Mean Time Repair. The mean time duration of failures

OC Outage Cost ($/Mwh).

PDF Probability distribution function.

RD Residual Demand.

SOC Average Outage Cost per customer.

BIBLIOGRAPHY

Baron, D. P. (1981), "Price Regulation, Quality, and Asymmetric Information," American Economic Review, Vol.71, No.1, pp. 212-221.

Billinton, R., G. Walker, and R. K. Eubramuniam. "Factors Affecting The Development of a Commercial Customer Damage Function." IEEE Transactions on Power System PWRS1, No.5 (November 1986): 28-33.

----------."Factors Affecting the Development of a Residential Customer Damage Function." IEEE Transactions on Power System PWRS2, No.1 (February 1987): 204-208.

Booth, R.R.,"Power System Simulation Model Based on Probability Analysis," IEEE Trans., Vol. PAS 91,No.1, pp 62-69, 1972.

Caramanis, M., Stremel, J., Fleck, W., and Daniel, S., " Probabilisitic Production Costing," International Journal of Electric Power and Energy Systems, Vol.5, pp. 75-86, 1983.

Chamberlain, J. (1985), "Pricing and Incentives," IEEE Proceedings Special Issue on Load Management (October).

Chao, Hung-po, Robert B. Wilson (1976), "Priority Service: Investment and Market Organization," American Economic Review, Vol. 77, No.5, pp. 899-916.

Chao, Hung-po, Shmuel S. Oren, Stephen A. Smith and Robert B. Wilson (1986), "Multilevel Demand Subscription Pricing for Electric Power," Energy Economics, pp. 199-217.

-------------- (1987), "Selected Papers on Priority Service Methods," EPRI P-5350. Electric Power Research Institute.

-----------, (1989), "Priority Service: Market Structure and Competition," The Energy Journal, Vol 9, pp. 77-103.

Cramer, H., "Mathematical Methods of Statistics," Princeton Univ. Press, 1974.

Crew, M.A., P.R. Kleindorfer,"The Economics of Public Utility Regulation," first edition, MIT Press, Cambridge, 1986.

Crew, M.A., P.R. Kleindorfer, "Reliability and Public Utility Pricing," The American Economic Review, Vol. 68, No. 1, 1978, pp. 31-40.

Dansby, R. E. (1979), "Multi-period Pricing with Stochastic Demand," J.E.,9,223-37.

Electric Power Research Institute,"Environmental and Socioeconomic Consequences of a Shortage in Installed Generating Capacity," EPRI EA-2462, Project 1374, June, 1982.

EPRI,"Customer Demand for Service Reliability: A Synthesis of Outage Cost Literature," EPRI report P-6510, Project 2801-1, Sep. 1989.

EPRI,"Environmental and Socioeconomic Consequences of a Shortage in installed Generation Capacity," EPRI Report EA-2462, June 1982.

EPRI,"Priority Service: Unbundling the Quality Attributes of Electric Power," EPRI Report EA 4851, Nov. 1986.

EPRI,"The Value of Service Reliability to Consumers," EPRI Report, EA-4494, May 1986.

Fisinger, J. (1980), "Peak-Load Pricing and Rationing Policies," Journal of Economics, Vol 40, No. 2, pp. 169-182.

Guldmann, J.M, "A Chance-Constrained Programming Approach to Natural Gas Curtailment Decisions", Resources and Energy, 3(1981) 133-161.

Guldmann, J. M., "Supply, Storage, and Service Reliability Decisions by Gas Distribution Utilities: A Chance-constrained Approach", Management Science, Vol.29, No.8, August 1983.

Harris, M. and A. Raviv (1981),"A Theory of Monopoly Pricing Schemes with Demand Uncertainty," American Economic Review, Vol 71, No.3, pp.347-365.

IEEE Reliability Test System Report of Task Force on APM Subcommittee, IEEE PAS-98, No.6 (November/December 1979) : 2047-54.

Joskow, P. L., and Schmalensee, R. (1986), "Incentive Regulation for Electric Utilities," Yale Journal on Regulation, 4, pp. 1-50.

Manove, Michael (1983),"Provider Insurance," Bell Journal of Economics, Vol 14, No.2 (Autumn), pp. 489-96.

Kendall and Stuart,"The Advanced Theory of Statistics," Vols. 1,2, MacMillan Publishing Co., Inc., New York, 1977.

Marchand, Maurice G. (1974), "Pricing Power Supplied on an Interruptible Basis," European Economic Review, Vol 5, pp.263-274.

Mazumdar, M. and Yin, C.K. "Variance of Power Generating System Production Costs," Sp88-68, Summer Meeting of the IEEE power Engineering Society, Portland OR., July 1988.

Mirman, Leonard, and David Sibley (1980), "Optimal Nonlinear Prices for Multiproduct Monopolies," Bell Journal of Economics, 11:659-70.

Muansinghe, M. (1981),"Principles of Modern Electricity Pricing," Proceedings of IEEE, 69 NO. 3, (March), pp.332-48.

Mussa, Michael, and Sherwin Rosen (1978), "Monopoly and Product Quality," Journal of Economic Theory, 18:301-17.

Nakamura, S.," A Review of Electric Production Simulation and Capacity Expansion Planning Programs," Energy Research, Vol. 8, pp. 231-240, 1984.

Nakamura, S., and Brown, J.,"Accuracy of Probabilistic Simulation Using Fourier Expansions," Proceedings of 2nd WASP Conference, Aug. 1977, Ohio State University.

Nakamura, S., Tzemos, S., and Harunuzzaman, M.," CERES: Capacity Expansion and Reliability Evaluation System," The National Regulatory Research Institute NRRI,1981.
Vardi, J., and Avi-Itzhak, B.,"Electric Energy Generation," MIT Press, Cambridge, 1981.

Oren, Shmuel, Stephen Smith, "Design of Curtailable Electricity Service to Reduce Annual Peaks," EPRI/UERG, Symposium at Univ. of Calif. Berkeley, June, 1990.

Oren, Shmuel, Stephen Smith, and Robert Wilson (1985), "Capacity Pricing," Econometrica, 53:545-66.

Oren, Shmuel, Joseph A. Doucet (1990), "Interruption Insurance for Generation and Distribution of Electric Power," Journal of Regulatory Economics.

Pitbladdo, Richard (1985),"Interruptible Service Under Individual Uncertainty and

Imperfect Communication," Phd Dissertation, Dept. of Economics, Stanford University.

Rau, N.S., and Schenk, K.F.,"Application of Fourier Method for Evaluation f Capacity outage Probabilities," IEEE Winter Summer Meeting, Feb., 1979.

Rau, N.S. and Necsulescu, C. "Probability Distributions of incremental cost of production and production cost," IEEE Trans. on PA and S, dec. 1985, pp.3493-3500.

Rau, N.S., And Necsulescu, C. "An Investigation of Two Methods for the Probabilistic Energy Production Simulation." IEEE Trans. on PA and S, Vol. PAS 102, No.8, Aug,1983, pp.2543-2551.

Rau, N.S., Toy, P., and Schenk, K.F.,"Expected Energy Production Costs by the Method Of Moments," IEEE Trans., Vol. PAS 99, No.5, pp. 1908-1917, 1980.

Rau, N.S. and Y. Hegazy (1989),"The Variance of Production Cost," IEEE Transactions on Power System, PES, 89 SM 706-3 PWRS, July.

Rau, N.S., and Hegazy, Y., "The Variance Of Production Cost," IEEE Trans. on Power System, Vol.5, No.2, pp.367-376, May, 1990.

Rau, N.S. and Y. Hegazy (1990),"Reliability Differentiated Pricing of Electricity Service," NRRI, NR-90-5, March.

Rosenfield, A., D. Bulliet, and R. Peddie, (1985), "Smart Meters and Spot Pricing: Experiments and Potential," Lawrence Berkeley Lab. Working Paper 20507.

Schenk, K.F., Misra, R.B., Vassos, S., and Wen, W.," A New Method for the Evaluation of Expected Energy Generation and Loss of Load Probability," IEEE Trans., Vol. PAS 103, NO.2, pp.294-304, 1984.

Schmalensee, R.,"Another Look at the Social Valuation of Input Price Changes," American Economic Review, Vol.66, N.1, March 1976, PP.239-243.

Smith, Stephen A. (1989), "Efficient Menu Structure for Pricing Interruptible Electric Power Service," Journal of Regulatory Economics, Vol 1, No. 3, pp. 203-224.

Stremel, J.P., Jenkins, R.T., Babb, R.A., and Bayless, W.D., "Production Costing Using the Cumulant Method of Representing the Equivalent Load Curve," IEEE Trans., Vol. PAS 99, No.5, pp. 1947-1956, 1980.

Telson, M. (1975), "The Economics of Alternative Levels of Reliability for Electric Power Generation Systems," Bell Journal of Economics, Vol. 6(Autumn), pp.679-694.

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

Tzemos, S., and Nakamura, S.,"Alternating use of Cumulants and Piecewise Polynomials in Probabilistic Simulation for Electric Power Systems," Proceedings of Electric Generation Expansion Analysis Conf., March 1981, Ohio State University.

Vickery, W. (9171), "Responsive Pricing of Public Utility Services," Bell Journal of Economics and Management Science, Vol 2, pp. 337-346.

Viseanathan, N. and E. Tse (1989), "Monopolistic Provision of Congested Service with Incentive Based Allocation of Priorities," International Economic Review, No. 1, pp. 153-174.

Wilson, Robert B. (1989),"Efficient and Competitive Rationing," Econometrica, Vol 57, No. 1, pp. 1-40.

Wilson, Robert B. (1989),"Ramsey Pricing of Priority Service," Journal of Regulatory Economics, Vol 1, No.3, pp. 189-202.