ELECTRIC UTILITY PRICING AND INVESTMENT DECISIONS
UNDER UNCERTAINTY

by

Randall Poor Ellis

B.A. Yale University (1976)
M.S. London School of Economics (1977)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

August 1981

© Randall Poor Ellis 1981

The author hereby grants to M.I.T. permission to reproduce and
to distribute copies of this thesis document in whole or in
part.

Signature of Author ____________________________

Department of Economics

Certified by ____________________________

Paul L. Joskow
Thesis Supervisor

Accepted by ____________________________

Richard S. Eckaus
Department Committee on Graduate Students

OCT 19 1981

LIBRARIES
Electric Utility Pricing and Investment Decisions

Under Uncertainty

by

Randall Poor Ellis

Submitted to the M.I.T. Department of Economics on August 19, 1981 in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

ABSTRACT

This thesis develops a theoretical model of socially optimal prices and investments for electric utilities in the presence of exogenous demand and input price uncertainty. Multiple demand periods and multiple user classes, as well as multiple realizations of a random variable (e.g., income or energy using capital) are collapsed into a single probability distribution for demand. This "load distribution curve" is then used as the basis of pricing and investment decisions. Optimal prices are shown to be price elasticity and demand weighted averages of operating and shortage costs. Investment strategies in cases where different technologies have different lead times and where generating capacity is perfectly durable are considered. The optimal capital stock of the most capital intensive generating technology (base load capacity) is shown to decrease as uncertainty increases.

Four extensions of the theoretical model are explored. It is shown that uncertain fuel prices can be desirable if fuel prices are negatively correlated with the level of demand. Interfuel substitutability is examined in terms of random operating costs and it is demonstrated that the ability to change the loading order of different generating technologies may make a technology attractive even if its fuel cost has a higher expected value than some technology which does not allow interfuel substitutability. Average cost pricing is shown to exacerbate consumption uncertainty; and in both the average cost pricing and risk averse cases it is no longer optimal to minimize expected total costs.

A simulation model is used to demonstrate how the theoretical results derived here can be implemented on a practical basis. Using a simplified model with one price and two types of generating capacity successive levels of complexity are introduced: a price feedback loop, demand uncertainty, sequential investment and pricing decisions, and durable capital. Without the assumption of durable capital, random
fuel costs and average cost pricing are also considered. Preliminary findings suggest that it is important to consider the price feedback mechanism, but that demand and fuel cost uncertainties have only a modest impact on desired investments.

The theoretical framework should provide a useful approach for empirical planning and forecasting models.

Thesis Supervisor: Paul L. Joskow
Title: Professor of Economics
Massachusetts Institute of Technology
ACKNOWLEDGEMENTS

I owe an enormous debt of gratitude to my two thesis supervisors, Paul Joskow and Dick Schmalensee. Each provided helpful and extensive comments on several earlier drafts, and equally important, the encouragement to continue pursuing the subject at several junctures.

I am also grateful to Marty Zimmerman who stimulated early interest in the topic of this thesis, and Mike Caramanis at the M.I.T. Energy Laboratory for his continuing intellectual support. My fellow students were an important source of stimulation, and I benefitted immensely from discussions with my friends Roger Bohn, Jon Hamilton, Allen Jacobs, David Levine and Carl Shapiro. I would also like to say a special word of thanks to Ted Roth, whose late night computer messages and sense of humor made the thesis writing process more tolerable.

My parents provided me with the support, both financial and emotional, that made my graduate studies possible. For this I cannot be grateful enough.

I am grateful to the Department of Economics Energy Economics Research Fund and the Sloan Foundation for financial support while writing this thesis.

Finally, I wish to thank Alice Sanderson for her superb job typing and preparing this manuscript. Her pleasant acceptance of my tight schedule, obscure handwriting and repeated revisions deserves special note.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>2</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>4</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>5</td>
</tr>
<tr>
<td>List of Tables</td>
<td>7</td>
</tr>
<tr>
<td>List of Figures</td>
<td>8</td>
</tr>
<tr>
<td>I. Introduction</td>
<td>10</td>
</tr>
<tr>
<td>II. Overview</td>
<td>13</td>
</tr>
<tr>
<td>Footnotes to Chapter II</td>
<td>24</td>
</tr>
<tr>
<td>III. The Model</td>
<td>25</td>
</tr>
<tr>
<td>A. The Load Duration Curve</td>
<td>25</td>
</tr>
<tr>
<td>B. The Load Distribution Curve</td>
<td>31</td>
</tr>
<tr>
<td>C. Generation Technologies</td>
<td>34</td>
</tr>
<tr>
<td>D. Shortage Costs</td>
<td>36</td>
</tr>
<tr>
<td>E. Welfare Criterion</td>
<td>39</td>
</tr>
<tr>
<td>Footnotes to Chapter III</td>
<td>42</td>
</tr>
<tr>
<td>IV. Static Results</td>
<td>44</td>
</tr>
<tr>
<td>A. Optimal Investment Rules</td>
<td>44</td>
</tr>
<tr>
<td>B. Optimal Pricing Rules with a Single Price</td>
<td>49</td>
</tr>
<tr>
<td>C. Elasticities of Demand, Optimal Pricing and Profitability</td>
<td>53</td>
</tr>
<tr>
<td>D. Optimal Pricing with Two Prices</td>
<td>56</td>
</tr>
<tr>
<td>E. Effects of Increasing Uncertainty</td>
<td>62</td>
</tr>
<tr>
<td>Footnotes to Chapter IV</td>
<td>68</td>
</tr>
<tr>
<td>V. Dynamic Results with One Period Capital</td>
<td>71</td>
</tr>
<tr>
<td>A. Effect of Lengthening Lead Times</td>
<td>71</td>
</tr>
<tr>
<td>B. Sequential Pricing and Investment Decisions</td>
<td>77</td>
</tr>
<tr>
<td>C. Sequential Investment Decisions</td>
<td>85</td>
</tr>
<tr>
<td>Footnotes to Chapter V</td>
<td>89</td>
</tr>
<tr>
<td>VI. Dynamic Results with Durable Capital</td>
<td>90</td>
</tr>
<tr>
<td>A. Rapidly Growing Demand Case</td>
<td>90</td>
</tr>
<tr>
<td>B. Irreversibility Effect</td>
<td>94</td>
</tr>
<tr>
<td>Footnotes to Chapter VI</td>
<td>100</td>
</tr>
</tbody>
</table>
Table of Contents (continued)

VII. Extensions of the Model

A. Uncertain Input Prices ........................................... 101
B. Interfuel Substitutability ................................. 110
C. Average Cost Pricing ............................................ 116
D. Risk Aversion .................................................. 135
   Footnotes to Chapter VII ...................................... 143

VIII. Simulation Results Incorporating Pricing Effects ......... 146

A. Ten Modeling Frameworks .................................... 148
B. The Simulation Model ........................................... 154
C. Price Effects: Frameworks I and II ....................... 161
D. Uncertainty and Simultaneous Decision Making:
   Framework IV .................................................. 169
E. Sequential Decision Making: Framework V ............... 174
F. Durable Capital: Framework X ................................ 183
G. Sensitivity Analysis ............................................ 189
H. Random Fuel Costs and Average Cost Pricing ............ 198
I. "One Third Less Base Load Capacity" ....................... 204
J. Summary of Simulation Results ............................ 210
   Footnotes to Chapter VIII .................................. 212

IX. Summary and Conclusions ..................................... 219

Appendices

A. Derivation of Consumer Surplus Measures ................. 229
B. Inverse Representation of Total Generation Costs ....... 234
C. Profitability Results of Alternative Pricing
   Specifications .................................................. 237
D. Costs of Holding the Wrong Capacity Mix ................. 249
E. Detailed Description of Simulation Model ................. 253
F. Derivation of "Excess Generating Cost" Approximation
   of Option Values .............................................. 265

Bibliography ..................................................................... 278

Biographical Note .......................................................... 287
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-1</td>
<td>Notation</td>
<td>26</td>
</tr>
<tr>
<td>8-1</td>
<td>Alternative Investment Modeling Frameworks</td>
<td>149</td>
</tr>
<tr>
<td>8-2</td>
<td>Data Assumptions Used in Simulations</td>
<td>155</td>
</tr>
<tr>
<td>8-3</td>
<td>Simple Pricing Effects with No Uncertainty</td>
<td>163</td>
</tr>
<tr>
<td>8-4</td>
<td>Effects of Increasing Uncertainty with Simultaneous Pricing and Investment Decisions</td>
<td>171</td>
</tr>
<tr>
<td>8-5</td>
<td>Sequential Decision Making: Base Load Capacity, Peaking Capacity and Prices All Chosen with a Ten Year Lead Time</td>
<td>176</td>
</tr>
<tr>
<td>8-6</td>
<td>Sequential Decision Making: Base Load and Peaking Capacity with a Ten Year Lead Time, Prices Chosen in the Last Period</td>
<td>177</td>
</tr>
<tr>
<td>8-7</td>
<td>Sequential Decision Making: Base Load Capacity Chosen with a Ten Year Lead Time, Peaking Capacity and Prices Chosen in the Last Period</td>
<td>178</td>
</tr>
<tr>
<td>8-8</td>
<td>Sequential Decision Making: Base Load Capacity, Peaking Capacity and Prices All Chosen in the Last Period</td>
<td>179</td>
</tr>
<tr>
<td>8-9</td>
<td>Sequential Decision Making with Durable Capital</td>
<td>188</td>
</tr>
<tr>
<td>8-10</td>
<td>Sensitivity Analysis: Nonstochastic Parameters</td>
<td>190</td>
</tr>
<tr>
<td>8-11</td>
<td>Sensitivity Analysis: Demand Growth and Demand Uncertainty Parameters</td>
<td>194</td>
</tr>
<tr>
<td>8-12</td>
<td>Uncertain Fuel Costs</td>
<td>200</td>
</tr>
<tr>
<td>8-13</td>
<td>Average Cost Pricing</td>
<td>202</td>
</tr>
<tr>
<td>8-14</td>
<td>One Third Less Base Load: Marginal Cost Pricing</td>
<td>207</td>
</tr>
<tr>
<td>8-15</td>
<td>One Third Less Base Load: Average Cost Pricing</td>
<td>209</td>
</tr>
<tr>
<td>E-1</td>
<td>Programs and Subroutines for Simulation Model</td>
<td>263</td>
</tr>
</tbody>
</table>
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1</td>
<td>Derivation of Load Distribution Curve from Two Load Duration Curves</td>
<td>17</td>
</tr>
<tr>
<td>3-1a</td>
<td>Pattern of Demand on a Typical Summer Weekday</td>
<td>29</td>
</tr>
<tr>
<td>3-1b</td>
<td>Load Duration Curve</td>
<td>29</td>
</tr>
<tr>
<td>3-2a</td>
<td>Derivation of Load Distribution Curve from Conditional Density Functions</td>
<td>33</td>
</tr>
<tr>
<td>3-2b</td>
<td>Contrast Between Load Distribution Curve and Load Duration Curve with the Same Mean</td>
<td>33</td>
</tr>
<tr>
<td>4-1a</td>
<td>Screening Curves</td>
<td>48</td>
</tr>
<tr>
<td>4-1b</td>
<td>Load Distribution Curve</td>
<td>48</td>
</tr>
<tr>
<td>4-2a</td>
<td>Change in Short Run Costs as a Result of a Price Increase</td>
<td>52</td>
</tr>
<tr>
<td>4-2b</td>
<td>Change in Long Run Costs as a Result of a Price Increase</td>
<td>52</td>
</tr>
<tr>
<td>4-3a</td>
<td>Constant Price Elasticity Case</td>
<td>54</td>
</tr>
<tr>
<td>4-3b</td>
<td>Price Elasticity Decreasing with the Level of Demand</td>
<td>54</td>
</tr>
<tr>
<td>4-3c</td>
<td>Price Elasticity Increasing with the Level of Demand</td>
<td>54</td>
</tr>
<tr>
<td>4-4</td>
<td>Effects of Increasing Uncertainty While Holding Price Fixed</td>
<td>65</td>
</tr>
<tr>
<td>5-1</td>
<td>Effects of the Resolution of Uncertainty Over Time</td>
<td>75</td>
</tr>
<tr>
<td>5-2</td>
<td>Contrast Between the Horizontal and Vertical Expectations of the Load Distribution Curves</td>
<td>80</td>
</tr>
<tr>
<td>5-3</td>
<td>An Example of Sequential Decisions</td>
<td>87</td>
</tr>
<tr>
<td>7-1</td>
<td>Capacity Choices Ignoring Operating Cost Uncertainty</td>
<td>107</td>
</tr>
<tr>
<td>7-2a</td>
<td>Average Cost Pricing with Locally Increasing Returns</td>
<td>120</td>
</tr>
<tr>
<td>7-2b</td>
<td>Average Cost Pricing with Locally Decreasing Returns</td>
<td>120</td>
</tr>
<tr>
<td>7-2c</td>
<td>Average Cost Pricing with Locally Constant Returns</td>
<td>120</td>
</tr>
<tr>
<td>7-3</td>
<td>Average Cost Pricing with Demand Uncertainty</td>
<td>124</td>
</tr>
</tbody>
</table>
List of Figures (continued)

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7-4a</td>
<td>Case Where AC Pricing Exacerbates Operating Cost Uncertainty</td>
<td>127</td>
</tr>
<tr>
<td>7-4b</td>
<td>Case Where AC Pricing Is Equivalent to MC Pricing</td>
<td>127</td>
</tr>
<tr>
<td>7-4c</td>
<td>Case Where AC Pricing Mitigates Operating Cost Uncertainty</td>
<td>127</td>
</tr>
<tr>
<td>7-5</td>
<td>Average Cost Pricing with Capital Cost Uncertainty</td>
<td>129</td>
</tr>
<tr>
<td>8-1</td>
<td>Linear Load Duration Curve</td>
<td>159</td>
</tr>
<tr>
<td>8-2</td>
<td>Effect of Intermediate Capacity on Base Load Use</td>
<td>197</td>
</tr>
<tr>
<td>B-1</td>
<td>Diagrammatic Representation of Total Generating Costs when Capacity is Chosen Optimally</td>
<td>236</td>
</tr>
<tr>
<td>D-1</td>
<td>Cost of the Wrong Generating Mix</td>
<td>251</td>
</tr>
<tr>
<td>E-1</td>
<td>Alternative Representations of Load Duration Curve</td>
<td>254</td>
</tr>
<tr>
<td>E-2</td>
<td>Example of Deriving a Load Distribution Curve from Three Load Duration Curves</td>
<td>257</td>
</tr>
<tr>
<td>E-3</td>
<td>Illustration of a Simple Algorithm for Finding $P^<em>$ and $K^</em>$</td>
<td>261</td>
</tr>
<tr>
<td>F-1</td>
<td>Costs of Excess Capacity and Approximation Used for Simulations</td>
<td>271</td>
</tr>
<tr>
<td>F-2</td>
<td>Proportionality of Optimal Base Load Capacity to $n$</td>
<td>272</td>
</tr>
</tbody>
</table>
I. Introduction

In sharp contrast with the predictable prices, consumption and costs of the 1950's and 60's, electric utility forecasts in the 1980's are characterized by substantial uncertainty. This uncertainty poses a considerable challenge to electric utilities who must plan capacity additions as far as ten or twelve years into the future. The problem is further complicated by the fact that utilities are adopting (or at least being asked to adopt) new and different pricing schemes which have important ramifications for the desired level and pattern of capacity additions. This thesis develops a consistent theoretical framework for analyzing optimal pricing and investment decisions by electric utilities in the presence of uncertainties. A simulation model is then used to demonstrate how this theoretical framework can be implemented on a practical basis.

The basic results on welfare optimal pricing and investments in a static model with demand uncertainty are well established, and are presented in Turvey and Anderson (1977) and Crew and Kleindorfer (1979). The approach developed here clarifies, generalizes and extends these previous results in several important directions.

By collapsing both the time varying nature of demand and long-term uncertainties into a single probability distribution of demand we are able to characterize quite simply the effect of long-term uncertainties on desired prices and investments. We then use this technique to examine the impact of sequential investment decisions, durable capital, uncertain input prices and risk aversion.
The pricing rules developed here are more general than those which have been previously derived. No restrictions need be placed on cross price elasticities, and optimal pricing rules are derived not only for the case where there are different demand periods (e.g., peak and off peak) but also for cases where there are different classes of users (e.g., residential and industrial) or different dimensions of product service (e.g., peak use and total use). The technique used to derive these results is to look at the welfare effects resulting from price changes rather than the traditional approach of considering the welfare effects of quantity changes. Short run marginal cost pricing, long run marginal cost pricing and average cost pricing are each considered, as well as the feedback effect of prices on desired investments.

A simulation model is used to demonstrate how the theoretical results derived here can be implemented on a practical basis. Using a simplified model with one price and two types of generating capacity successive levels of complexity are introduced: a price feedback loop, demand uncertainty, sequential investment and pricing decisions, and durable capital. Without the assumption of durable capital, random fuel costs and average cost pricing are also considered. Preliminary findings suggest that it is important to consider the price feedback mechanism, but that demand and fuel cost uncertainties have only a modest impact on desired investments.

The thesis is organized into 9 chapters. Chapter II provides an overview of the theoretical issues and major results of this thesis. The model and notation used throughout are described in Chapter III.
In Chapter IV the basic results on optimal investment and optimal pricing are derived in a static model where capital only lasts one period. Dynamic questions involving the resolution of uncertainty over time are considered in Chapter V while still maintaining the assumption of one period capital, and the effects of lead times are highlighted. Both the timing of pricing decisions and the effect of varying lead times for different generating technologies are considered.

Chapter VI considers optimal investments in a dynamic setting with durable capital. A few basic results are derived showing the importance of the "irreversibility effect," i.e., the fact that utilities cannot disinvest in capital as easily as they can invest. Chapter VII extends the model to consider uncertain input prices, interfuel substitutability, average cost pricing and risk aversion.

Chapter VIII develops a simulation model which is used to implement many of the theoretical results which we have shown. Uncertain demand, sequential decision making, durable capital, random fuel costs and both marginal and average cost pricing rules are examined under a variety of modeling assumptions. Concluding remarks are contained in Chapter IX.
II. Overview

In recent years major advances have been made in both theoretical and empirical models of pricing and investment decisions by electric utilities. Building upon the seminal works of Hotelling (1938), Boiteux (1949) and Steiner (1957), theoretical work has explored the pricing and investment implications of indivisibility constraints (Williamson, 1966) increasing returns (Mohring, 1970), shifting peaks (Bailey and White, 1974) stochastic demand (Meyer, 1975; Crew and Kleinöhrer, 1976, 1979; and Turvey and Anderson, 1977) second best pricing (Baumol and Bradford, 1970; Sherman and Visscher, 1976) and intraperiod time varying demand (Dansby, 1975b, 1976, 1977a, and Crew and Kleinöhrer, 1979). Over the same period empirical investment planning and forecasting models have also become increasingly sophisticated. Many current models use nonlinear programming and dynamic programming algorithms to incorporate uncertain hydro generation, unconventional generating sources, discrete unit sizes, reliability constraints, and various further complications (Farrar and Woodruff, 1973; Noonan and Giglio, 1974; Joy and Jenkins, 1974; Rowse, 1980).

Despite these advances, there remains a significant gulf between the theoretical models, which have been developed largely by economists, and the empirical models, which have strong origins in the engineering and operations research literature. 1 Three serious shortcomings of both sets of models are readily identifiable. First, although economists place a great deal of emphasis on price
elasticities of demand in their theoretical work, most of the empirical models are essentially cost-minimization models rather than welfare maximization models: prices are largely ignored. Economists bear much of the blame for this shortcoming. Their traditional approach to the multi-period pricing problem (peak, off-peak, etc.) using inverse demand functions with constant price derivatives within each period, and assuming zero cross-price elasticities of demand between different demand periods is inappropriate for most real world applications. Furthermore, economists have paid relatively little attention to the related problems of efficient pricing with multiple classes of users (residential, industrial, etc.) as well as multiple dimensions of electricity use (peak kilowatt use, total kilowatt use, etc.). When all three of these pricing dimensions are confronted simultaneously, the difficulties of including prices in a manageable empirical model are indeed quite formidable.

A second major shortcoming of the theoretical and empirical models is their incorporation of uncertainties. Several theoretical papers have dealt with the issue of demand uncertainty (Meyer, 1975; Crew and Kleindorfer, 1976, 1979; Turvey and Anderson, 1977). Typically the demand period is divided up into a number of discrete intervals with the level of demand in each interval assumed to be stochastic. Although a useful theoretical simplification, little attention is focused on how this approach can be implemented on a practical basis. Several theoretical papers have also considered fuel cost uncertainties (Fuss and McFadden, 1971; Stewart, 1978; and Perrakis, 1980). Such
efforts have not considered the pricing implications of such input cost uncertainties, however, and the author is aware of no work which considers average cost pricing in the presence of input price uncertainty. The failure of theoretical models to develop pricing and investment rules under uncertainty which can be readily employed on a practical basis may explain why empirical models, although often simulated for different "scenarios", are essentially all designed to choose optimal investments for a single deterministic load duration curve and set of factor prices. Unfortunately, as we will show, this leads to suboptimal investment strategies.

A third shortcoming of most of the models is their treatment of dynamics. Virtually all of the theoretical results have been derived in a static framework, ignoring the complexities introduced by the fact that capital lasts more than one period, different types of generating capacity require different lead times, and pricing and investment decisions need not be made at the same time. In this area many of the engineering and operations research models are more advanced than those of economists. Out of necessity most empirical models incorporate some dynamic elements. Frequently, however, the optimization technique is essentially a sequence of static optimizations. Furthermore, since uncertainty is omitted, the importance of the resolution of uncertainty over time is ignored.

The present study is an effort to overcome some of these shortcomings. A general framework is developed incorporating multiple prices and multi-dimensional uncertainties in a theoretical model oriented towards choosing optimal generating capacity investments in a
dynamic setting where uncertainty is resolved slowly over time. Following historical tradition, generating technologies are all assumed to be fixed proportion, constant returns to scale processes with no indivisibilities. The objective function is taken to be expected consumers' plus producers' surplus, and for analytic simplicity, social shortage costs are assumed to be a constant cost per kilowatt-hour of demand not met. Also in the mainstream of tradition, issues of transmission and distribution costs are ignored.

Breaking with historical tradition, at the heart of our model is a very general, continuous representation of the load duration curve, a curve indicating the fraction of a period that demand is below a given level, x. For this paper, the load duration curve is represented by a probability distribution function, G(x|P, n), which is the conditional distribution of x given a vector of prices, P, and the realization of an exogenous random vector n. This load duration curve incorporates all of the usual information about short-term demand fluctuations, whether they be due to predictable daily, weekly and annual consumption patterns, or due to uncertain short-term fluctuations, such as those arising from weather or economic activity variations. Long-term uncertainties, such as uncertainty about future income or the level of energy-using capital, are represented in uncertainty about n. Two load duration curves, corresponding to two different realizations of n are shown with broken lines in Figure 2-1.

An important innovation of this paper is its manner of summarizing
In the case where \( f(\eta_1) = f(\eta_2) = \frac{1}{2} \) the load distribution curve is simply the horizontal average of the two equally likely load duration curves.

\[
\theta = \frac{1}{2} G(x|P, \eta_1) + \frac{1}{2} G(x|P, \eta_2)
\]

\( \theta = fraction of a year that demand is below level x \)

(Note reversal of axes)
both long-term and short-term uncertainty in a single curve, $H(x|P)$, which we call the load distribution curve. Whereas the load duration curve represents the conditional distribution of $x$ given $\eta$, the load distribution curve is the marginal distribution of $x$. Equivalently, the load distribution curve can be viewed as the expected load duration curve, where expectations are taken over the uncertain vector $\eta$.

Figure 2-1 illustrates how a load distribution curve is derived from two load duration curves in the simple case where $\eta$ only takes on two equally likely outcomes. Here, and in the general case where $\eta$ can take on many possible values, the load distribution curve is the horizontal average (expectation) of the various possible load duration curves.\textsuperscript{6}

A number of interesting results are easily derived using this characterization of demand uncertainty. First, note that since the load duration curve conditional on $\eta$ will in general be a nonlinear function of $\eta$, the load distribution curve (expected load duration curve) will not be simply the load duration curve conditional on the expected value of $\eta$, $G(x|P, \mathbb{E}_\eta(\eta))$. Investment planning models which ignore demand uncertainty will not choose investments optimally. This simple but important result has been noted previously by many authors.\textsuperscript{7} More can be said, however. One effect of increasing uncertainty on predictions of future demand levels is to increase the likelihood of both very high and very low levels of demand. In terms of our representation of demand, this means simply that the load distribution curve becomes steeper and demand dispersed over a wider
range of possible levels. This result is formalized here in terms of mean preserving spreads, and it is shown that as uncertainty increases, it is optimal to decrease investments in baseload capacity and increase peaking (and total) capacity. The increase in the total level of capacity which is optimal has frequently been noted in the literature, but the change in the optimal mixture of generating capacity has not been generally noted. 8

Because the treatment of prices and uncertainty used here differs from previous efforts, the traditional results on optimal investments and pricing in a static one-period capital model are rederived here using our notation. It is shown that the pricing and investment rules are formally the same with and without uncertainty. The optimal investment rules are the same as those found by Turvey (1968) and others: on the margin the additional capital costs incurred by replacing a small amount of one type of capacity with the next most capital-intensive alternative should just equal the expected reduction in operating costs resulting from that change.

The framework adopted here permits the consideration of any number of different prices, whether they be for different demand periods (daytime versus nighttime), different classes of users (residential versus industrial), or different dimensions of product use (peak versus total use). Optimal prices are derived by considering the welfare change resulting from a price change. Thus we look at the marginal cost of providing the additional consumption demanded as a result of a price change.
In the case where cross-price elasticities of demand are all zero, optimal prices are shown to be those such that if any price, \( p_j \) is perturbed the resulting change in expected total generating plus shortage costs, divided by the change in total kilowatt hours sold at price \( p_j \) should equal \( p_j \) itself. If price elasticities of demand are constant for each user during each pricing period across all possible states of nature, then expected profits are slightly positive. When cross-price elasticities of demand are not zero the pricing rules need to be changed to reflect the change in revenue collected during periods for which different prices are charged.

Issues of sequential decision making are considered in Chapter V in a dynamic framework where it is assumed that additional information is gained each period about future demand. The focus is on the effect which this resolution of uncertainty over time has on optimal pricing and investment decisions. For simplicity we continue to assume capital lasts only one period. In the case where prices and investments are set simultaneously, it is shown that increasing the lead time needed to construct plants has an impact similar to that of increasing the level of uncertainty, and hence the previous results of increasing uncertainty on the capacity mix and prices continue to hold.

The ability to adjust prices after the uncertainties reflected by \( \eta \) have been realized reduces demand uncertainty and allows a greater capital intensity than when prices and investments must be chosen simultaneously. When more capital intensive generating technologies have longer lead times, the static investment rules continue to apply,
as long as the load distribution curve on which decisions are based is updated for each decision.

A few basic results on optimal pricing and investment are derived in a dynamic setting with perfectly durable capital. It is shown that if demand is growing sufficiently fast that one adds additional generating capacity of each type in every period, then the investment rules derived in the static case continue to hold. If, as appears more likely, there is a possibility that unusually low demand growth can result in excess capacity lasting more than one period, then the "irreversibility effect" comes into play, and capacity decisions are affected by the fact that it is very costly to reduce capacity. In the simple case of a single type of durable generating capacity, we show that the "irreversibility effect" unambiguously reduces the optimal level of capacity below the level that is optimal in the static case.

Some of the most interesting theoretical results are contained in Chapter VII which extends the one period capital model in four new directions. In the presence of fuel cost uncertainty, the covariance of fuel costs with the level of demand turns out to be a critical variable. Compared to basing decisions solely on expected fuels costs, it is desirable to invest more heavily in those technologies which have negative covariance with consumption and less heavily in those which have a positive covariance. Interfuel substitutability is examined as an extension of the principles of fuel cost uncertainty. The ability to switch around the loading order of different generating technologies in response to fuel price fluctuations is seen to improve the
attractiveness of generating technologies which have the capability of utilizing more than one fuel.

The third topic introduced in Chapter VII is average cost pricing. Consumption is shown to be more highly variable under average cost pricing than under marginal cost pricing, and capital cost uncertainty is seen to have an impact on prices which is absent under marginal cost pricing. Of great practical significance is the fact that with uncertainty it is no longer socially optimal to minimize expected costs under a regime of average cost pricing. Instead less capacity than the cost minimizing level should be built in an effort to move (suboptimal) prices closer to marginal costs.

Risk aversion is the fourth topic considered in Chapter VII. The not unexpected finding with risk aversion is that one should invest less in those technologies which are most responsible for the uncertainties, whether they be fuel cost, capital cost or demand uncertainties. In the case where consumers are more risk averse than producers, increasing consumer risk aversion while holding producer risk aversion constant, increases price variability and increases desired capital intensity.

Chapter VIII uses a simulation model to demonstrate how the theoretical results of this paper can be applied on a practical basis and examine the empirical magnitude of those results. The simulation model differs from previous efforts in that: 1) it uses a closed loop model which explicitly takes into account the effect of capacity additions on prices and hence the levels of demand for which the
capacity is being chosen; 2) it incorporates uncertainty about both the level of future electricity demand and fuel prices by choosing capacity which is optimal for a probability weighted average of 100 different "scenarios;" 3) it explicitly models the sequential nature of investment and pricing decisions, whereby different decisions are made with differing amounts of information; and 4) in a few simple situations it examines the impact of the "irreversibility effect," i.e., the fact that generating capacity is durable and hence one should build less of it than one would choose if one could simply rent capacity for one year at a time.

Using this simulation model it is demonstrated that incorporating pricing effects has an important effect on desired investments. Although long-term demand and input price uncertainty affects desired investment by ten percent or more in some instances, for our base case assumptions uncertainty has a relatively small impact. We speculate that this small impact may be due to the fact that our simulations are based on only two different generating technologies; the absence of an "intermediate" generating capacity significantly affects the sensitivity of base load capacity choices to demand and fuel cost uncertainty.
Footnotes to Chapter II

1. For a recent review of various empirical electricity investment models see Lee, Stoughton and Badertcher (1978).

2. The case of neoclassical production technologies has been considered by Panzar (1976) and Dansby (1975b). Increasing returns are explored in Mohring (1970), while indivisibility constraints are considered by Williamson (1965) and Manne (1974). In the empirical literature the assumptions of fixed proportions and constant returns to scale are almost universally used (see Farrar and Woodruff, 1973; Joy and Jemkins (1974); Noonan and Giglio, 1974; Stoughton et al., 1980) although both continuous (e.g., Baughman et al., 1979) and discrete (e.g., Joy and Jenkins, 1974) characterizations of capacity choice have been considered.

3. This assumption is discussed in Section III.E, below. The implications of introducing risk aversion are considered in Chapter VII.

4. This assumption has been used by Turvey and Anderson (1977), Balasko (1974), Munasinghe and Gellerson (1979), EPRI (1978), and Anderson and Perl (1980). Various generalizations are discussed in Section III.D of this paper.

5. Transmission and distribution costs are incorporated in Turvey and Anderson (1977), Scherer (1977), and Munasinghe (1979).

6. That it is the horizontal and not the vertical average (expectation) of the various possible load duration curves should not be surprising given that for capacity and pricing decisions it is the expected fraction of a year that given level of capacity will be used and not the expected capacity that will be used a given fraction of the year which is of interest. This intuitive result cannot be easily seen using the traditional representation of demand in a given period as $x = X(P, n)$ (or worse $P = P(X, n)$), where demand is a function of price $P$ and random variable $n$. Given this approach the only expectation which is easily taken is the expected demand (or price).

7. See especially Brown and Johnson (1968), Sherman and Visscher (1976), Meyer (1975), and Crew and Kleindorfer (1976).

8. The one exception being Fuss and McFadden (1971) who make this point briefly in a model without prices.

9. This result has important practical significance in that if true, then average cost prices will be very close to the welfare optimal prices. See Section IV.C for further discussion of this issue.
III. The Model

In this chapter we develop the basic model which is used in deriving the subsequent results. The notation used is summarized in Table 3-1. Although in the dynamic setting essentially all of the variables are dated with time subscripts, we omit time subscripts here for clarity.

III.A. The Load Duration Curve

One of the most fundamental characteristics of electricity is that its usage varies substantially over time. Figure 3-1a illustrates the pattern of demand facing an electric utility on a typical summer weekday. The pattern during weekends or on weekdays during other seasons is quite different. The patterns of consumption throughout an entire year can usefully be summarized by a load duration curve which, as illustrated in Figure 3-1b, shows the fraction of the period that demand is below a given level. Thus in Figure 3-1b demand is below $x_1$, for fraction $\theta_1$, of the period and above $x_1$, for $1-\theta_1$.¹

We choose to represent the load duration curve by a continuous function $G(x)$, the fraction of the year that demand is below level $x$. $G(x)$ has all of the usual properties of a probability distribution function. In this case $G(0) = 0, G(\infty) = 1, \frac{dG}{dx} \geq 0$. In particular we will commonly use $g(x) = \frac{dG(x)}{dx}$, the analog of the density function corresponding to $G(x)$.

In order to maintain complete generality, we make the load duration curve a function of a vector of prices, $P$, which includes all
Table 3-1

NOTATION

\(i = 1, 2, \ldots, n\)  
plant type index, in order of decreasing capital intensity. (If only two technologies, then 1 = base load, 2 = peak)

\(k_i\)  
capacity of type \(i\) (kilowatts)

\(k = \{k_1, k_2, \ldots, k_n\}\)  
vector of capacities of each type

\[K_i = \sum_{j=1}^{i} k_j\]  
total capacity of plants of types up through \(i\). Note \(K_0 = 0, K_1 = k_1\)

\(K = \{K_1, K_2, \ldots, K_n\}\)  
vector of all cumulative capacities

\(c_i\)  
capital cost per kilowatt of capacity of type \(i\)

\(m_i\)  
operating cost per kilowatt per period of capacity of type \(i\)

\(\alpha\)  
constant shortage cost per kilowatt per period of unmet demand

\(p_j\)  
price charged for \(j\)th service

\(P = \{p_1, p_2, \ldots, p_J\}\)  
vector of all prices

\(\theta_i\)  
fraction of period that generating capacity of type \(i\) is used at less than full capacity

\(x\)  
level of demand (kilowatts)

\(\beta = \frac{1}{1 + \tau}\)  
constant discount factor
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\eta$</td>
<td>random vector affecting load duration curve (e.g. income or energy using capital stock)</td>
</tr>
<tr>
<td>$f(\eta)$</td>
<td>density function of $\eta$, defined over $S$</td>
</tr>
<tr>
<td>$S$</td>
<td>space in $\mathbb{R}^n$ over which $f(\eta)$ is defined</td>
</tr>
<tr>
<td>$G(x</td>
<td>P, \eta)$</td>
</tr>
<tr>
<td>$g(x</td>
<td>P, \eta) = \frac{dG(x</td>
</tr>
<tr>
<td>$h(x</td>
<td>P) = \int_S g(x</td>
</tr>
<tr>
<td>$H(x</td>
<td>P) = \int_0^x h(z</td>
</tr>
<tr>
<td>$D(P) = \int_0^{\infty} x h(x</td>
<td>P) dx$</td>
</tr>
<tr>
<td>$\tilde{D}(P)$</td>
<td>total expected served demand, i.e., demand excluding shortages</td>
</tr>
<tr>
<td>$UCS(P, \eta)$</td>
<td>unrationed consumers' surplus</td>
</tr>
<tr>
<td>$REV(P, \eta)$</td>
<td>total unrationed revenue</td>
</tr>
<tr>
<td>$FC(k)$</td>
<td>total fixed costs</td>
</tr>
<tr>
<td>$OC(P, k, \eta)$</td>
<td>total operating costs</td>
</tr>
</tbody>
</table>
Table 3-1 continued

\[ SC(P,k,n) \quad \text{total shortage costs} \]

\[ C(k,P) = \int_S [FC(k) + OC(P,k,n) + SC(P,k,n)]f(n)d(n) \quad \text{total costs} \]

\[ \overline{C}(k,P) \quad \text{total costs excluding shortage costs} \]

\[ U(P,k) = USC(P) + REV(P) - C(P,k) \quad \text{total welfare measure} \]

\[ s,t \quad \text{used as time subscripts in dynamic setting; preceding subscripts denote the time at which decisions are made, following subscripts the time at which outcomes are realized} \]

\[ s^C_t \quad \text{costs expected at time } s \text{ for period } t \]

\[ s^f_t(n) \quad \text{probability density function of } n \text{ for period } t \text{ from the perspective of } s \]

\[ s^{V_t}(K_t, K_{t-1}, n_s) \quad \text{discounted expected value of costs in period } t \text{ and thereafter from the perspective of period } s \]

\[ \delta^{s+1}(K) \quad \text{cutoff function for } n \]

\[ \omega(n) \quad \text{weighting function (used in Chapter VII)} \]

\[ E(\ ) \quad \text{expectation operator} \]
Figure 3-1a
Pattern of demand on a typical summer weekday

Figure 3-1b
Load Duration Curve

\[ x = x(P,t,\eta) \]

\[ \theta = G(x|P,\eta) \]

or \[ x = G^{-1}(\theta|P,\eta) \]

fraction of a year that demand is below level \( x \)
types of electricity prices. In particular, the elements of $P$, the $p_j$, might include prices charged for different demand periods (day, night, summer, winter, etc.) prices charged to different users (residential, commercial, industrial, etc.) prices charged for different dimensions of electricity use (peak kilowatt use, total kilowatt use, hookup charges, etc.) or prices for various permutations of the above uses (the peak kilowatt charge to industrial uses on a summer weekdays, for instance). An important point is that when one of these prices is changed it will in general change the shape of the load duration curve. Later, when deriving optimal prices, we will wish to consider the change in costs resulting from a change in prices, and relate this change in costs to the appropriate price.

It is useful to distinguish two different types of uncertainty which affect the demand for electricity: short-term and long-term uncertainty. Uncertainty which is due to the weather or stochastic fluctuations in the level of economic activity is reasonably classified as short-term in that as time progresses there is essentially no resolution of this type of uncertainty until immediately before its realization. Long-term uncertainty, on the other hand, has the property of being gradually resolved as one approaches its realization. For example, income and the choice of energy-using capital (heating and cooling equipment, appliances, machinery, etc.) both affect electricity consumption, and both have the property that the closer we are to their realization, the less the degree of uncertainty about them.
For this paper short-term uncertainty is assumed to be already reflected in our load duration curve, $G(x)$. Long-term uncertainty on the other hand is represented by a random vector $\eta$ with a density function $f(\eta)$, defined over $S$ in $\mathbb{R}^n$ and in subsequent chapters we allow $f(\eta)$ to change over time to reflect improved information. The random vector $\eta$ should be thought of as representing all of the random variables affecting demand, such as total income and the aggregate stock of energy-using capital. In Chapter VI it will prove more convenient to treat $\eta$ as a scalar, however in general it should be thought of as a full vector of random variables.

The technique used here for incorporating both prices and long term uncertainty is to make the load duration curve conditional on a vector of prices, $P$, and the realization of the random vector $\eta$. We use $G(x|P,\eta)$ to represent the conditional distribution of $x$ given $P$ and $\eta$. To simplify notation, some or all of the arguments of $G$ will often be suppressed, however the sense in which we use the load duration curve should always be understood.

III.B. Load Distribution Curve

Our representation of uncertainty and the load duration curve leads to a particularly convenient simplification. Defining

$$h(x|P) = \int_S g(x|P,\eta) f(\eta) \, d\eta \quad (3-1)$$

we can take the expectation of the load duration curve as follows:
\[ E_n[G(x_0|P, n)] = \int_S G(x_0|P, n) f(n) dn = \int_0^{x_0} \int_S g(x|P, n) f(n) dx dn \]

\[ = \int_0^{x_0} \int_S g(x|P, n) f(n) dn dx \]

\[ = \int_0^{x_0} h(x|P) dx \]

\[ = H(x_0|P) \quad (3-2) \]

We call \( H(x|P) \) the load distribution curve to distinguish it from \( G(x|P, n) \), the load duration curve. Whereas \( G(x|P, n) \) is the conditional distribution of \( x \) given \( n \), \( H(x|P) \) is the marginal distribution of \( x \). The importance of the load distribution curve is that it incorporates the uncertainty about \( n \), and hence reflects the full range over which demand may possibly vary.

In Figure 1-1 a load distribution curve is derived holding prices fixed in the simple case where \( n \) can only take on two equally likely outcomes. \( f(n_1) = f(n_2) = \frac{1}{2} \) and \( H(x|P) = \frac{1}{2} G(x|P, n_1) + \frac{1}{2} G(x|P, n_2) \) in that case. The load distribution curve is simply the horizontal expectation of the load duration curves. In the more general case where \( n \) can take on many different values it is easiest to derive the marginal distribution of \( x \) from the conditional density functions of \( x \). Figure 3-2a illustrates how the marginal density function of \( x \), \( h(x|P) \), could be derived while holding prices constant using the conditional density functions \( g(x|P, n) \). The resulting load distribution curve \( H(x|P) \) is shown in Figure 3-2b and contrasted with a
**Figure 3-2a**

Derivation of Load Distribution Curve from Conditional Density Functions

- \( f(\eta) \) marginal density function of \( \eta \)
- \( h(x|P) \) marginal density function of \( x \)

\[ g(x|P, \eta_{max}) \]
\[ g(x|P, \eta_0) \]
\[ g(x|P, \eta_{min}) \]

**Figure 3-2b**

Contrast between Load Distribution Curve and Load Duration Curve with the Same Mean

\( \theta = H(x|P) \)
\( \theta = G(x|P, \eta_0) \)
load duration curve conditional on $\eta_0$ which has the same mean as $H(x|P)$. The key property of the load distribution curve derived here, and one which will apply to most reasonable density functions $g(x|\eta)$ and $f(\eta)$ representing demand, is that for a given set of prices the load distribution curve has a greater dispersion than the load duration curve with the same mean. This is true for the same reason that the marginal density function of a random variable will in most cases have a greater dispersion (e.g. variance or range) than its conditional density function.³ The significance of this relationship for pricing and investment decisions is explored in Chapter IV, Section E.

At this point it is worth noting several conventions in notation which will be used. If both $g$ and $f$ are continuous density functions, then both $G(x|P, \eta)$ and $H(x|P)$ will be continuous, monotonic functions of $x$, so that each will have a well defined inverse. Letting $\theta = G(x|P, \eta)$ we will occasionally find it convenient to use $x = G^{-1}(\theta|P, \eta) = X(\theta, P, \eta)$, the inverse of the load duration curve. In a similar manner we will sometimes use $X = H^{-1}(\theta|P) = X(\theta, P)$, the inverse of the load distribution curve. We will occasionally find it convenient to consider total expected electricity consumption during a period, $D(P)$. We therefore define

$$D(P) = \int_S \int_0^\infty xg(x|P, \eta)dx f(\eta)d\eta \quad (3-3)$$

III.C. Generation Technologies

In order to provide service to a time varying demand, we model
electric utilities as choosing among \( n \) different fixed proportion, constant returns to scale, continuous generating alternatives which vary most importantly in their capital intensity. Since we are assuming that the technologies are fixed proportion technologies, we can order the \( n \) alternatives by their decreasing capital cost per kilowatt of capacity, \( c_i: c_1 > c_2 > \ldots > c_n > 0 \). If we let \( m_i \) be the annual operating costs per kilowatt from operating capacity of type \( i \) at full capacity, and we eliminate any technologies strictly dominated by any other (e.g., \( c_i > c_j \) and \( m_i > m_j \)) then our ordering also implies increasing operating costs: \( m_1 < m_2 < \ldots < m_n \).

Let the amount of capacity of type \( i \) be denoted \( k_i \), and define \[ K_i = \sum_{j=1}^{i} k_j. \] Cost minimization requires that the utility always uses available capacity of each type in ascending order. If demand, \( x \), lies between \( K_{i-1} \) and \( K_i \) then all capacity of types up through \( i-1 \) will be used completely, with the remainder, \( x-K_{i-1} \) generated by capacity of type \( i \). Using the notation for the load duration curve developed above, the total operating costs, \( OC \), from using capacity \( k_i \) efficiently will be:

\[
OC(k_i, P, n) = \int_{K_{i-1}}^{K_i} m_i(x-K_{i-1}) g(x | P, n) \, dx
\]

\[
+ \int_{K_i}^{\infty} m_i k_i g(x | P, n) \, dx \tag{3-4}
\]

Fixed costs, \( FC \), during any period are simply the product of the cost per unit of capacity \( (c_i) \) and the quantity of new capacity
additions. Starting from a point of zero capacity of type \( i \), we have:

\[
FC(k_i) = c_i k_i
\]

The total cost of electricity of all types in this period would then be

\[
FC(k) + OC(k, P, n) = \sum_{i=1}^{n} [FC(k_i) + OC(k_i, P, n)]
\]

Throughout this paper we will use \( k = \{k_1, k_2, \ldots, k_n\} \), \( K = \{K_1, K_2, \ldots, K_n\} \).

III.D. Shortage Costs

Given that generating capacity is costly, it will in general not be optimal to build capacity adequate to serve all possible levels of demand.\(^4\) Shortages are therefore likely and some mechanism for evaluating their social cost is necessary. It should be noted that the usual market solution, allowing prices to rise sufficiently to eliminate the shortage, is generally infeasible to do after the resolution of demand uncertainty.\(^5\)

Modeling the social cost of electricity shortages is a difficult undertaking because of the many factors influencing these costs. First, and perhaps most important, the time at which a shortage occurs has an important impact on its cost. Shortages due to insufficient capacity being available do not occur randomly; they occur precisely at times when demand is highest, e.g., hot summer weekdays.\(^6\) The cost of power shortages on winter weekends, or even cool summer weekdays may be very different. Second, utilities have some ability to ration shortages differentially among users. If users are rationed in order
of increasing value of use, then average shortage costs may increase with the size of the shortage. Third, the duration of shortages influences social costs. Fourth, unless demand is perfectly inelastic, total and average shortage costs will vary as prices vary. Lower prices increase the total value attached to electricity, but decrease the average value per kilowatt attached to electricity consumed. Fifth, the realization of long-term uncertainty variables may affect social shortage costs. As income increases, for instance, the total value attached to consumption will increase, although whether the average value per kilowatt-hour increases or decreases is indeterminate. Finally, knowledge that shortages are possible itself affects planned consumption. Back-up energy systems or plans to shift unmet demand into alternative, non-rationed time periods may substantially lower the direct costs of shortages but they imply significant indirect shortage costs.

Given the complexity of modeling the social cost of shortages, a number of different approaches have been adopted. The classic approach has been to base investments on either some exogenous reserve margin (e.g., Boyd and Thompson, 1980; Turvey and Anderson, 1977, p. 253) or loss of load probability (e.g., Brown and Johnson, 1969; Meyer, 1975; Rowse, 1980). These approaches essentially impose a reliability constraint on the pricing and investment problem and do not allow this to be an outcome of the welfare maximization problem.

An alternative and somewhat more appealing approach is to choose some fairly simple representation of social shortage costs. The
simplest and most common assumption is that social shortage costs are a constant cost per unit of unmet demand. (See, e.g. Balasko, 1974, Turvey and Anderson, 1977, p. 301; Munasinghe and Gellerson, 1979; Electric Power Research Institute, 1978; Anderson and Perl, 1980.) This assumption implies that total social shortage costs are a linear function of excess demand.

Three generalizations of this linear shortage cost function have been used in the literature. One generalization has been to make the linear shortage cost depend on the demand period during which shortages actually take place. A second generalization has been to allow social shortage costs per kilowatt to increase as the size of the shortage grows (e.g., Crew and Kleindorfer, 1979, p. 71). This approach captures the effect that selective rationing by the electric utility might have on social shortage costs, and makes shortage costs a non-linear function of excess demand. A third alternative generalization has been to leave the shortage cost function linear, but make its slope a function of prices. Brown and Johnson (1969) assumed the shortage cost per kilowatt was the price, where Crew and Kleindorfer (1976) and Dansby (1975b) use consumer surplus plus revenue divided by total demand as a measure of the shortage cost per kilowatt.

For this paper we represent social shortage costs as a constant cost per unit of unmet demand, \( \alpha \). \( \alpha \) should be thought of as the maximum charge per unit of consumption which consumers facing a shortage would be willing to pay in order to consume their normal
consumption of electricity. The above three generalizations to shortage costs could be introduced into the present framework, however they would complicate the analysis without adding further insight. The importance of our linearity assumption is noted when it becomes critical for our results.

The analytical simplifications which result from assuming $\alpha$ is constant are substantial. Given this representation of shortage costs it is possible to view the serving of demand via shortages as another constant returns to scale, continuous production technology with zero capital costs and operating cost of $\alpha$ per kilowatt per period. The cost minimization problem is then one of choosing among $n+1$ technologies to meet a fluctuating demand. Total expected shortage costs given $n$, $SC(K_n, P, n)$ can be written

$$SC(K_n, P, n) = \int_{K_n}^{\infty} \alpha(x - K_n) g(x | P, n) dx$$

III.E. Welfare Criterion

In order to meaningfully consider welfare optimal prices and investment, we must choose a relevant measure of welfare. For most of the paper we assume risk neutrality on the part of all agents and use expected consumers' plus producers' surplus as our welfare measure. The implications of risk aversion are examined in Chapter VII. In the present model welfare is measured by
Welfare = [unrationed consumers' surplus + unrationed total revenue - fixed costs - operating costs - shortage costs]

This measure of consumers' surplus, estimated via the integral under a demand curve is for a good which is used primarily as an intermediate factor of production, has frequently been used as basis of welfare calculations. In keeping with historic tradition, and for lack of any simple alternative, consumers' plus producers' surplus will be used here despite its shortcomings.

It should be noted that we have subtracted off shortage costs from the measure of unrationed consumers' surplus, and in so doing have implicitly assumed that the existence of occasional electricity outages does not affect the utility of consumption when the consumer is unrationed. The principal justification for this is that since outages due to transmission failure (which are not considered here) are far more common than those due to capacity shortages, the marginal impact of capacity shortages on unrationed consumer surplus may be small. This independence of unrationed consumers' surplus from the probability of capacity shortages may not be too unrealistic an assumption.

In evaluating consumers' surplus, economists typically use the inverse demand function and integrate over consumption under the demand curve. Here we wish to consider time varying demand, so instead we choose the more convenient alternative of integrating over all prices. The problem is complicated however by the fact that in addition to summing up consumer surplus across different individuals,
we must also sum up the surplus across time. It is shown in Appendix A that this can be accomplished simply by integrating the area under the load duration curve for various prices. In the case of a single price the expected consumer surplus at $P^0$ is

$$UCS(P^0) = \int_0^{\infty} \int_0^{\infty} xg(x|p,n) \, dx \, dp \, f(n) \, dn = \int_{P^0}^{\infty} D(p) \, dp \tag{3-6}$$

The appropriate expressions in cases where there are multiple prices are derived in Appendix A and considered in Section IV.C.

Before continuing it is worth stating several assumptions which have not been explicitly mentioned. In the model presented here we assume that there are no transmission and distribution costs interacting with generating costs. We also assume there are no forced outages or uncertainties about fuel availability. Lead times are known with certainty, and the complications introduced by uncertain hydroelectric generation are also ignored. These and other types of supply uncertainty could be added on to the present model, however, we ignore such considerations here. The implications of introducing supply side uncertainties are discussed briefly in the concluding chapter.
Footnotes to Chapter III

1. In using a load duration curve, one is implicitly assuming that daily demand fluctuations and capacity scheduling problems are unimportant, and the only relevant dimension of capacity use is the fraction of the year that the capacity is actually in use. Of course cycling and peaking capacity are designed differently from some base load capacity in order to accommodate short term demand fluctuations; however, these refinements are ignored here, as they are elsewhere. For a discussion of load duration curves, see Turvey and Anderson (1977), pp. 287-289.

2. P could include non-electricity prices as well; however, since we are using a partial equilibrium approach and holding these fixed, they will be omitted for simplicity.

3. It is of course easy to construct examples where the density functions of a random variable conditional on certain realizations of another random variable (including its mean) have a greater dispension than its marginal density function; however, we argue that these examples are empirically implausible in the case of electricity.

4. This statement is true even for completely deterministic load duration curves as long as there exists a demand level \(x_0\), which will be reached less than fraction of \(\delta\) of a year, for all \(\delta > 0\). Shortage costs should be considered even in the deterministic demand case, although in the deterministic case the difference between building the optimal level of capacity and the level of capacity sufficient to meet all possible demand levels may be negligible.

5. One exception to this is the possibility of "ripple" controlled pricing whereby the price to certain large users is determined by the concurrent level of demand. See Dansby (1977b) and Bohn et al. (1981).

6. This is a slight simplification in that if the capacity shortfall is severe enough, then shortages may occur even during periods when demand is well below peak, but capacity is off line for repairs or maintenance.

7. Although Munasinghe (1979) discusses the duration issue, none of the theoretical or empirical models are well designed to handle the effect of the duration of shortages.

8. A serious weakness in this use of consumers' surplus plus revenues as a measure of the value of foregone electricity consumption is that it amounts to assuming that expected and unexpected
electricity shortages both have the same welfare implications. The actual burden of shortages may be greater or less than this measure would indicate. This is one of the reasons that this measure was not used here.

9. I am indebted to Jean Tirole for this insight.


11. With the exception of Fuss and McFadden (1971), all of the literature we have seen which examines welfare optimal utility pricing under uncertainty assumes risk neutrality and uses expected consumers' plus producers' surplus as a measure of welfare. We examine risk aversion in Chapter VII.

12. To pursue this issue a little further, if capacity shortages were the only cause of blackouts, then all of the costs of guarding against and coping with blackouts (energy back-up systems, increased insulation, etc.) would legitimately have to be subtracted from consumers' plus producers' surplus. Due to comparatively frequent transmission failure, however, only the incremental blackout costs of capacity shortages need be subtracted off from the welfare measure. We would argue that these costs may be fairly small.

IV. Static Results

In this chapter rules for optimal investment and pricing are derived in a model where capital lasts only one period and pricing and investment decisions must be made simultaneously. To simplify presentation, we will frequently consider a model with only one price, $p^0$, and only two generating technologies: base load ($k_1$) and peaking ($k_2$). Most results in this simple case are easily generalized to the case of many generating technologies and many prices.

IV.A. Optimal Investment Rules

Using the notation developed above, the objective function, $U(k,P)$, expected consumer plus producer surplus, in the one period case can be written:

$$U(k,P) = E_n \left[ \text{unrationed consumer surplus + total unrationed revenue} - \text{fixed costs} - \text{operating costs} - \text{shortage costs} \right]$$

$$= E_n \left[ \text{UCS}(P,n) + \text{REV}(P,n) - FC(k) - OC(k,P,n) - SC(k,P,n) \right]$$

(4-1)

Rules for optimal pricing and investment can be derived by partially differentiating the above expression with respect to prices, $P$, and capacities, $k$. Here we focus on optimal investment rules for any given set of prices, $P$, deferring until the next section the consideration of optimal pricing.

Notice that in our formulation the first two terms, which represent the total value of unrationed electricity consumption, are independent
of the mix of generating capacity chosen. The welfare optimal
investments are simply the cost minimizing investments, where costs
include fixed costs, expected operating costs, and expected shortage
costs. Letting \( C(k, P) \) be the sum of these three costs, we can
write

\[
C(k, P) = \int_S \left[ FC(k) + OC(k, P, n) + SC(k, P, n) \right] f(n) dn
\]

\[
= \int_S \left\{ \sum_{i=1}^{n} c_i k_i + \int_{K_{i-1}}^{K_i} m_i (x-K_{i-1}) g(x|P, n) dx + \int_{K_i}^{\infty} m_i k_i g(x|P, n) dx \right\} f(n) dn
\]

\[
+ \int_{K_n}^{\infty} \alpha(x-K_n) g(x|P, n) dx \right\} f(n) dn \tag{4-2}
\]

To simplify exposition we will consider the case of only two generating
technologies and omit the price argument. It will also be convenient to
write total costs as a function of the \( K_i \), cumulative capacities, instead
of the \( k_i \). Making these changes and simplifying, we have:

\[
C(K_1, K_2) = c_1 K_1 + c_2 (K_2 - K_1)
\]

\[
+ \int_S \left\{ m_1 \int_0^\infty x g(x|n) dx + (m_2 - m_1) \int_{K_1}^{\infty} (x-K_1) g(x|n) dx \right\}
\]

(cont.)
\[(\alpha - m_2) \int_{K_2}^{\infty} (x - K_2) g(x|n) \, dx \left\{ f(n) \, dn \right\}
\]

At this point we use our notation that \(\int_S g(x|n)f(n) \, dn = h(x)\), and write

\[C(K_1, K_2) = c_1 K_1 + c_2 (K_2 - K_1)\]

\[+ m_1 \int_0^{\infty} x h(x) \, dx + (m_2 - m_1) \int_{K_1}^{\infty} (x - K_1) h(x) \, dx\]

\[+ (\alpha - m_2) \int_{K_2}^{\infty} (x - K_2) h(x) \, dx\]  \hspace{1cm} (4-3)

Differentiating \(C\) with respect to \(K_1\) and \(K_2\) we get. 2

\[\frac{\partial C}{\partial K_1} = (c_1 - c_2) - (m_2 - m_1) \int_{K_1^*}^{\infty} h(x) \, dx = 0\]

\[\frac{\partial C}{\partial K_2} = c_2 - (\alpha - m_2) \int_{K_2^*}^{\infty} h(x) \, dx = 0\]

Using \(\int_{K_1}^{\infty} h(x) \, dx = 1 - H(K_1)\) we can write:

\[c_1 + m_1 [1 - H(K_1^*)] = c_2 + m_2 [1 - H(K_1^*)]\]  \hspace{1cm} (4-4)

\[c_2 + m_2 [1 - H(K_2^*)] = \alpha [1 - H(K_2^*)]\]  \hspace{1cm} (4-5)
These first order conditions imply that optimally, each type of capacity should be constructed up to the point where the fixed plus expected operating costs of the last unit of capacity are equal to the fixed plus expected operating costs of replacing that marginal capacity by the next highest ordered technology. In the case of peaking capacity the trade-off is between an additional unit of peaking capacity which will be used fraction \( [1-H(K_2)] \) of the year or additional reliance on the use of the shortage "technology" which has zero capital cost and variable cost \( \alpha \).\(^3\)

These investment rules can also be derived diagrammatically using what are called screening curves in the engineering literature. This technique emphasizes the fact that the minimum fraction of the year that each type of capacity should be used is independent of the pattern of demand throughout the year, and hence the load distribution curve. In Figure 4-1 the cost curves for each of the three technologies being considered are drawn on one set of axes. The intersections of the three cost curves indicate the optimal switch-over points between different technologies. Projecting these points down on the load distribution curve reveals the choice of generating technologies which will be optimal if the utility faces the pictured load distribution curve.

The investment rules derived here are formally the same as those derived by Boiteux (1949), Steiner (1957) and others for deterministic load duration curves. The are also similar to the investment rules derived with uncertainty by Meyer (1975), Crew and Kleindorfer (1976)
Figure 4-1a
Screening Curve

cost per kw per unit of capacity

\[ c_1 \]

\[ c_2 \]

\( \theta^* \)

\( \theta^* \)

\( \theta^* \)

\( \theta^* \)

fraction of period capacity is used

Figure 3-1b
Load Distribution Curve

\[ \theta = H(x|p) \]

\( k_2^* \)

\( K_2^* \)

\( k_1^* \)

\( K_1^* \)

\( \theta = \text{fraction of a period that demand is below level } x \)
and Turvey and Anderson (1977). The principal difference is that here we have collapsed multiple users, multiple demand periods and both long-term and short-term uncertainty into a single probability distribution for demand, \(H(x)\). As we will show, this simplifies the analysis of a number of interesting static and dynamic results.

IV.B. Optimal Pricing

In deriving optimal pricing rules, we first consider the case where all electricity is sold at a single price, \(p^0\). This conveniently allows us to focus on cases where price elasticities of demand are not constant within a single pricing period. Extensions to cases where electric utilities can price discriminate between different users (residential and industrial), different demand periods (day and night) or different product dimensions (peak use and total use) are deferred until Section IV.D below.\(^4\)

Using equations (3-6), (4-1) and the result derived in Appendix A, the objective function to be maximized in the case where a single price, \(p^0\), is charged for all electricity consumed can be written:

\[
U(k, p^0) = \mathbb{E}[UCS(p^0) + REV(p^0) - C(k, p^0)]
\]

\[
= \int_{p^0}^{\infty} D(p) \, dp + p^0 D(p^0) - C(k, p^0) \tag{4-6}
\]

It should be remembered that \(D(p)\) and \(C(k, p)\) are expected total demand and expected total costs, respectively. Differentiating with respect to \(p^0\),
\[
\frac{aU}{ap^o} = -D(p^o) + D(p^o) + p^o \frac{aD}{ap^o} - \frac{aC}{ap^o} \\
= p^o \frac{aD}{ap^o} - \frac{aC}{ap^o} = 0
\]

This can be written as:

\[
p^o = \frac{aC/ap^o}{aD/ap^o}
\]

(4-7)

This last expression shows that the optimal price should be the ratio of the change in total expected costs resulting from a price change divided by the change in total expected consumption resulting from the price change. The optimal pricing rule is that the price should equal the expected cost of serving the change in demand resulting from a small deviation of a price from its optimum. 5

If demand were constant throughout the demand period and there were no uncertainty, then the above pricing rule would simplify to say that prices should equal marginal costs. Because demand varies within a year, however, marginal costs will also vary, and the optimal price should be a weighted average of these marginal costs. The optimal weights for each marginal cost turn out to be the fractions of the change in demand generated at each marginal cost.

The manner in which costs change as prices are changed depends on whether capital costs are treated as fixed costs or whether investments can be reoptimized after a price change. This is equivalent to the question of whether short- or long-run marginal cost pricing is appropriate. 6 In this chapter we are assuming that pricing and
investment decisions are made simultaneously, so that long-run marginal costs are equivalent to short-run marginal costs. In the next chapter, however, sequential decisions are considered where short-run marginal costs are optimal. For this reason both methods of calculating marginal costs are considered.

Figure 4-2a illustrates how the change in costs would be calculated for a discrete price change in the case where capital investments are fixed. In the short run, the relevant change in costs is the change in usage of each existing technology (the horizontal areas in Figure 4-2a) multiplied by the operating cost of that technology.

In the long run, capacity can be reoptimized after the price change. In order to represent expected long-run costs it is useful to use a result alluded to by Boiteux (1949) and used by Turvey and Anderson (1977) and others, which is that the first order conditions for optimal investment can be substituted back into the total cost function C(K*,P) to find a convenient algebraic and graphical representation of total fixed, operating and shortage costs, C(ω*,P) which does not depend of K*. It is shown in Appendix B that if investments are always chosen optimally, then total expected generating plus shortage costs are

\[
C(\omega^*, P) = \int_{0}^{\omega^*_1} m_1 H^{-1}(\omega|P) \, d\omega \\
+ \int_{\omega^*_1}^{\omega^*_2} m_2 H^{-1}(\omega|P) \, d\omega
\]

(cont.)
Figure 4-2a
Change in Short Run Costs as a Result of a Price Increase of $\epsilon$

- Generating capacity held fixed

Figure 4-2b
Change in Long Run Costs as a Result of a Price Increase of $\epsilon$

- Generating capacity reoptimized after price change
\[ + \int_{\Theta^*_{2}}^{1} aH^{-1}(\theta|p) \, d\theta \quad (4.8) \]

This result is expressed diagramatically in Figure 4-2b. The vertical areas under the load distribution curve, multiplied by the appropriate operating cost or shortage cost sum up to the expected total of all generating plus shortage costs.\(^7\)

Using this notation for representing long-run costs, the change in long-run costs is the sum of the changes in vertical areas between the \(\Theta^*_{1}\)’s weighted by the appropriate operating costs. From a comparison of Figures 4-2a and 4-2b it should be evident that if capacity is chosen optimally then the expected short-run and long-run marginal costs of an infinitesimal price change will be equivalent.

IV.C Elasticities of Demand, Optimal Pricing and Profitability

The approach developed here facilitates the analysis of the effect of elasticities of demand on optimal pricing. We present here the intuitive content of pricing results which are proven rigorously in Appendix C. Consider first the case where price elasticities of demand are constant across time and across different realizations of the uncertain variable, \(n\).\(^8\) In this case the load distribution curve is shifted downward (upward) equiproportionally by price increases (decreases). Intuitively, and as shown by Figure 4-3a, total generating plus shortage costs will change by the same proportion as total demand, hence the optimal price is simply the average cost (including shortage costs) of serving that expected demand.\(^9\) This
Figure 4-3a
Constant Price Elasticity
Case: curve shifted down
equipropotionally

Figure 4-3b
Price Elasticity Decreasing
with Level of Demand: curve
shifted down proportionally
more at low levels

Figure 4-3c
Price Elasticity Increasing with
with Level of Demand: curve
shifted down proportionally
more at high levels
optimal price will be slightly higher than the price based on fixed plus expected operating costs divided by the expected demand which is served, the single price which is in principle the basis of most actual electric rates. The difference arises because the latter measure excludes expected social shortage costs from the cost base. The interesting result in this case is that if price elasticities of demand are constant, and if prices and investments are simultaneously set optimally, then total expected revenues will be greater than fixed costs plus expected operating costs, or in other words, expected profits will be positive.10

Consider next the case in which as prices increase the load distribution curve shifts down proportionately less at high demand levels than at low demand levels. This case, which is shown in figure 4-3b, would result if for any given realization of η, demand is less price elastic during high demand periods than during low demand periods. It would also result if price elasticities of demand, while constant throughout the demand period for any given realization of η, were lower for realizations of η associated with high levels of demand. In this case price changes will have a proportionately greater impact on base load costs than on shortage costs, the optimal price will tend to be lower than average expected costs, and expected profits will tend to be negative.11

The third case, in which as a price increases the load distribution curve shifts down proportionately more at high demand levels than at low demand levels is the symmetrically opposite case from the previous
case. The effect of a price change on total generating plus shortage costs is depicted in Figure 4-3c. Clearly the optimal price will be above average expected costs and expected profits will be positive.

Note that in deriving these results we have had to say essentially nothing about how the uncertain vector \( \eta \) affects the demand curve. Previous authors have focused their attention on whether uncertainty affects the demand curves additively (e.g., Brown and Johnson, 1968; Sherman and Visccher, 1978; Crew and Kleindorfer, 1976, 1979, pp. 69–84) or multiplicatively (Carlton, 1977; Turvey and Anderson, 1977; Crew and Kleindorfer, 1979, pp. 86–87). It is argued in Appendix C, however that the critical factor is not whether errors are additive or multiplicative, but rather how price elasticities of demand vary with different realizations of \( \eta \) (and within a given year). If errors enter in multiplicatively, then price elasticities remain constant for different realizations of \( \eta \). If errors are additive, then price elasticities decrease as \( \eta \) increases. With the further assumption that demand elasticities are constant over time for any given realization of \( \eta \), examples using multiplicative and additive errors reduce to cases of elasticities which are constant and decreasing, respectively, as \( \eta \) increases. Focusing on price elasticities of demand should prove more fruitful both theoretically and empirically than focusing on the specification of the error structure affecting demand.

IV.D Optimal Pricing with Two Prices

Thus far we have focused on the case where all electricity is sold
at a single price. Here we extend the analysis by looking at cases where there are two prices being charged, \( p_1 \) and \( p_2 \). These prices are allowed to be those charged for any two dimensions of electricity service. For example, the two prices could correspond to day and night periods, summer and winter use, industrial and residential users, or peak KW use and total KW-hr. use. The fully general case, where there are numerous different prices is readily generalized from the simple case considered here.

In Appendix A it is shown that if \( D_1 \) and \( D_2 \) are total expected consumption of units sold at prices \( p_1 \) and \( p_2 \) respectively, then the appropriate welfare measure can be written in the form.\(^{12}\)

\[
U(k, p_1, p_2) = \int_{p_1}^{\infty} D_1(p, p_2) \, dp + \int_{p_2}^{\infty} D_2(\infty, p) \, dp \\
+ p_1 D_1(p_1, p_2) + p_2 D_2(p_1, p_2) - C(k, p_1, p_2) \tag{4-9}
\]

Maximizing with respect to \( p_1 \) and \( p_2 \), we get

\[
p_1 \frac{\partial D_1}{\partial p_1} + p_2 \frac{\partial D_2}{\partial p_1} = \frac{\partial C}{\partial p_1} \tag{4-10}
\]

\[
p_1 \frac{\partial D_1}{\partial p_2} + p_2 \frac{\partial D_2}{\partial p_2} = \frac{\partial C}{\partial p_2} \tag{4-11}
\]

These two equations in two unknowns implicitly define optimal prices \( p_1 \) and \( p_2 \).\(^{13}\) Their conceptual content is readily seen by rewriting the first equation as
\[ p_1 = \frac{\frac{\partial C}{\partial p_1}}{\frac{\partial D_1}{\partial p_1}} - p_2 \frac{\frac{\partial D_2}{\partial p_1}}{\frac{\partial D_1}{\partial p_1}} \]  

(4-12)

This expression indicates that the change in revenue collected from \( D_2 \) as a result of an increase in \( p_1 \) should be subtracted off from the decrease in expected total costs (short or long run) before dividing this total by the decrease in \( D_1 \). Note that \( \frac{\partial C}{\partial p_1} \) also captures the cost consequences of the change in \( D_2 \). Optimal pricing should take into account not only the pattern of change within any demand category in response to a price change, but also between different demand categories.

Because the derivations are somewhat tedious, we again present here only the intuitive content of results which are shown rigorously in Appendix C. Consider momentarily the case where cross price elasticities of demand are all zero. In this case \( \frac{\partial D_2}{\partial p_1} = \frac{\partial D_1}{\partial p_2} = 0 \). Furthermore \( \frac{\partial D_1}{\partial p_1} = \frac{\partial D}{\partial p_1} \), i.e., the change in consumption of service (period or user) 1 is the same as the change in total demand. The above pricing rule reduces to the same formula as the rule derived for a single price. The optimal price, \( p_1 \), is above, equal to, or below expected average costs depending on whether the elasticity of total expected demand with respect to \( p_1 \) is increasing, constant or decreasing as the level of demand increases, i.e., as one moves up the load distribution curve. The profitability result derived in the isoelastic case (Case 1) of a single price generalizes to the case where there are many prices: if cross price elasticities are all zero and own price elasticities of demand are
constant (but not necessarily identical) for each user and each demand period over all possible realizations of the uncertain variable \( n \), and if prices and investments are both determined simultaneously, then expected profits will be (slightly) positive. Heuristically the justification for this result lies in the fact that if all own price elasticities are constant, then each user price or period price will be simply the expected average cost of meeting that demand. Since these expected average costs include social shortage costs, expected profits which exclude these costs will be positive.

These results continue to hold under certain restrictions even when cross price elasticities are nonzero. If when \( p_1 \) is changed the change in revenue collected from sales at \( p_2 \) is equal to the change in costs resulting from serving the increased sales at \( p_2 \), then the cross price effects of the change in \( p_1 \) will cancel. In the simple case where both cross price and own price elasticities are constant (though not necessarily identical) throughout each demand period for each user over all possible realizations of \( n \), then optimal prices will still result in (slightly) positive profits.\(^{14}\)

Thus far we have only considered cases where price elasticities are constant across different users and demand periods, however the approach adopted here facilitates the consideration of optimal prices when demand elasticities may vary both within and across different demand periods. Four different examples ignoring long-term uncertainty highlight the usefulness of our approach. Consider first a world where there are only two prices: a single industrial price, \( p_1 \), and a price
for all other users, $p_2$. (Note that in this case it is reasonable to assume $\frac{aD_2}{aP_1} = 0$, i.e. no cross price effect.) It is possible that although the industrial users' consumption pattern contributes heavily to peak and shortage costs, when its uniform price is raised, its primary response may be to reduce nighttime lighting, which affects only base load operating costs. The optimal industrial price in this case should be based only on the base load operating cost, not the shortage and peaking costs. This example emphasizes that optimal pricing may be very imperfectly related to the total or average costs of providing service to a user.\textsuperscript{15}

Consider next a utility with a price $p_1$ for nighttime electric heat and price $p_2$ for all other uses. Suppose that this utility is in a region where daytime heating contributes significantly to peak demand and hence shortage costs. Because $p_2$ is charged for many uses besides daytime heating, all of which will respond to changes in $p_2$, then $p_2$ is likely to be below the average cost of providing daytime heating. Lowering the nighttime heating price, $p_1$, will certainly increase nighttime, base load generating costs, but it will also lower daytime heating, which it has been argued are costing more than the price charged. The inability to charge the appropriate daytime heating price means that the optimal nighttime heating price, $p_1$, should be below even the operating cost of base load capacity. This example highlights the importance of considering both inter- and intra-period demand shifts.

At the opposite extreme, consider a utility with summer peaks which has reached an agreement with some customers to charge a price $p_1$ per
kilowatt on the ten hottest hours of the year when shortages are occurring, and a uniform price \( p_2 \) for all other times of the year. Suppose price elasticities of demand are such that \( p_2 \) is optimally close to the average cost of producing electricity at all non-peak hours. Raising \( p_1 \) will clearly reduce peak demand and hence social shortage costs, valued at \( \alpha \) per kilowatt per period. Raising \( p_1 \) may also reduce consumption at near peak hours (e.g., through improved insulation or air conditioning efficiency) when marginal generating costs are still above average generating costs, and hence \( p_2 \). As \( p_1 \) is raised, the generating cost savings during non-peak hours are greater than the loss of revenue during these hours. As a result the above pricing rule indicates that \( p_1 \) should be greater than \( \alpha \), i.e., greater than even the social cost of shortages. As in the previous example, this is a second-best pricing result arising because of the (assumed) inability to vary prices continuously with the level of demand. Together, this example and the previous example illustrate that if there are constraints on the number of different demand period prices which can be charged, then it is possible that welfare optimal prices may be below the operating cost of the least expensive technology or above the average social cost of shortages. With nonzero cross elasticities of demand, optimal prices are not bounded by these costs.

Finally, consider the case where \( p_1 \) is the premium charged for peak kw use by each user, \( p_2 \) a flat kw-hr charge for all usage. Raising \( p_1 \) will very likely change the peak use by different customers, but whether it affects total generating costs depends on the extent to
which these peak uses coincide. If there are many users and if their peak usages occur at different times, then flattening out individual firm peaks may have a very slight effect flattening out the aggregate load duration curve. In this example the expression derived above will indicate that a very low value for $p_1$ is optimal.

The results derived here are very general and can be used to consider a wide variety of pricing schemes. They differ from the conventional approach in that we represent price changes via changes in the shape of the load distribution curve, we have not had to specify how uncertainty enters into the demand functions, and we do not impose any restrictions on the own or cross elasticities of demand between different users, demand periods or quality dimensions. Perhaps the most interesting new result which appears from our approach is that if optimal prices and investments are chosen simultaneously, and if own and cross price elasticities of demand are constant (although not necessarily identical) for each user class and within each demand period over all possible realizations of $n$, then expected revenues will be slightly greater than expected generating costs. The difference arises only because the utility does not internalize the social cost of the power shortages resulting from insufficient capacity being available.

IV.E Effects of Increasing Uncertainty

The approach developed here enables us to consider what happens as uncertainty about future demand increases. Here we focus on the
effects of increasing uncertainty with a given lead time between the pricing and investment commitments and the eventual realization of uncertainty. We defer until the next chapter examining the impact of lengthening lead times.

There are several ways in which increasing uncertainty could be modeled using our framework. Specific functional forms for the conditional density function of $x$, $g(x \eta)$ or the marginal density functions of $x$ or $\eta$ ($h(x)$ and $f(\eta)$, respectively) could be assumed and the effect of increasing the variance of one of these probability distributions could be examined. This approach suffers from the dependence of the results on the specific functional form which is used, and the imprecision of variance as an indicator of the degree of uncertainty.

An alternative characterization of increasing uncertainty which does not require the specification of probability distributions is to use mean preserving spreads, an approach developed by Rothschild and Stiglitz (1970). The heuristic content of their definition of increasing risk or variability is that the cumulative distribution function $F^2$ is more risky than $F^1$ if both distributions have the same mean, but $F^2$ attaches greater weight to the tails of the distribution than $F^1$.

For the purposes of this paper the most convenient representation of increasing uncertainty is in terms of mean preserving spreads of the marginal distribution function of $x$, $H(x \ P)$. We therefore assume that each load distribution curve differs from a load duration curve with
the same mean by a sequence of mean preserving spreads. While increasing uncertainty about \( \eta \) might have an effect focused on either peak or low demand levels, we restrict our attention to mean preserving spreads which increase the probability both that demand is below low levels and above high levels. Figure 4-4 illustrates a load duration curve, \( G(x) \), together with two load distribution curves \( H^1(x) \) and \( H^2(x) \), corresponding to increasing uncertainty about \( \eta \) with these properties. We also note the empirical fact that fixed and operating costs are such that baseload capacity is only optimal for steady and hence low levels of demand, while peaking capacity is optimally used to serve even high and infrequent levels of demand.

If increasing uncertainty is characterized as we have depicted it in Figure 4-4, then we get three important implications about the effects of increasing uncertainty on investments, costs and pricing. If the probability that demand is below level \( K_1 \), i.e., \( H(K_1 P) \), increases as uncertainty increases, then \( K_1 \) must be decreased as uncertainty increases in order to hold this probability constant and satisfy the optimal investment rules derived above. Conversely if \( H(K_1 P) \) decreases as uncertainty increases then \( K_2 \) must be increased in order to hold \( H(K_2 P) \) constant. We thus have the important result that if prices are held fixed, then increasing uncertainty as described above decreases the optimal quantity of base load capacity and increases the optimal quantity of non base load and total capacity. This result was previously derived by Fuss and McFadden (1971) in a framework without prices or long-term uncertainty, however it appears
Figure 4-4

Effects of Increasing Uncertainty while holding prices fixed

\( x \) level of demand
\( \theta = H^2(x) \)

Increasing total capacity

\( \theta = H^1(x) \)

\( \theta = G(x) \)

decreasing base load capacity

\( m_2 \)

\( m_1 \)
to have escaped notice in the rest of the literature on utility investments under uncertainty. This may be because it is not readily apparent unless demand uncertainty is collapsed into a single load duration or distribution curve as is done by Fuss and McFadden and in this paper.

A second result which is apparent from our characterization of increasing uncertainty is that if prices are held constant, then total and average expected generating and social shortage costs increase as uncertainty increases. This is readily seen by comparing the vertical areas under the load distribution curves shown in Figure 4-4, which as shown in equation (4-8) is one way of representing total expected costs. This result was derived by Rothschild (1969) in a somewhat more general framework which did not restrict itself to fixed proportion production technologies. It is worth repeating here because again it appears to have escaped notice in the rest of the literature on utility investments under uncertainty.

If prices and investments must both be chosen simultaneously, then a third result follows directly from the second and from the optimal pricing rules, which are derived in Appendix C and discussed in Section IV.C. There it was shown that if price elasticities of demand are constant over time and over n then the optimal price is simply total expected costs divided by total expected demand. If all electricity is sold at a single price, and price elasticities of demand are constant over the year and either constant or increasing with realizations of n corresponding to high levels of demand, then optimal prices increase as
uncertainty increases. If price elasticities decrease with \( n \), then it is ambiguous as to whether the single price increases or decreases, although there remains a strong tendency for prices to increase. This result has obvious extensions to cases where there is more than a single price. If prices are increased then this will feed back upon investment decisions, further reducing baseload investments and mitigating the tendency to increase total generating capacity. Since the load distribution curves will also shift down as prices are increased, total expected costs may or may not increase as uncertainty increases.

It would be nice if more definitive statements about the effects of increasing uncertainty could be made, however we believe that the above statements are as precise as can be made without being substantially more explicit and less general about the probability distribution of \( n \) and the way in which \( n \) affects demand.
Footnotes to Chapter IV

1. It should be noted that since these shortage costs will not be born by the utility, some outside agent will be needed to influence investments to guarantee that these costs are considered.

2. It is easily verified that the matrix of second derivatives of $C$ is diagonal with elements $(m_{i+1} - m_i) g(K_i)$ down the diagonal $(m_{n+1} = 0)$. This matrix is therefore positive semidefinite over all possible values of $K_i$, $C(K)$ is globally convex, and the solution we have found must be a global minimum (possibly not unique in degenerate cases).

3. This result is, of course, not new. It is contained in Boiteux (1949) and is an essential feature of virtually all investment planning models. It is derived here for completeness and to facilitate subsequent analysis.

4. The approach we adopt here is the opposite of the traditional approach, where it is generally assumed that there are many different pricing periods, and the problem of choosing optimal prices is one of constraining prices over these many periods into a few demand periods. The resulting lagrange multipliers are difficult to implement on a practical basis, and this approach is not well suited for handling multiple user classes or product dimensions.

5. Unlike costs, which are globally convex in capital, $K$, nothing can be said in general about the properties of $U(K,P)$ in our very general formulation. We simply assume that $U(K,P)$ is concave in prices and ignore the possibility of multiple equilibria.

6. Note that here we are actually talking about the marginal cost of price changes (or precisely the marginal cost of the demand shifts induced by price changes) rather than the more conventional marginal cost of output changes. The long versus short run characterization is still appropriate.

7. Although as presented this result depends on our linearity assumption about $a$, the model could be extended to incorporate the three generalizations discussed previously. The first generalization, letting $a$ vary with the demand period, can be accommodated by using the vertical regions under the load distribution curves appropriate for each demand period, with different $a$ for each period. The second generalization, letting $a$ depend nonlinearly on the size of the shortage can be handled by replacing the region weighted by $a$ with a region weighted by the (increasing) first derivative of the loss function. The third generalization is readily accommodated by making $a$ and $g_i$ functions of the price. Although not as analytically convenient, the qualitative nature of our long versus short run
pricing results and profitability conclusions will continue to hold with these generalizations.

8. This will be true only if the demand function, \( X(P,n,t) \) can be written in the form \( x = x^1(P)x^2(n,t) \). See Appendix C.

9. Note that the pricing formula is in terms of the expected change in costs divided by the expected change in demand, which due to non-linearities will not be the same, in general, as the expected change in average costs.

10. If we define \( Z = \int_{K_n}^{\infty} (x - K_n)h(x)dx \), the expected total number of kilowatt-hours of unmet demand, the expected profits with optimal pricing and investments would be \( (\alpha - P_0)Z \). Empirically this is likely to be small relative to total revenue.

11. We can only say "tend to be negative" here because in the previous case, with constant price elasticities, optimal expected profits were shown to be slightly positive. Clearly there will be some small range with decreasing elasticities that will still have positive or zero profits, but with more pronounced elasticity decreases expected profits will be negative.

12. This is one of many possible line integral measures of expected consumer surplus. We assume the second derivatives of the indirect utility functions are symmetric so that all such line integrals are equivalent. See Appendix A.

13. Once again both short run and long run interpretations are possible.

14. Interestingly enough, as the number of demand periods increases, it appears that expected profits will decrease, the reason being that as more demand periods are allowed, prices during peak periods will rise toward \( \alpha \), and the expected size of shortages will decline, so that the expression derived in the footnote 35 for expected profits will decline.

15. If this result seems somewhat disconcerting, it should be noted that 1) in the long run, daytime conservation efforts may occur which should also be taken into account in \( P_1 \) and 2) substantial efficiency gains may be possible through the use of more prices, i.e. greater price discrimination.

16. An alternative formulation is to use mean preserving spreads of the marginal distribution of \( \eta \), \( F(\eta) \), and examine the expectation over \( \eta \) of our conditional load duration curve, \( G(XP,\eta) \). Our results can then be derived by assuming \( G(XP,\eta) \) is locally convex in \( \eta \) for low
values of $X$ and locally concave in $\eta$ for high values of $X$. Although useful for some applications, this approach is not as intuitive as the one used in the main text.

V. Dynamic Results in a One Period Capital Model

In this chapter a one period capital model is used to examine welfare optimal pricing and investment decisions in a dynamic setting where long-run demand uncertainty is gradually resolved over time. The effects of increasing lead times are first examined when pricing and investment decisions are made simultaneously. We then consider the case where pricing and investment decisions are made sequentially, with prices set either before or after all decisions are made. Finally, we examine the implications of sequential decisions in a simplified model where prices are set exogenously, but base load investment decisions must be made before peaking capacity decisions.

V.A Effect of Lead Times

Up until this point nothing has been said about the timing of investment and pricing decisions. Of course, in either a deterministic world or a world in which there is no resolution of uncertainty over time until it is actually realized, then as long as decisions are always made optimally neither the order in which decisions are made nor the timing matters. With uncertainty which is resolved slowly over time, however, both the order and timing of pricing and investment decisions become critical. This is because as time progresses and uncertainty is resolved, better information about the future improves the quality of the decisions that are made. Here we focus on the effect which lengthy lead times have on optimal decisions when all decisions are made simultaneously, deferring for now the interesting
cases where decisions are made sequentially.

Now that issues of timing are being considered, attention must be
given both to the time at which decisions are being made, and to the
time at which costs and demand will be realized. The following
notation is adopted. A preceding subscript is used to denote the time
at which a decision or forecast is being made, while a following
subscript is used to denote the time at which the outcome is actually
realized. Hence \( f_t(n_t) \) is the density function of the uncertain vector
\( n \) at time \( s \) for period \( t \). The load duration curve for period \( t \) becomes
\( G_t(x|n_t) \), while the load distribution curve at time \( s \) for period \( t \) can be
written:

\[
H_t(x) = \int_S G_t(x|n_t) f_t(n_t) \, dn_t. \tag{5-1}
\]

Virtually all of the other variables are also dated. However in this
chapter, since only the case of one period capital is being considered,
the subscripts on these other variables will be omitted, it always
being understood that the costs correspond to period \( t \) in each case.

In addition to dating the variables, some representation of the
information available at time \( s \) is needed. We assume that all of the
relevant new information gained in any period \( s \) is reflected in the
realization of \( n_s \). Since \( n \) is the only type of long-term uncertainty
allowed in our model this is both the simplest and the most natural
representation of the new information which becomes available each period.

Up until this point we have spoken of the load distribution curve as
if there were a single such curve for period \( t \). In a dynamic framework
where uncertainty is resolved slowly there will actually be a whole family of such curves: \( s_{-1}H_t, sH_t, s+1H_t, \ldots tH_t \), corresponding to demand in period \( t \). We can usefully express the load distribution curve at time \( s \) as a function conditional on \( n_s \), hence \( sH_t(x|n_s) \).

Note that \( tH_t(x|n_t) = G_t(x|n_t) \) i.e., the load distribution curve in the final period is simply the usual load duration curve. For the remainder of this paper we will often use \( H_t \) rather than \( G_t \), and not distinguish this curve as the load duration curve. Also note that there are several equivalent ways of representing a load distribution curve for a given set of prices, \( P \).

\[
\begin{align*}
    sH_t &= \int_S s_{+1}H_t(x|n_{s+1}) s_{s+1}(n_{s+1}) \, dn_{s+1} \\
    &= \int_S s_{+2}H_t(x|n_{s+2}) s_{s+2}(n_{s+2}) \, dn_{s+2} \\
    \quad \vdots \\
    &= \int_S tH_t(x|n_t) s_{t}(n_t) \, dn_t
\end{align*}
\]

The issue we wish to address here is how welfare optimal prices and investments will differ when they are based on \( s_{-1}H_t(x|P, n_{s-1}) \) rather \( sH_t(x|P, n_s) \), i.e., what is the impact of lengthening the lead time between pricing and investment decisions and the ultimate realization of the uncertain demand. In many respects the answer to this question is the same as the answer to the previously addressed question: what are the effects of increasing uncertainty. Clearly, in our model there
is greater uncertainty at time s-1 than at time s. The difference however, is that except for one or a few realizations of \( \eta_s \), the probability distributions represented by \( s_{-1}H_t \) and \( s_{+1}H_t \) will have different means. The two distributions cannot be readily analyzed in terms of mean preserving spreads.

In order to make qualitative statements about the effect of lead times on decision making, Figure 5-1 contrasts \( s_{-1}H_t(x|p^0, \eta_{s-1}) \) with the load distribution curves \( s_{+1}H_t(x|p^0, \eta_s) \) conditional on three different realizations of \( \eta_s \): \( \eta_s^0 \), \( \eta_s^1 \) and \( \eta_s^e = \int_s \eta_s s_{-1}f_s(\eta_s) \), and the same vector of prices \( p^0 \). For the same reason that we previously argued that the marginal distribution of a random variable (here \( s_{-1}H_t(x) \)) will in most interesting cases display a greater dispersion that its conditional distribution (here \( s_{+1}H_t(x|\eta_s) \)) we argue that \( s_{-1}H_t \) will have a greater dispersion than \( s_{+1}H_t \) for most if not all values of \( \eta_s \). Comparing our representation of total expected generating plus shortage costs via vertical areas in Figure 5-1 it should be clear that if our dispersion argument is true empirically and if prices are held constant, then having to make investment commitments one period earlier will result in
1) less reliance on base load capacity as a fraction of total capacity,
2) higher expected average generating costs and 3) larger expected capacity shortages.

If we allow prices to be chosen optimally in the two different years then these conclusions would have to be modified to incorporate price as well as uncertainty induced changes in the load distribution curves. Unless price elasticities of demand are strongly negatively
Figure 5-1

Effects of the Resolution of Uncertainty over time

Load Dist. Curve from the Perspective of Period $s-1$

$\theta = s_{Ht}(x|\eta_s)$

Three Possible Load Distribution Curves from the Perspective of period $s$

Fractions of a year that demand is below level $x$
correlated with \( n \), then optimal prices will also tend to be higher as lead times lengthen.\(^1\) Higher prices may magnify, diminish, or perhaps even reverse the three effects mentioned above, depending on price elasticities of demand. If they have a neutral effect on the shape of the load distribution curves, shifting them up and down equiproportionally, then the above three results will continue to hold.

Although the above results are based on empirical relationships between probability distributions that we have asserted but not shown to be true, we believe that they have important ramifications for empirical planning and simulation models. We note here two applications in which the failure to consider lead times results in inappropriate conclusions being reached.

In empirical investment planning models long-run uncertainty is frequently considered by studying three (or more) different demand growth scenarios: high, medium, and low growth trajectories are extrapolated and optimal investment strategies for each of these deterministic scenarios are calculated. Policy recommendations and investment plans are then usually based on the medium growth trajectory. The weakness of this approach is that these investment strategies are based on the assumption of zero lead times between investment commitments and the realization of demand. Once lead times are acknowledged then investments should be based on our load distribution curves: relatively less base load and more peaking capacity should be chosen, expected shortage costs should be higher, and prices will in general be higher than the deterministic, medium
growth trajectory would suggest.

The assumption of zero lead times has also been used in simulation models trying to determine the optimal reserve margins and loss of load probabilities for utilities with different demand and supply side characteristics. The most interesting conclusion of the 1978 Electric Power Research Institute (EPRI) study of the net costs of alternative reserve margins is how little difference the reserve margin appears to make over a fairly wide range (i.e., reserve margins of 15-30 percent). This conclusion may in part be due to the fact that they use a deterministic demand growth model, which is equivalent to choosing capacity with zero lead times. If it is recognized that even peaking units may require three years to construct, then the load distribution curves on which these calculations should be based will be more dispersed, expected shortage costs will be greater, and total costs will be more sensitive to reserve margins or the loss of load probabilities than when deterministic load duration curves are used.

V.B Sequential Pricing and Investment Decisions

Up until this point it has been assumed that all pricing and investment decisions are made simultaneously. As we have show, this implies short- and long-run marginal costs are equivalent. Here we examine the implications of making all investment decisions before or after all pricing decisions. We continue to use a one period capital model where uncertainty is gradually resolved over time. The case where investment decisions are themselves made sequentially over time
is considered in Section C.

Consider first the case where investment decisions are made after prices are set. Although not particularly realistic for electric utilities, the example might be relevant for other industries such as price regulated airlines. Specifically, assume that all investment commitments for period $t$ are made in period $t-1$ while prices must be set in period $t-2$. By period $t-1$, prices, $\bar{P}$, will be fixed for period $t$, so that the investment problem remains one of minimizing the cost of meeting $t-1H^t_t(x|\bar{P})$. Using our representation of total costs in terms of the inverse of the load distribution curve, $t-1H^{-1}_t$, i.e., in terms of vertical areas under the load distribution curve, we can write

$$
t-1C_t(\theta^*, \bar{P}, n_{t-1}) = \sum_{i=1}^{n} m_i \int_{\theta^*_{i-1}}^{\theta^*_i} t-1H^{-1}_t(\theta|\bar{P}, n_{t-1}) d\theta
+ a \int_{\theta^*_n}^{1} t-1H^{-1}_t(\theta|\bar{P}, n_{t-1}) d\theta \quad (5-5)
$$

Note that at time $t-1$, $t-1H^{-1}_t$ is a function of $n_{t-1}$, which is known at that time. From the perspective of period $t-2$, $n_{t-1}$ is unknown.

Measures of expected consumer surplus and revenues can be based on $t-2H^t_t$. The relevant measure of expected costs from the perspective of period $t-2$, however, is

$$
\int_S t-1C_t(\theta^*, \bar{P}, n_{t-1}) t-2f_{t-1}(n_{t-1}) dn_{t-1} \quad (5-6)
$$

Unfortunately these are not the same, in general, as $t-2C_t(\theta^*, \bar{P}, n_{t-2})$. 
This is so because in general

\[ \int_S t^{-1}H^{-1}(\theta_i^*, \bar{P}, n_{t-1}) t^{-2}f_{t-1}(n_{t-1})d\nu_{t-1} \neq t^{-2}H^{-1}(\theta_i^*, \bar{P}) \] (5-7)

The expected value of the inverse of a function is not, in general, equal to the inverse of the expected value.

Whereas our usual expectation of the load distribution curve \( E_n(H(x|P, n)) \) corresponds to the horizontal expectation (average) of the curves, the above expression, the expectation of \( H^{-1}(\theta|P, n) \), corresponds to the vertical expectation (average) of the curves. Figure 5-2 contrasts these two different expectations in the simple case where \( n_{t-1} \) can only take on two possible values, \( n_1 \) and \( n_2 \). Here and in the general case where \( n_{t-1} \) can take on many different values, the horizontal expectation will be more dispersed than the vertical expectation.

As we have previously argued in the case of increasing lead times, costs will be lower when investments can be deferred from period t-2 to period t-1. For most demand specifications optimal prices set in period t-2 will be lower when investments are made in period t-1 than in period t-2. We thus have the interesting result that optimal prices should not be based on the degree of uncertainty existing at the time prices are set, but rather should be based on the degree of uncertainty which it is expected will exist at the time that cost determining commitments (e.g., investments) are made. Prices can be set lower to reflect the prospective reduction in uncertainty which will occur before investments need be made.

The case where prices are set before investment decisions are made
Figure 5-2

Contrast between the horizontal and vertical expectations of the load distribution curves, corresponding to $E_\eta(H(x|\eta))$ and $E(H^{-1}(\theta|\eta))$. 

Fraction of a period that demand is below level $x$. 

x

level
of
demand
(kilowatts)

Horizontal expectation of Load Distribution Curves

$E_\eta(H(x|\eta))$

$H(x|\eta_1)$

$H(x|\eta_2)$

Vertical Expectation of Load Distribution Curves

$E_\eta(H^{-1}(\theta|\eta))$
has been considered at length because it illustrates how optimal behavior in one period should take into account the effect of this behavior on subsequent decisions and eventual outcomes. The opposite and considerably more interesting case will now be examined where investment decisions must be made before pricing decisions.

Consider the case where all investment decisions for period \( t \) are made in period \( t-2 \), with pricing decisions made in \( t-1 \). At time \( t-1 \) capacity is fixed so the pricing problem is one of maximizing expected consumer surplus, \( UCS(P) \), plus expected total revenue, \( REV(P) \) minus expected short run generating plus shortage costs, \( C(K,P) \). These short run costs will depend on the level of capacity investments previously made, \( K \). Since \( t-1H_t \) depends in part on \( n_{t-1} \) (which is known) then we can write the optimal price chosen in period \( t-1 \) for period \( t \) as \( t-1_p^*(K, n_{t-1}) \). Instead of simply minimizing the expected costs at time \( t-2 \), the firms objective function has become

\[
\max \int S \quad t-1UCS_t(p^*(K, n_{t-1}), n_{t-1}) + t-1REV_t(p^*(K, n_{t-1}), n_{t-1}) - t-1C_t(K, p^*(K, n_{t-1}), n_{t-1}) t-2^f(t-2(n_{t-1})dnt-1
\]

Differentiating with respect to \( K_i \), we get

\[
\frac{aU}{aK_i} = \int S \left[ \sum_{j=1}^J \left[ \frac{aUCS_j}{ap_j} \frac{ap_j^*}{aK_i} + \frac{aREV}{ap_j} \frac{ap_j^*}{aK_i} - \frac{aC}{ap_j} \frac{ap_j^*}{aK_i} \right] - \frac{aC}{aK_i} \right] t-2^f(t-1(n_{t-1})dnt-1 = 0
\]

(5-8)
The group of terms inside the summation sign cancel due to the first order conditions for optimal pricing, this leaves

\[ \int_S \frac{a_{t-1}C_t(K, n_{t-1})}{aK_i} t-2 f_{t-1}(n_{t-1})dn_{t-1} = 0 \quad (5-9) \]

This is not the same expression as minimizing long run costs from the perspective of period t-2, which would be:

\[ a_{t-2}C_t(K, n_{t-2}) = 0 \]

While the above expression implies minimize expected generating plus shortage costs, the previous expression implies that one should try to minimize capital costs plus the expected value of short run costs while taking into account the effect of uncertainty about \( n_{t-1} \) on \( P^* \). Due to nonlinearities, there is no reason to suppose expected short run operating costs will equal long run operating costs. In contrast with a single price chosen concurrently with investments, short-run marginal cost pricing allows price to be lowered for realizations of \( n \) corresponding to low levels of demand and increased for high demand scenarios. Relative to setting a fixed price for all scenarios, this reduces demand variability. A simple application of the results of Section IV.E implies a greater dependence on more capital-intensive technologies, and a reduced dependence on less intensive technologies. This suggests that relative to setting a fixed price for all scenarios, short-run marginal cost pricing in the last period increases desired capital intensity. Whether the capital choices are above or below
the levels chosen when there is no demand uncertainty is indeterminate.

Choosing investments conditional on prices subsequently being set equal to short-run marginal costs is sufficiently complex that it is difficult to say very much about optimal investment strategy. A number of important and interesting issues still remain. So that we may continue to investigate optimal investment strategy, for the remainder of this chapter and all of the next we assume pricing is based on long run rather than short run marginal costs. Short run marginal cost pricing is reconsidered in Chapter VII and dealt with explicitly in various simulations of Chapter VII.

Long run marginal cost pricing is not without some merit for consideration here anyway. Both Sweden and France have already adopted long run marginal cost pricing as the basis of their electricity pricing.\(^3\) Whatever its theoretical deficiencies as a welfare maximizing strategy the assumption is at least of practical relevance. Furthermore, short run marginal cost pricing will lead to greater price fluctuations than long run marginal cost pricing. With consumer misperceptions or imperfect insurance short run marginal cost pricing may not be optimal.

Assuming long run marginal cost pricing simplifies but does not eliminate the significance of sequential investment and pricing. Suppose investments are made in period \(t-2\) for period \(t\), while pricing decisions are made in-period \(t-1\). Whereas before the optimal price, \(P^*\), was a function of \(K\) and \(n_{t-1}\), it is now a function only of \(n_{t-1}\).
which may be thought of as the information available in \( t-1 \). The costs to be minimized from period \( t-2 \) are therefore

\[
t_{-2} C_{t-1}(K) = \int_S t_{-1} C_t(K, t_{-1} P^*_t(n_{t-1}), n_{t-1})
\]

\[
t_{-2} f_{t-1}(n_{t-1})dn_{t-1} \tag{5-10}
\]

This is substantially more tractable than the previous formulation in that the load duration curve is not a function of \( K \). However this formulation still has the property that the history of \( n \) before it reaches \( n_t \) has an effect on the outcome. \( n_{t-1} \) has an effect on \( P^* \) and hence on realized demand and costs. It is not enough to know the distribution of \( n_t \) and its effect on demand.

To eliminate this complication we make the further assumption that long run prices are themselves chosen in period \( t \), when \( n_t \) is known. This simplifies the analysis considerably since under this assumption the history of \( n \) between \( n_{t-1} \) and \( n_t \) is no longer relevant in the one period capital case.\(^4\) Notice that now the load distribution curve for period \( t \) can be simplified to

\[
G_t(x|P^*, n_t) = G_t(x|P^*(n_t), n_t)
\]

\[
= \bar{G}_t(x|n_t) \tag{5-11}
\]

The price argument can also be omitted from load distribution curves. We shall do so for the remainder of this chapter without any distinguishing change in notation for the new sense in which we are using load distribution curves.
V.C. Sequential Investment Decisions

The assumption that prices are based on long run marginal costs and are set once long term demand uncertainty is realized means that prices are exogenous. This transforms the problem of welfare maximization into the much simpler problem of cost minimization. Here the issues of sequential investment decisions due to differing lead times are explored.

Consider the case where there are only two generating technologies available, peak and base load capacity, where base load capacity \((K_1)\) must be chosen two periods in advance while peaking capacity \((k_2 = K_2 - K_1)\) need only be chosen one period in advance. We continue to assume that capital lasts only one period deferring until the next chapter the consideration of durable capital. The firm's objective function in period \(t-1\) for period \(t\) given that \(K_1\) has already been chosen can be written

\[
t_{-1}^C t(K_2, K_1)
\]

\[
= c_1 K_1 + c_2 (K_2 - K_1) + \int_0^\infty m_1 x t_{-1} h_t(x) dx
\]

\[
+ \int_{K_1}^\infty (m_2 - m_1) (x - K_1) t_{-1} h_t(x) dx
\]

\[
+ \int_{K_2}^\infty (\alpha - m_2) (x - K_2) t_{-1} h_t(x) dx
\]
Minimizing this with respect to $K_2$ yields

$$c_2 - (a - m_2)(1 - t_1 H_t(K_2)) = 0$$

(5-12)

This is of course the same result as that derived previously, (4-5). The point to emphasize here, however, is that the previous choice of $K_1$ does not affect the optimal level of $K_2$. $K_1$ itself can therefore be chosen in the previous period while disregarding its impact on $K_2$. The optimal choice of $K_1$ can be written

$$(c_1 - c_2) - (m_2 - m_1) [1 - t_2 H_t(K_1)] = 0$$

The only distinguishing feature of these two expressions from the previous static results is that the two investment rules are based on two different load distribution curves reflecting the different amounts of information which are available at different times. Figure 5-3 illustrates a possible sequence of optimal investments. In period $t-2$ base load capacity $k_{t-2}^*$ is chosen based on $t_2 H_t$. By period $t-1$ part of the uncertainty has been resolved, so that $t_1 H_t$ will in general have less dispersion than $t_2 H_t$. In period $t-1$ peaking capacity $k_2$ should be chosen so that total base load plus peaking capacity is equal to $K_2^*$. Notice that at $t-1$, in the figure drawn $K_1^*$ is already revealed to be less than would be optimal from the perspective of period $t-1$. Since it is too late to add base load capacity, the discrepancy must be made up by peaking capacity. Finally in period $t$, all long term uncertainty is resolved and the load duration curve $G_t(x | n_t)$ is realized. In the sequence pictured it is revealed 	extit{ex post} that too little base load and too much total capacity were installed. A different
An Example of Sequential Investment Decisions

\( K_1^* \) chosen at time \( t-2 \) based on \( t-2H_t(x|\eta_{t-2}) \) with \( \eta_{t-2} \) known.

\( K_2^* \) chosen at time \( t-1 \) based on \( t-1H_t(x|\eta_{t-1}) \) with \( \eta_{t-1} \) and \( \eta_{t-2} \) known.

\( G_t(x|\eta_t) \) realized at time \( t \) when \( \eta_t \) is known.

\[ \theta = t-2H_t(x|\eta_{t-2}) \]

\[ \theta = t-1H_t(x|\eta_{t-1}) \]

\[ \theta = G_t(x|\eta_t) \]

fraction of a year that demand is below level \( x \)
sequence of π's could have led to a different outcome, however.

It has been shown here that even though investments are made sequentially, the same static investment rules derived previously may be used, the only distinction being that the load distribution curve which is used must be continually updated to reflect the resolution of uncertainty over time. In the one period capital model considered here the load distribution curve at any time contains all of the relevant information needed to make investment decisions. Unfortunately this is no longer the case when capital is durable, for then capacity decisions in one period influence decisions in subsequent periods. With durable capital one needs to be concerned about the information which will be available in each subsequent period.
Footnotes to Chapter V

1. This result follows from our previous result that for constant elasticities of demand, prices should optimally be set equal to average costs which we have argued will increase as uncertainty increases.

2. One problem with this expression is that even if \( \frac{a_{t-1}H_t}{aK_i} \) is non-negative, here is no guarantee that \( \frac{d_{t-1}H_t}{dK_i} \) will be non-negative, since the price effect may outweigh the direct effect. This implies that the objective function may not be globally convex, as in the simple case, so that local extrema are possible. There may be multiple solutions to \( K^* \) which need to be searched for the global optimum.


4. This is no longer true with durable capital, as Section VI makes clear.

5. The case where \( K_2 \) is constrained by the size of \( K_1 \) is ignored; however, this type of possibility is considered briefly when we examine the dynamic, durable capital case.
VI. Dynamic Results with Durable Capital

In this chapter the implications of perfectly durable capital are examined. It is argued that if demand is growing sufficiently rapidly, then the static rules derived above continue to hold. In the case of a single generating technology it is shown that if capacity is perfectly durable and illiquid, then a firm should optimally choose less of it than would be chosen if the capacity were liquid.

IV.A. Rapidly Growing Demand Case

Continuing the assumptions of the preceding chapter, it is assumed that pricing is based on long-run marginal costs instead of short run costs, so that prices are independent of capacity choices. This means that the problem of cost minimization can be considered instead of the broader issue of welfare maximization. Since costs in more than one period will now be considered, we will now date all variables, continuing to use a preceding subscript for the period at which decisions are made and a following subscript for the time at which outcomes are realized. It is assumed that all types of capacity have lead times $t - s$, so that $t$ related subscripts refer to operation dates, $s$ to commitment dates. A constant discount factor, $\beta = 1/(1+r)$, will be used.

Continuing to let $sC_t$ be total fixed, operating and shortage costs expected for period $t$ from the perspective of period $s$, we define $sV_t$ to be the discounted expected value of all $sC_t$, taking into account the fact that capital investments are not all made simultaneously, but rather are
made sequentially for different periods. In the cost minimization problem which we have formulated, both $sV_t$ and $sC_t$ are functions of previous period commitments for capital levels, $K_{t-1}$, current capital choices, $K_t$, and the information available in period $s$, $n_s$. The cost minimization problem from the perspective of period $s$ can then be set up as a forward recursive dynamic programming problem of the form

$$sV_t(K_t, K_{t-1}, n_s) = sC_t(K_t, K_{t-1}, n_s)$$

$$+ \beta sE_{s+1}[s+1V_{t+1}(K_{t+1}^{**}, K_t, n_{s+1})]$$

where $K_{t-1} \leq K_t \leq K_{t-1}^{**}$. The double asterisks on $K^{**}$ are used to indicate that this vector of capacity will subsequently be chosen optimally using the dynamic investment rules derived here.

The central issue with durable capital is the fact that decisions made in any period $s$ constrain the choices which can be made in all subsequent periods, $s+1, s+2...$. Under the present assumptions, once investments $K$ are made, i.e. once one is committed to the fixed cost of renting $K$ forever, it is no longer possible to opt to rent less $K$ in subsequent periods. This irreversibility of investment decisions causes one to invest in less capacity than would be purchased if investments could be costlessly reversed or rental contracts cancelled. This result has been labeled the "irreversibility effect" and has been considered previously by Henry (1973, 1974), Arrow and Fisher (1974) Freixas and Laffont (1979) and Bernanke (1979). The contribution here is that we examine this effect in the context of electric utilities, we use a continuous representation of demand
uncertainty, and we subsequently employ our results in our simulation model.

Before examining the impact of the irreversibility effect on utility investments, it is useful to examine first the case where demand is growing very rapidly, so that new capacity of each type is added in every period.\footnote{1} If this is true, then capacity decisions in period $s$ never constrain decisions in period $s+1$. The only effect which an additional unit of capacity in period $t$ has on costs in period $t+1$ is its reduction of capital costs. It will be shown that the relevant cost of capital is simply the rental cost of each type of capacity, and once this change is made the static investment rules derived earlier continue to hold.

If capacity decisions for period $t$ must be made at time $s$, then the expected costs incurred for period $t$, $sC_t$, can be written (using the conventions $K_0 = 0$, $K_{n+1} = \infty$, $c_{n+1} = 0$, $m_0 = 0$, $m_{n+1} = a$):

$$sC_t(K_t, K_{t-1}, n_s) = \sum_{i=1}^{n+1} c_{i,t}[(K_{i,t} - K_{i-1,t}) - (K_{i-1,t} + K_{i-1,t-1})]$$

$$+ \int_{K_{i-1,t}}^{K_{i,t}} (m_{i,t} - m_{i-1,t})(x - K_{i-1,t})_x h_t(x|n_s)dx \quad (6-2)$$

Differentiating with respect to $K_{i,t}$ and $K_{i,t-1}$ and simplifying, one gets

$$\frac{\partial sC_t}{\partial K_{i,t}} = (c_{i,t} - c_{i+1,t}) - (m_{i+1,t} - m_{i,t})[1 - sH_t(K_{i,t})] \quad (6-3)$$

$$\frac{\partial sC_t}{\partial K_{i,t-1}} = -(c_{i,t} - c_{i+1,t}) \quad (6-4)$$
Since it is assumed here that new capacity of each type is added every period, then $K^{**}_{t+1}$, the optimal vector of capacity choices for period $t+1$ chosen at $s+1$, is independent of $K^{*}_t$. The only effect of $K^{*}_t$ on future costs is a reduction of capital costs. Hence

$$\frac{a_{s+1}V_{t+1}}{a_{K_{i,t}}} = \frac{a_{s+1}C_{t+1}}{a_{K_{i,t}}} = -(c_{i,t+1} - c_{i+1,t+1})$$

We are interested in finding an expression for $K^{**}_t$. Maximizing $sV_t$ over $K_t$ we get

$$\frac{a_sV_t}{a_{K_{i,t}}} = \frac{a_sC_t}{a_{K_{i,t}}} + \beta \frac{a_sE_{s+1}(s_{i+1}C_{t+1})}{a_{K_{i,t}}}$$

$$= (c_{i,t} - c_{i+1,t}) - (m_{i+1,t} - m_{i,t})[1 - sH_t(K^{**}_i,t')]$$

$$- \beta sE_{s+1}(c_{i,t+1} - c_{i+1,t})$$

$$= [(c_{i,t} - \beta c_{i,t+1}) - (c_{i+1,t} - \beta c_{i+1,t+1})]$$

$$- (m_{i+1,t} - m_{i,t})[1 - sH_t(K^{**}_i,t')]$$

$$= 0 \quad (6-5)$$

This last expression is identical to the results, (4-4) and (4-5), derived in the one period capital case only here, instead of $(c_i - c_{i+1})$ we have $[(c_{i,t} - \beta c_{i,t+1}) - (c_{i+1,t} - \beta c_{i+1,t+1})]$. Each of the two terms in this expression is the difference in discounted costs between an additional unit of capacity in period $t$ rather than in period $t+1$.

This is simply the rental value of capacity in period $t$, its one period capital cost. The important and not too surprising result here is that if demand is growing sufficiently rapidly that one never regrets the
choice of capacity the following period, then the static, one period
investment rules derived previously continue to hold even with durable
capital.

VI.B. The Irreversibility Effect

Having considered the rapidly growing demand case, we consider here
the more realistic case where demand grows slowly; so that as a result
of an unexpected slowdown in demand growth new capacity may not be
added in every period. We derive and interpret here what has been
found to be a very general result, namely that with durable capital and
uncertain future demand it is in general optimal to choose less
capacity than a static investment model using rental costs would
suggest. The results derived here are subsequently used in the
simulation model of Chapter VII.

To simplify the analysis $\eta$ is assumed to be a scalar random
variable, with high realizations of $\eta$ corresponding to high demand
levels. A single generating technology is assumed to be available.
We can thus dispense with the plant type subscripts and speak of
$c_t$ and $m_t$ as the capital and operating costs, respectively, per unit of
capacity, $K_t$, in period $t$. Expected costs in period $t$ from the
perspective of $s$ can now be written as

$$sC_t(K_t, K_{t-1}, n_s) = c_t(K_t - K_{t-1})$$

$$+ m_t \int_0^{\infty} x s h_t(x, n_s) \, dx + (a_t - m_t) \int_{K_t}^{\infty} (x - K_t) s h_t(x, n_s) \, dx \quad (6-6)$$

Unlike in the rapidly growing demand case where the utility is always
adding new capacity, with slowly growing demand the utility needs to be concerned about how choices made in period s will constrain choices in period s+1. Clearly, the more capacity which is purchased at time s, the more likely the utility will find itself constrained at s+1. We may formalize this by defining a non-decreasing function $\delta^{s+1}(K_t)$ with the property that if $n_{s+1} < \delta^{s+1}(K_t)$, which corresponds to unfavorable information about future demand levels, then the utility will find itself constrained to hold $K_{t+1} = K_t > K_{t+1}^{**}$. If $n_{s+1} > \delta^{s+1}(K_t)$, which corresponds to favorable information about future demand, then the utility will choose $K_{t+1} = K_{t+1}^{**} > K_t$, i.e., the utility is unconstrained.² Note that if $n_{s+1} = \delta^{s+1}(K_t)$ then $K_{t+1}^{**} = K_t$.

From the perspective of period s, expected costs in period t and thereafter can be written as

$$s^V_t(K_t, K_{t-1}, n_s) = s^C_t(K_t, K_{t-1}, n_s)$$

$$+ \beta \int_{-\infty}^{\delta^{s+1}(K_t)} s^V_{t+1}(K_t, K_t, n_{s+1}) s^{f_{s+1}}(n_{s+1}) dn_{s+1}$$

$$+ \beta \int_{\delta^{s+1}(K_t)}^{\infty} s^V_{t+1}(K_{t+1}, K_t, n_{s+1}) s^{f_{s+1}}(n_{s+1}) dn_{s+1}$$

(6-7)

where $K_{t-1} \leq K_t \leq K_{t+1}^{**}$.

We are interested in finding an optimal rule for choosing $K_t^{**}$ in the case where the firm is unconstrained (the constrained case not being very interesting). Differentiating $s^V_t$ with respect to $K_t$ and setting equal to zero:
\[
\frac{a_s V_t}{a K_t} = \frac{a_s C_t}{a K_t} + \beta \int_{-\infty}^{\delta^{s+1}} \frac{a_{s+1} V_{t+1}(K_t, K_t, \eta_{s+1})}{a K_t} s f_{s+1}(n_{s+1}) d\eta_{s+1} \\
+ \beta \int_{\delta^{s+1}}^{\infty} \frac{a_{s+1} V_{t+1}(K_{t+1}, K_t, \eta_{s+1})}{a K_t} s f_{s+1}(n_{s+1}) d\eta_{s+1} \\
+ \beta \frac{a\delta^{s+1}}{a K_t} s+1 V_{t+1}(K_t, K_t, \delta^{s+1}) f(\delta^{s+1}) \\
- \beta \frac{a\delta^{s+1}}{a K_t} s+1 V_{t+1}(K_{t+1}, K_t, \delta^{s+1}) f(\delta^{s+1}) \\
= 0
\]

(6-8)

The last two terms cancel since \( K_t = K_{t+1}^{**} \) if \( n_{s+1} = \delta^{s+1}(K_t) \). It is easily seen that

\[
\frac{a_s C_t}{a K_t} = c_t - (a_t - m_t)[1 - s H_t(K_t^{**})]
\]

and

\[
\frac{a_{s+1} V_{t+1}(K_{t+1}^{**}, K_t, \eta_{s+1})}{a K_t} = \frac{a_{s+1} C_{t+1}}{a K_t} = -c_{t+1} \quad \text{for} \quad n_{s+1} > \delta^{s+1}(K_t)
\]

Substituting these in the first order condition above and rearranging gives us the following condition:

\[
\frac{a_s V_t}{a K_t} = c_t - \beta c_{t+1} - (a_t - m_t) [1 - s H_t(K_t^{**})]
\]
\[ + \beta \int_{-\infty}^{\delta^{s+1}(K_t^{**})} \left[ c_{t+1} - \frac{a_{s+1} V_{t+1}(K_t^{**}, K^{***}, n_{s+1})}{a K_t} \right] s f_{s+1}(n_{s+1}) dn_{s+1} \]

\[ = 0 \quad (6-9) \]

The first half of the above expression is exactly the same as in the previous model when demand was growing rapidly. As argued above since capital lasts for more than one period, the relevant capital cost is the incremental cost of bringing an additional unit of capacity on line at time \( t \) instead of \( t+1 \). This is exactly \( c_t - \beta c_{t+1} \), which can also be thought of as the rental cost of capital. Offsetting this capital cost is the familiar shortage cost savings \( (a_t - \alpha) [1 - s H_t(K_t^{**})] \) which results when an additional unit of capacity is added. If these were the only two terms in the equation, then investments could be chosen using the same static investment rules developed earlier, where the rental cost of capital is used in place of the one year purchase cost.

The new term in the equation is the final expression. This expression represents the increase in total costs that may result in those cases where \( K_t \) constrains \( K_{t+1} \) to be greater than it would optimally be chosen if unconstrained. Notice that this can only arise if \( n_{s+1} \) is less than \( \delta^{s+1}(K_t) \).

The sign of the final expression determines whether optimal dynamic investment should be greater or less than the static rule suggests. The assumptions we have made require that the term be positive, as can be seen via the following argument. Suppose \( K_t^{**} \) satisfies the static
investment rule, so that the first two terms vanish. The sign of
\[
\frac{a_s V_t(K_t^*, K_{t-1}, n_s)}{\partial K_t}
\]
will therefore be the same as the sign of

the last term. By assumption, \(n_{s+1} < \delta^{s+1}(K_t)\) implies that the
firms optimal plan when confronted with \((K_t^*, n_{s+1})\) is not to
build any new capacity. This must mean that the capital cost of new
capacity \(c_{t+1}\) is greater than expected period \(t+1\) shortage cost
savings plus future capital cost and shortage savings. In short, capital
in period \(t+1\) when constrained is valued less than \(c_{t+1}\). This means

that
\[
\frac{a_{s+1} V_{t+1}(K_t^*, K_t^*, n_{s+1})}{\partial K_t}
\]
is less than \(c_{t+1}\) for \(n_{s+1} < \delta^{s+1}(K_t^*)\).

Hence the final term is positive, and \(\frac{a_s V_t}{\partial K_t}(K_t^*, K_{t-1}, n_s)\) must also be
positive. Expected costs will be less if capacity is reduced below \(K_t^*\),
the capacity based on static optimization rules.

We have shown in a model with exogenous prices and a single type of
generating capacity that if capital is perfectly durable, and if
uncertainty is resolved slowly over time, then utilities should invest
in less generating capacity than is suggested by considering
investments in a static model using rental costs as the one period
capital cost. A similar type of result could be shown in the case
where capital has a finite life, only then the rental cost would have
to include depreciation and replacement costs as well.

The same result will also hold with more than one type of
generating capacity. In choosing base load capacity the above direct
effect will continue to hold, and there will be the added effect that increasing $K_1$ also constrains the choice of $K_2 = K_1 + k_2$ and higher order generating alternatives. A similar argument will hold for $K_2$, etc.

All of the $K_i$ should be biased downward by the irreversibility effect. It is worth noting, however that although all levels of cumulative capacity, $K_i$, will be biased downward, we can say nothing in general about the capacity of individual generating types other than $k_1$, for instance, $k_2 = K_2 - K_1$. It is also worth noting that this irreversibility effect will continue to be significant even when different types of generating capacity require different lead times.
Footnotes to Chapter VI

1. This might be true, for instance, if the probability of $n$ declining, is zero, and costs $c_i$ and $m_i$, for all $i$, are not changing so rapidly that the price effects on demand outweigh the growth effects of $n$. Further assumptions are needed to guarantee not only that $s+1H_{t+1}$ will always lie above $H_t$, but also that the former will be more dispersed than the latter. To be rigorous substantially more would have to be said about the behavior of demand over time.

2. In the more general case where $n$ is a vector of random variables, we could define $\delta^{s+1}(K)$ to be a subset of $S$ and speak of $n_{s+1} \in \delta^{s+1}(K)$. 
VII. Extensions of the Model

Up until this point we have examined only exogenous demand uncertainty. In this chapter we extend the model in order to explore the significance of uncertain input prices (i.e., capital and operating costs), interfuel substitutability, average cost pricing and risk aversion. Throughout this chapter we return to the assumptions that all investments are made simultaneously, a fixed lead time before uncertainty is resolved and prices are set. To simplify exposition we also assume that all electricity is sold at a single price, and there are only two generation technologies available: base load and peaking capacity.

VII.A Uncertain Input Prices

Uncertainty about input prices has been examined by Fuss and McFadden (1971), Stewart (1978) and Perrakis (1980). The general finding has been that with risk neutrality, variable proportions technologies, non-random output price and quantity, uncertainty results in relatively greater use of the riskless inputs than is used in the certainty case with input prices set equal to their expectations.¹

If the more risky input prices are those of variable inputs, this means that uncertainty results in relatively greater reliance on the more capital-intensive technologies. Note that this is the opposite effect to that of output uncertainty, which we (and others) have shown reduces reliance on more capital-intensive technologies. Perrakis (1980) demonstrates that with both input cost and output uncertainty,
the direction of the distortion in input choices is indeterminate, and depends on the cost and demand functions. By making use of our assumptions that each generating technology is characterized by fixed proportions we are able to derive conditions under which input price uncertainty may increase reliance on the more risky technologies.

Suppose that all capital and operating costs now depend on the random vector \( \eta \), so that we can write \( c_i(\eta) \) and \( m_i(\eta) \), \( i = 1,2 \). (In line with our previous treatment of shortage costs, \( (\alpha) \), as a type of operating cost we could also treat \( \alpha \) as uncertain, however we choose not to here for simplicity.) Let \( \bar{c}_i \) and \( \bar{m}_i \) be mean capital and operating costs, respectively, for technology \( i \), so that for instance we can write

\[
\bar{m}_i = \int_S m_i(\eta) f(\eta) \, d\eta
\]

(7-1)

We are interested in finding the optimal choices of base load capacity, \( k_1 \), and peaking capacity \( k_2 \), given uncertainty about both the level of demand and input prices \( c_i(\eta) \) and \( m_i(\eta) \), \( i = 1,2 \). As we have done previously, it is more convenient to work with cumulative capacities, \( K_i \), and hence we parameterize the choices in terms of \( K_1 = k_1 \), and \( K_2 = k_1 + k_2 \).

Now that there is input price uncertainty, output prices should be adjusted to reflect not only capacity choices and the direct effect of \( \eta \) on demand, but also to reflect changing cost parameters. Let the optimal price be \( P^*(K,\eta) \) chosen after all \( \eta \) is known. We can still find the optimal choice of capacity \( K_i \) by the solution of equation (5-8) in Chapter 5.
\[
\frac{aU}{aK_i} = \int_S \sum_{j=1}^J \left( \frac{aUCS_j}{ap_j} \frac{ap^*}{aK_i} + \frac{aREV}{ap_j} \frac{ap^*}{aK_i} - \frac{aC}{ap_j} \frac{ap^*}{aK_i} \right) - \frac{aC}{aK_i} f(n) dn = 0
\]

(7-2)

If prices are set in the socially optimal manner conditional on \( n \), the group of terms inside the square brackets will cancel, yielding

\[
\int_S \frac{aC(K, P(K^*, n), n))}{aK_i} f(n) dn = 0
\]

(7-3)

Since we are looking forward a fixed lead time while choosing capacity, we ignore the time subscripts. Total expected costs with uncertain fuel costs can then be written

\[
C(K) = \int_S \left[ c_1(n)K_1 + c_2(n)(K_2 - K_1) + m_1(n) \int_0^\infty x g(x|P(K, n), n) dx 
+ (m_2(n) - m_1(n)) \int_{K_1}^\infty (x - K_1) g(x|P(K, n), n) dx 
+ (\alpha - m_2(n)) \int_{K_2}^\infty (x - K_2) g(x|P(K, n), n) dx \right] f(n) dn
\]

(7-4)

Differentiating with respect to \( K_1 \) and \( K_2 \) yields equations (7-5) and (7-6)

\[
\frac{aC}{aK_1} = \int_S \left[ (c_1(n) - c_2(n)) - (m_2(n) - m_1(n)) \int_{K_1}^\infty g(x|P(K, n), n) dx \right] f(n) dn
\]

(7-5)
\[
\frac{3C}{aK_1} = \int_{S} \left[ (c_2(n) - (a - m_2(n))) \int_{K_2}^{\infty} g(x|P(K, n) n) dx \right] f(n) dn \tag{7-6}
\]

Using \( \bar{c}_i \) and \( \bar{m}_i \) (i = 1,2) as well as the notation

\[
\int_{S} g(x|P(K, n), n) f(n) dn = 1 - H(K),
\]

these can be rewritten as

\[
\bar{c}_1 + \bar{m}_1 [1 - H(K_1)] + \int_{S} [m_1(n) - \bar{m}_1] [1 - G(K_1|P(K, n), n)] f(n) dn
\]

\[
= \bar{c}_2 + \bar{m}_2 [1 - H(K_1)] + \int_{S} [m_2(n) - \bar{m}_2] [1 - G(K_2|P(K, n), n)] f(n) dn
\tag{7-7}
\]

\[
\bar{c}_2 + \bar{m}_2 [1 - H(K_2)] + \int_{S} [m_2(n) - \bar{m}_2] [1 - G(K_2|P(K, n), n)] f(n) dn
\]

\[
= a[1 - H(K_2)] \tag{7-8}
\]

These equations are very similar to equations (4-4) and (4-5) of Chapter 4, only they use mean capital and operating costs, and three additional terms have appeared. These additional terms are the covariance of each operating cost, \( m_i \), with the fractions of the year demand is above the levels of capacity \( K_1 \) and \( K_2 \). Although by no means identical, these covariance terms are very closely related to the covariances between total demand and operating costs. High levels of demand will almost certainly be positively correlated with high values of \( [1 - G(K_i|P(K, n), n)] \), hence a positive covariance between \( m_i \) and \( [1 - G(K_i|P(K, n), n)] \) strongly suggests a positive covariance between operating costs and demand.

Two results follow immediately from equations (7-7) and (7-8).
First, in a one period capital framework with prices set equal to short-run marginal costs, capital cost uncertainty does not affect optimal capacity choices by a risk neutral utility. The same choices will be made based on the mean realizations of capital costs, $\bar{c}_i$, as will be made by considering the entire probability distribution of capital costs, $c_i(n)$. This result is not surprising: by the time all capital cost uncertainty is resolved, capital costs are fixed costs, and hence irrelevant for marginal cost pricing decisions.  

A second and more interesting result follows for operating costs, since these affect prices and hence the level of demand. Suppose that there is only one operating cost, $m_i$, which is uncertain. The rigorous statement which can be made is the following. If operating costs of only one type of generating capacity are positively (negatively) correlated with the fraction of the year that that capacity is used at full capacity then less (more) of that capacity should be constructed than is indicated by using expected fuel costs. The same is true if operating costs are negatively (positively, respectively) correlated with the fraction of the year that the capacity is not used at all. A more intuitive but somewhat less precise statement is that one should rely more heavily on generating technologies whose operating cost is negatively correlated demand and less heavily on generating technologies where operating costs are positively correlated with demand.

This result may be clarified through an example where pricing effects are ignored. Suppose $m_2(n)$, i.e., peaking fuel costs, are the
only uncertain input price, and there are only two equally likely realizations of \( n \), \( n^1 > n^2 \). Suppose base and peaking capacity are chosen while ignoring fuel cost uncertainty, using the mean fuel cost \( m_2 \). The two load duration curves \( G^1 \) and \( G^2 \) corresponding to \( n^1 \) and \( n^2 \) respectively and the implied load distribution curve \( H = \frac{1}{2} G^1 + \frac{1}{2} G^2 \) are shown in Figure 7-1 together with the choices \( K^1 \) and \( K^2 \). If \( m_2(n) \) were uncorrelated with \( n^1 \) and \( n^2 \), then these would be the optimal choices of \( K^1 \) and \( K^2 \). If instead \( m(n) \) is positively correlated with \( n \), then high fuel costs correspond to those times when peaking capacity is used the most, and unused the least. Since peaking capacity is costing the most when it is most needed, expected total costs (including shortage costs) can be reduced by relying less on peaking capacity and more on base load and shortages. Clearly, the opposite holds when peaking fuel costs are negatively correlated with \( G(x|P(K, n), n) \).

The above example was constructed while ignoring pricing effects. With marginal cost pricing, higher fuel costs will tend to be reflected in higher prices, and hence lower demand. Unless fuel costs are positively correlated with uncertain variables affecting the level of demand, this pricing effect will automatically tend to create a negative correlation between fuel costs and \( G(x|P(K, n), n) \), and hence one should optimally invest more in those types of generating capacity most responsible for pricing fluctuations. This is a very intuitive result.

Whether or not particular probability distributions of fuel costs are positively or negatively correlated with the probability
Figure 7-1

Capacity Choices Ignoring Operating Cost Uncertainty

Load Duration Curve
\[ G^1 = G(x|P(K, \eta^1), \eta^1) \]

Load Distribution Curve
\[ H(x) = \frac{1}{2} G^1 + \frac{1}{2} G^2 \]

shortages

peaking capacity \( (k_2) \)

base load capacity \( (k_1) \)

\[ x \]

fraction of the year that demand is below level \( x \)
distributions of future electricity demand may depend on many factors. Electricity is a substitute for many other fuels, therefore one might expect to find a positive correlation. Electrical generation is also a major purchaser of many of these same fuels, which would further tend to suggest the correlations may be positive.

It is important to realize that the relevant correlations are not the \textit{ex post} or realized correlations between fuel costs and demand, but rather the correlations in the \textit{ex ante} probability distributions of $m_t(n)$ and $G(x|P(K, n), n)$. To emphasize this point, note that real fuel prices and electricity demand are both likely to grow over the next ten years. This "positive correlation" has nothing to do with the correlations we are speaking of. Similarly, the cross-price elasticity of demand between electricity and oil is typically found to be positive. This "positive correlation" also has very little to do with the type of forecasting error correlations we are dealing with here. Both of these factors affect predictions about the \textit{levels} of demand and input prices, but they tell us nothing about the correlations between errors in our predictions of these future levels. The latter is what the covariance terms in equations (7-7) and (7-8) are based upon.

Here and throughout this paper so far we have been examining marginal cost pricing, which may result in highly variable profits. An alternative pricing rule such as pricing so as to recoup all fixed plus operating costs, i.e., average cost pricing, would result in very different pricing outcomes. With average cost pricing prices will respond to both fixed and operating cost uncertainty, and the above
derived formulas will be incorrect. We explore this issue in the next section.

These results have been derived under the assumption of risk neutrality. The introduction of risk aversion would presumably have the usual effect of motivating capacity choices towards decisions involving less risk. We defer a consideration of this issue until Section VII.D below.

Our assumption that each technology is a fixed proportion technology yields stronger results than would be derived with variable proportions technologies. When input costs have a zero correlation with $G(x; P(K, n), n)$ then uncertainty has no effect on capacity choices with fixed proportions technologies. Signing the correlations we speak of is therefore sufficient for showing the direction in which uncertainty biases capacity choices. With variable proportions technologies, one does not get a cost function which is linear in capital and operating costs, but rather one gets a cost function concave in these input prices. Even with zero correlations between demand and input prices, uncertainty has an effect on capacity choices. As Perrakis has demonstrated, with variable proportions uncertainty biases one towards the less risky inputs. As a result, the signs of the input price-demand correlations are not sufficient to determine the direction of the bias in capacity choices. In the next section we examine the impact of interfuel substitutability on optimal capacity choices in a simple framework which builds upon our representation of uncertain costs.
VII.B. Interfuel Substitutability

It is readily observed that many electric utility generating units are designed so as to be capable of burning multiple fuels. This can be seen as a mechanism for reducing both the expected value and the variance of fuel cost. Such interfuel substitutability is generally only feasible with a sacrifice of heat efficiency, as well as additional capital expenditures. In this section we examine both the ex post decision of how to operate capacity so as to take advantage of such interfuel substitutability, and the ex ante decision about whether or not to invest in such flexible capacity rather than choose conventional single fuel capacity. Fuss and McFadden (1971) derived many of the results we find below. We expand upon their work here and demonstrate that interfuel substitutability is readily incorporated into our model through the introduction of new generating technologies.

Suppose that in addition to the second generating technology, which has fixed and operating costs $c_2$ and $m_2(n)$ respectively, there is an alternative technology which has a higher capital cost $c_2'$ but the capability of burning either the same fuel as the second technology or a different fuel, which has a price of $m_2'(n)$. Let $\text{eff}(\rho)$ be the heat efficiency of the alternative technology relative to the second technology when fraction $\rho$ of the heat is generated from the different fuel. The fuel cost of this alternative technology can then be written

$$m_2(n) = \frac{(1 - \rho)m_2(n) + \rho m_2'(n)}{\text{eff}(\rho)} \quad 0 \leq \rho \leq 1$$

(7-9)
\( \rho \) is a choice variable of the electric utility and should be chosen once \( \eta \) is known so as to minimize \( \tilde{m}_2(\eta) \) for all realizations of \( \eta \). If \( \text{eff}(\rho) \leq 1 \) for all \( \rho \) then a necessary condition for the alternative technology to be adopted is for

\[
\frac{m_2^*(\eta)}{\text{eff}(\rho)} < m_2(\eta) \text{ for some } \eta \text{ and } \rho \text{ optimal given } \eta.
\]

For those realizations of \( \eta \), \( \rho \) should be increased and a larger fraction of the alternative technology used. If the unit is designed for complete specialization in one fuel or another, then \( \rho \) can take on values only of 0 and 1. If inputs are continuously variable then \( \rho \) should be chosen between 0 and 1 so as to minimize (7-9).

Several results follow immediately from this characterization of interfuel substitutability. The first and perhaps most important is that this alternative technology with interfuel substitution possibilities can now be incorporated into our investment process just like any other technology, with capital and uncertain operating (fuel) costs \( c_2 \) and \( \tilde{m}_2(\eta) \) respectively.

Most investment modeling frameworks ignore fuel cost uncertainty, and hence have a great deal of difficulty incorporating interfuel substitution in the decision making process. As Fuss and McFadden (1971) demonstrate, uncertainty is the only reason why a firm would choose a technology with both higher capital costs and lower fuel efficiency than the cost minimizing capacity choices for any level of fuel costs. The flexibility to adjust fuel usage in response to fuel prices may be worth the efficiency loss and additional capital cost.
Interfuel substitutability will tend to lower both the mean and the variance of fuel price uncertainty. The ability to shift into an alternative fuel if it should become the lower-cost alternative effectively truncates the probability distribution of the primary fuel price. As a result both the mean and the variance of fuel prices are reduced. Unfortunately, the lower operating efficiency tends to offset this effect. As we show below, it is possible that even if the mean of \( \bar{m}_2(\eta) \) is above the mean of \( m_2(\eta) \) it may still be attractive under certain not implausible circumstances.

Interfuel substitutability focuses attention on the possibility of switches in the order in which different generating technologies are loaded, which are not only possible but likely with interfuel substitution. It should be remembered that when we originally arranged our generation technologies, they were placed in order of increasing operating cost and decreasing capital cost. This hierarchy then determined the order in which generating capacity should be loaded. With fuel cost uncertainty the fuel loading order may be dependent on the realization of \( \eta \).

Consider \( m_2 \) and \( \bar{m}_2 \) and suppose an electric utility possesses some of each type of capacity. By design \( m_2(\eta) \) is below \( \bar{m}_2(\eta) \) for certain realizations of \( \eta \), since the alternative technology has a lower operating efficiency. For the rest of the distribution of \( \eta \), \( m_2(\eta) \) is above \( \bar{m}_2(\eta) \). In the former instance the second technology should be loaded before the alternative technology which allows interfuel substitutability. In
the latter instance the reverse is true, and generation should be more intensive from the alternative generating technology. The ability to switch the loading order of different generating technologies increases the attractiveness of interfuel substitutability; one can rely more heavily on alternative, less expensive fuels when conditions are favorable, yet rely less on that capacity when its lower operating efficiency makes it less attractive. This is a result which can only be seen when uncertainties are considered explicitly. Different behavior for different realizations of \( n \) make the consideration of expected values alone misleading.

So far we have shown that a technology with interfuel substitution possibilities can be treated like essentially any other technology with uncertain fuel costs. When will it be desirable to invest in such a technology? Suppose that for a single generating unit with capacity \( I \) kw, a utility is choosing between the second, single fuel generating technology and the alternative technology we have developed here. Let \( K \) and \( \tilde{K} \) be the vectors of capacity in the two respective cases. Recognizing the switching in the loading order, the operating cost savings that can be realized by choosing the alternative generating capacity can be derived from equation (7-4) to be

Operating Cost Savings =

\[
\int_{\tilde{m}_2(n) < m_2(n)} \left\{ \int_{K_1}^{K_1 + I} \left[ m_2(n)(x - K_1)g(x|P(K, n), n) - \tilde{m}_2(n)(x - K_1)g(x|P(\tilde{K}, n), n) \right] dx \right\} f(n)dn
\]
\[ + \frac{-\tilde{m}_2(n) - \tilde{m}_2(n)}{m_2(n)} \begin{cases} \int_{K_2-I}^{K_2} (m_2(n)(x - K_2)g(x|\mathcal{P}(K, n), n) \\ - \tilde{m}_2(n)(x - K_2 - I)g(x|\mathcal{P}(\tilde{K}, n), n))dx \end{cases} \ f(n)dn \]

\[ (7-10) \]

The above expression can be simplified appreciably if we ignore the fact that the two different investments have different pricing implications \([\mathcal{P}(K, n) \text{ versus } \mathcal{P}(\tilde{K}, n)]\). Suppose we also simplify to the case where complete fuel specialization for the alternative technology is always optimal and fuel efficiency is constant for different choices of \(\rho\). With these simplifications we can write

**Operating Cost Savings =**

\[ \frac{-\tilde{m}_2(n) - \tilde{m}_2(n)}{m_2(n)} \left[ \int_{K_1}^{K_1+I} (x - K)g(x|\mathcal{P}(K, n), n)dx \right] f(n)dn \]

\[ + \frac{-\tilde{m}_2(n) - \tilde{m}_2(n)}{m_2(n)} \left[ \int_{K_2-I}^{K_2} (x - K_2 - I)g(x|\mathcal{P}(K, n), n)dx \right] f(n)dn \]

\[ (7-11) \]

\[ = \frac{-\tilde{m}_2(n) - \frac{\tilde{m}_2(n)}{\text{eff}}}{m_2(n)} [X(n)] \int_{K_1}^{K_1+I} f(n)dn \]

\[ + \frac{-\tilde{m}_2(n)(1 - \frac{1}{\text{eff}})}{m_2(n)[K_2-I]} [X(n)] \int_{K_2-I}^{K_2} f(n)dn \]

\[ (7-12) \]
where the $[X(n)]_{a}^{b}$ terms are as implicitly defined in the equation, i.e.,
expected generation from units operating between capacity levels $a$ and $b$.

Equation (7-12) clearly differentiates the two possible uses of the
alternative generating technology. If $\bar{m}_2 < m_2$, so that the alternative
technology is attractive, then that technology is used relatively
intensively, $[X]_{K_1}^{K_1+I} > [X]_{K_2}^{K_2-I}$ and the fuel cost savings per KW-year are
the differences between $m_2$ and $\frac{m_2}{\text{eff}}$. When $\bar{m}_2 \geq m_2$ then the alternative
technology is used less intensively and an excess cost occurs because
the alternative technology uses the primary fuel less efficiently.

Offsetting these operating cost savings are the additional capital
costs incurred in constructing capacity with interfuel substitution
capability. These costs are simply $(c_2 - c_2)I$. These additional capital
costs can be contrasted with the operating costs as derived in equation
(7-10) or (7-11) to decide whether interfuel substitutability is
desirable.

It is interesting to note that there are two possible cases where
it can be desirable to use a generating technology which allows
interfuel substitution even if the expected value of its fuel price is
above that of a single fuel generating technology. The first case is
if the loading order can be reordered so as to use this capacity very
intensively when conditions for its use are most favorable, and less
intensively when conditions are less favorable. The ability to buy and
sell electricity to neighboring utilities may make wider factor
intensity switches possible than could be realized by any individual
utility. The second case follows from our results in Section VII.A
above. If the alternative technology's fuel cost has a smaller covariance with the level of demand, and hence costs less to operate when demand is highest, then it may be desirable even if its fuel price has a higher mean and even if no reordering of the loading order is possible (e.g., when specialization in one technology or the other is optimal).

In this section we have examined interfuel substitutability by treating capacity which allows such a tradeoff as simply one more type of generating capacity with uncertain fuel costs. Substitution possibilities tend to lower both the mean and variance of fuel costs, however these savings are offset by both lower operating efficiency and higher capital costs. The ability to switch around loading orders favors technologies with interfuel substitution possibilities, and may make them attractive even if they appear to have a higher mean fuel cost than single fuel generation technologies. Negative correlations between secondary fuel prices and demand may also favor such flexible capacity.

VII.C Average Cost Pricing

Up until this point we have considered only two different pricing rules: short-run marginal cost pricing and long-run marginal cost pricing. As shown in Chapter IV, Section C, both marginal cost pricing rules may result in either positive or negative profits, depending on demand elasticities, capital and operating costs and actual capacity choices. Although marginal cost pricing principles are given some
attention when regulatory agencies choose the structure of rates to be set, present regulatory practices do not permit the substantial fluctuations in profits that might result under strict marginal cost pricing. Instead prices are based on somewhat vague minimum and maximum allowable rates of return.\(^4\) (See Joskow, 1973.) As an approximation of this regulatory practice we examine here the case where a single price is set in the last period, after any uncertain random variables have been realized, such that total revenues equal total capital plus operating costs. This is a bit of a polar case, but an interesting one nonetheless.

We are not going to examine the Aversch-Johnson (1962) effect, in which capacity choices are distorted by electric utilities in response to too high a rate of return on capital.\(^5\) Prices and investments are chosen by an omniscient utility regulator constrained to price electricity at its average cost. These choices are contrasted with socially optimal (marginal cost based) choices and the effects of three different kinds of uncertainties are considered: demand uncertainty, operating cost uncertainty, and capital cost uncertainty. We start with simple model without any uncertainty which could be applied to any regulated monopolist. We then add uncertainty to this simple model, and consider the three types of uncertainties mentioned above. Finally, we consider the applicability of these average cost pricing results to electric utilities, for which the presence of time varying demand, multiple generating technologies, and shortage costs (rationing) modify our results.
Let \( D(P) \) be total demand sold at a single price, \( P \), and \( C(K, D(P)) \) be the total cost of serving that demand, given (scalar) capacity \( K \). Using consumers' plus producers' surplus as a welfare measure, social welfare, \( W \), can be represented as:

\[
W = \int_{0}^{\infty} D(p)dp + PD(P) - C(K, D(P))
\]  
(7-13)

Social welfare when constrained to achieve a zero profit can be maximized by constructing the following Lagrangian:

\[
\tilde{W} = \int_{0}^{\infty} D(p)dp + PD(P) - C(K, D(P)) + \lambda[PD(P) - C(K, D(P))]
\]  
(7-14)

The first order conditions for a welfare maximum are

\[
\frac{\partial \tilde{W}}{\partial K} = -\frac{aC}{aK} - \lambda \frac{aC}{aK} = 0
\]  
(7-15)

\[
\frac{\partial \tilde{W}}{\partial P} = P \frac{aD}{aP} - \frac{aC}{aD} \frac{aD}{aP} + \lambda[P \frac{aD}{aP} + D(P) - \frac{aC}{aD} \frac{aD}{aP}] = 0
\]  
(7-16)

\[
\frac{\partial \tilde{W}}{\partial \lambda} = PD(P) - C(K, D(P)) = 0
\]  
(7-17)

Equation (7-15) can be simplified to

\[
(1 + \lambda) \frac{aC(K, D(P))}{aK} = 0
\]  
(7-18)

Whether or not the profitability constraint is binding, equation (7-18) gives the result that without any uncertainty or rationing, it is optimal to choose capacity so as to minimize the expected capital plus operating costs of serving the demand resulting from average cost pricing. Lest this seem total trivial, the reader is forewarned that
this result does not hold when there is either uncertainty or rationing (shortages), as we demonstrate below.

Equations (7-16) and (7-17) can be used to find a useful expression for $\lambda$, the marginal welfare change from relaxing the revenue requirement by one dollar:

$$\lambda = \frac{-[P - \frac{aC}{aD}] \frac{aD}{aP}}{[P + D \frac{aP}{aD} - \frac{aC}{aD}] \frac{aD}{aP}} = \frac{AC(K, D) - MC(K, D)}{MC(K, D) - MR(D)} \quad (7-19)$$

where

$AC(.) = \text{average cost}$

$MC(.) = \text{marginal cost}$

$MR(.) = \text{marginal revenue}.$

The second order conditions for a welfare maximum require that $MC > MR$. Beyond using this fact we will ignore the second order conditions and assume that they are satisfied. Equation (7-19) implies that $\lambda$ is positive whenever $AC > MC$, negative if $AC < MC$ and zero when the two are equal. The three different cases are depicted in Figures 7-2a through 7-2c. In the first case price is above marginal costs and welfare can be improved by relaxing the revenue requirement. It is often argued that economies of scale make this case the correct one for electric utilities. In the second case price is below marginal cost, and welfare would be improved by increasing the revenue requirement above total generating plus operating costs. In the third case, which will be true if there are long run constant returns to scale, prices, $AC$ and $MC$ are all equivalent, and $\lambda$ will be zero. Note
Figure 7-2a
Average Cost Pricing with Locally Increasing Returns

\[ P = AC > MC > MR \]
\[ \lambda > 0 \]

Figure 7-2b
Average Cost Pricing with Locally Decreasing Returns

\[ MC > P = AC > MR \]
\[ \lambda < 0 \]

Figure 7-2c
Average Cost Pricing with Locally Constant Returns

\[ P = AC = MC > MR \]
\[ \lambda = 0 \]
that in deriving the expression for \( \lambda \) we did not use equation (7-15), and hence the expression holds regardless of whether capacity is chosen optimally or not.\(^9\)

Suppose now that capacity must be chosen in the presence of uncertainty about what the future may be. At this point we need not specify whether it is demand uncertainty, operating cost uncertainty or capital cost uncertainty. Let all agents be risk neutral, so that expected consumer plus producer surplus is a valid welfare measure. Prices are assumed to be chosen after the uncertain vector, \( \eta \), is realized, however \textit{ex ante} \( \eta \) has a probability distribution \( f(\eta) \) defined over the set \( S \). Expected welfare can be written as

\[
\tilde{E}_W = \int_S \int_{\mathcal{P}} D(p, \eta)dp + PD(P, \eta) - C(K, D(P, \eta), \eta) \\
+ \lambda [PD(P, \eta) - C(K, D(P, \eta), \eta)] f(\eta)d\eta
\]  

(7-20)

If prices are selected \textit{after} \( \eta \) is known, then the pricing problem is precisely the same as that derived previously, and the first order conditions (7-16) and (7-17) continue to be applicable. Moreover, our expression for \( \lambda \), which is now conditional on \( \eta \), remains correct.

The manner in which prices will subsequently be chosen affects desired investments. Taking the partial derivative of \( \tilde{E}_W \) with respect to \( K \) and setting equal to zero equals

\[
\frac{\partial \tilde{E}_W}{\partial K} = \int_S - \frac{\partial C}{\partial K} - \lambda \frac{\partial C}{\partial K} f(\eta)d\eta = 0
\]  

(7-21)

Substituting in our expression for \( \lambda \) and simplifying
\[
\frac{\Delta E}{\Delta K} = - \int_S (1 + \lambda(n)) \frac{\Delta C}{\Delta K} f(n)dn
\]

\[
= - \int_S \left[ 1 + \left( \frac{AC - MC}{MC - MR} \right) \right] \frac{\Delta C}{\Delta K} f(n)dn
\]

\[
= - \int_S \left( \frac{AC - MR}{MC - MR} \right) \frac{\Delta C}{\Delta K} f(n)dn
\]

\[
= - \int_S \left( \omega(n) \frac{AC(K, D(P, n), n)}{\Delta K} \right) f(n)dn = 0 \tag{7-22}
\]

where \( \omega(n) = \frac{AC - MR}{MC - MR} \)

The function \( \omega(n) \) is a weighting function; since \( AC(=P) > MR \) and as previously noted \( MC > MR \), it is strictly positive. Equation (7-22) indicates capacity should not be chosen to minimize expected costs, but rather to minimize a weighted average of expected costs. The nature of those weights depends on the type of uncertainty affecting the problem, which we explore below.

Before continuing it is worth contrasting this result with the marginal cost pricing result. Under marginal cost pricing, the first order conditions for optimal capacity and pricing choices under uncertainty are

\[
\int_S \frac{AC(K, D(P^*(n), n), n)}{\Delta K} f(n)dn = 0 \tag{7-23}
\]

\[
P^*(n) = \frac{AC}{\Delta D} \text{ for all } n \tag{7-24}
\]

The first of these equations implies that one should minimize the
expected cost of meeting the demand resulting from optimal marginal
cost prices. At first glance, minimizing expected total costs might
appear to be a desirable objective under average cost pricing as well.
As we have shown above, this is incorrect. Capacity choices should
take into account not only their effect on total costs, but also their
effect on prices through the profitability constraint. This can be
achieved by weighting total costs for each realization of \( n \), with the
weights as indicated by equation (7-22).

At this point we distinguish between three different types of
uncertainty which affect capacity choices under average cost pricing:
demand uncertainty, operating cost uncertainty, and capital cost
uncertainty. Consider first demand uncertainty, by which we mean
uncertainty which affects the position of the demand curve while
leaving the AC and MC curves unchanged. The case where there are only
three different realizations of \( n \) is depicted in Figure 7-3. The
average and marginal cost curves are based on the level of capacity
which would be chosen if expected costs for the middle scenario (\( n_2 \))
are minimized. The cost curves are thus not the ones which should be
chosen under either MC or AC pricing.

Using marginal cost pricing principles the socially optimal prices
would be \( P_1^*, P_2^*, P_3^* \). Under average cost pricing principles, prices
will instead be \( P_1^{AC} (> P^*), P_2^{AC} (=P^*) \) and \( P_3^{AC} (<P^*) \). Hence we have the
not surprising result that average cost pricing in the presence of
demand uncertainty results in prices which are too high for low demand
scenarios and too low for high demand scenarios. Of somewhat greater
Figure 7-3

Average Cost Pricing with Demand Uncertainty

\[ P^*_1 > P^*_3, \quad \lambda > 0, \quad \omega > 1 \]
\[ P^*_2 = P^*_2, \quad \lambda = 0, \quad \omega = 1 \]
\[ P^*_3 < P^*_3, \quad \lambda < 0, \quad \omega < 1 \]
interest, although still not too surprising, is the result that relative to short run marginal cost pricing, average cost pricing exacerbates fluctuations in consumption due to demand uncertainty. Consumption is higher in high demand scenarios and lower in low demand scenarios under average cost pricing than under short run marginal cost pricing.

What implications do these pricing and consumption results have for the ex ante choice of capacity. One of the major findings in Chapter IV was that under long run marginal cost pricing, it is usually desirable to reduce capital intensity as demand uncertainty increases. Under short run marginal cost pricing this effect is mitigated, and it was shown that desired capital intensity may either increase or decrease, although unless pricing effects are large, there remains a strong tendency for desired capital intensity to decline as uncertainty increases. Under average cost pricing similar statements apply, and it is possible for either more or less capacity to be desired with uncertainty than without, depending on the cost and demand structure. We have not been able to derive any general statements about the relationship between optimal capacity choices under average cost pricing and marginal cost pricing.

We are able to show however, that with constant returns to scale, demand uncertainty and average cost pricing, less capacity should be constructed than the level which minimizes expected costs. Note from the definition of \( \omega(\eta) \) that \( \omega(\eta) \) is greater than one when \( AC > MC \) and less than one when the reverse is true. Also note that
\( \frac{aC}{aK} \) is positive for low realizations of \( n \), and negative for high realizations of \( n \), with \( \frac{a^2C}{aK_d} \leq 0 \) (marginal costs decline with additional capacity). With constant returns to scale \( \frac{aC}{aK} = 0 \) at \( AC = MC \), where \( \lambda = 0 \) and \( \omega(n) = 1 \). From this it follows that

\[
\int_S \omega(n) \frac{aC(n)}{aK} f(n) \, dn > \int_S \frac{aC}{aK} f(n) \, dn
\]  

(7-25)

This inequality implies that \( \frac{aW}{aK} < 0 \) if \( \int_S \frac{aC}{aK} f(n) \, dn = 0 \) (see equation 7-21). Welfare can be improved by reducing total capacity, as noted above. Although we do not pursue the argument here, the result is not necessarily true under decreasing or increasing returns. 12

The intuition behind this result is straightforward. Suppose capacity is simply chosen so as to minimize expected costs. We have argued above that if there were no profitability constraint, one would like to raise prices for high levels of demand and lower it for low levels. One mechanism for partially achieving this result is to reduce the holding of capacity. This raises prices (and average costs) for high realizations of \( n \), while lowering them for low realizations. The burden of increased costs is an unfortunate side effect. 13

With uncertain operating costs the results are somewhat different. The demand curve is now known with certainty, however both the marginal and average cost curves are shifting up and down with different realizations of \( n \). Figures 7-4a through 7-4c depict three possible simple cases. (Composites of the three are possible.) To simplify the diagrams only two different scenarios are shown on each diagram. In
Figure 7-4a
Case Where AC Pricing Exacerbates Operating Cost Uncertainty

Figure 7-4b
Case Where AC Pricing is Equivalent to MC Pricing

Figure 7-4c
Case Where AC Pricing Mitigates Operating Cost Uncertainty
the first case AC pricing exacerbates operating cost uncertainty, setting prices too high when operating costs are high, and setting them too low when operating costs are low. The weights, \( \omega \), attached to \( \frac{\partial C}{\partial K} \) when choosing \( K \) are higher for high realizations of operating costs, and by an argument similar to that above, it can be shown that in this first case one should purchase less capacity than the level which minimizes expected costs.

In the second case MC and AC both increase such that \( MC = AC = P \) at the optimum. In this case \( \omega = 1 \) and minimizing expected costs while pricing at average costs achieves the unconstrained social optimum.

In the third case prices are less variable under average cost pricing than under marginal cost pricing. Of some interest is the fact that \( \omega(\eta) \) now decreases as operating costs increase, so that it is desirable to invest in more capacity than would result if expected costs are minimized.

A third type of uncertainty which we can consider is capital cost uncertainty. Since capital costs do not affect short-run marginal costs, under marginal cost pricing such uncertainty can be ignored, and expected capital costs used in selecting the level of capacity. With average cost pricing this is no longer the case. Figure 7-5 depicts how \( AC \) may vary with the realization of \( \eta \). Once again AC pricing lowers prices too far in favorable (low capital cost) outcomes and raises them too high when capital costs are high, and the optimal response is to reduce capacity below the level which minimizes expected costs.
Figure 7-5

Average Cost Pricing with Capital Cost Uncertainty

\[ \omega(\eta^2) > 1 > \omega(\eta^1) \]
We have now examined in a very simple model three different types of uncertainty and their effect on desirable price and investment choices by a regulated monopolist forced to adopt average cost pricing. The analysis has been an elaboration upon second best principles in the presence of uncertainty. Given that prices are not set optimally, capacity choices should not be chosen according to first-best cost minimizing rules, i.e., expected costs should not be minimized. Instead costs should be weighted by a function which takes into account the discrepancy between price and marginal cost. Except for two cases of operating cost uncertainty, the basic finding has been that (the same or) less capacity should be constructed under average cost pricing than the level which would be chosen to minimize expected costs.

The results that we have shown in this simple model carry over to the more complex case of electric utilities, however three new factors must be introduced: the rate of consumption varies over time, multiple generation technologies are available, and rationing (shortages) may be necessary (or in fact desirable). As we show below the fact that consumption is time varying gives our weighting function greater importance, while the possibility of rationing demand means that minimizing expected costs is no longer desirable even without any uncertainty, although the distortion affects only the least capital intensive generating technology.

Using the notation developed in Chapter III we can represent total demand, $D(P, \eta)$ as
(7-26)  
\[ D(P, \eta) = \int_0^\infty xg(x|P, \eta)dx \]

while total served demand, \( D(P, \eta) \) is

(7-27)  
\[ \tilde{D}(P, \eta) = \int_0^{K_n} xg(x|P, \eta)dx \geq D(P, \eta) \]

where \( K_n \) is total generating capacity of all types. As in previous chapters we will find it more convenient to work with cumulative capacities of each generating type, \( k_i \). Continuing to let \( C(K, P) \) be total fixed, operating and shortage costs, we represent total fixed plus operating costs, excluding shortage costs, by \( \tilde{C}(K, P) \). Ignoring uncertainty about \( \eta \) for the moment, we can write the welfare maximization problem in the case where all electricity is sold at a single price in the form of the Lagrangian

(7-28)  
\[ \tilde{W} = \int_0^\infty D(p)dp + PD(P) - C(K, P) + \lambda[PD(P) - C(K, P)] \]

Maximizing over \( P \) and \( k_i \), \( i = 1, ..., n \) yields first order conditions

(7-29)  
\[ \frac{\partial \tilde{W}}{\partial k_i} = \frac{\partial C}{\partial k_i} - \lambda \frac{\partial \tilde{C}}{\partial k_i} = 0 \quad i = 1, ..., n \]

(7-30)  
\[ \frac{\partial \tilde{W}}{\partial P} = P \frac{\partial D}{\partial P} - \frac{\partial C}{\partial P} + \lambda[\tilde{D} + P \frac{\partial \tilde{D}}{\partial P} - \frac{\partial \tilde{C}}{\partial P}] = 0 \]

(7-31)  
\[ \frac{\partial \tilde{W}}{\partial \lambda} = PD(P) - \tilde{C}(K, P) = 0 \]

It is readily seen that \( \frac{\partial C}{\partial k_i} = \frac{\partial \tilde{C}}{\partial k_i} \), \( i \neq n \) (only changes in total capacity
affect shortage costs) hence in the absence of any uncertainty, the generating capacity of all but the least capital intensive technology should be chosen so as to minimize generating costs. The choice of the nth technology is affected by shortage costs. Writing out \( \frac{\partial C}{\partial K_n} \) and \( \frac{\partial \tilde{C}}{\partial K_n} \) in our usual notation:

\[
c_n + (m_n - \alpha)[1 - H(K_n)] + \lambda[c_n + m_n[1 - H(K_n)]] = 0 \tag{7-32}
\]

which can be rewritten as

\[
(1 + \lambda)[c_n + m_n[1 - H(K_n)]] = \alpha[1 - H(K_n)] \tag{7-33}
\]

The desired choice of \( K_n \) will be less (greater) than the cost minimizing level whenever \( \lambda > 0 \) (\( \lambda < 0 \)).

The intuition behind this result is that shortage costs are an externality to the firm, not reflected in \( \tilde{C} \). Pricing at average cost distorts capacity choices not only because price does not equal to marginal cost (in the sense we have used previously), but also because the costs used in setting prices exclude this externality. The second best solution is to alter the choice of capacity to reduce this externality. Peaking capacity is the only capacity altered, since it is the least expensive means of achieving this result.

Using equation (7-30) we can again derive a useful expression for \( \lambda \).

\[
\lambda = \frac{(p \frac{\partial \tilde{D}}{\partial P} - \frac{\partial C}{\partial P})}{(\frac{\partial \tilde{C}}{\partial P} - \tilde{D} - p \frac{\partial \tilde{D}}{\partial P})} \tag{7-34}
\]

Once again, the second order conditions for an average cost pricing
optimum require that the denominator be negative, hence the sign of \( \lambda \) is the opposite of the sign of the numerator. The numerator has already been encountered in Chapter IV in our derivations of optimal prices.

The profitability results derived in Chapter IV, Section C, in terms of price elasticities of demand have a direct bearing on the sign of \( \lambda \). In that section (and in Appendix C) we argued that the optimal marginal cost price is above average costs if demand is more elastic during high demand periods than during low demand periods. The price should be below average cost if the reverse elasticity pattern is true. From this it follows that in the first case AC pricing will result in prices which are too low, the numerator in equation (7-34) will be negative and \( \lambda \) will be positive. In this case less peaking capacity should be build under AC pricing than the level which minimizes total costs including shortages.

Uncertainty can be introduced into the problem in a manner analogous to the previous simple model. Taking the expectation of equation (7-28); taking derivatives with respect to \( K_i, i = 1, \ldots, n \); substituting in our expression for \( \lambda \), and simplifying yields

\[
\int_S \omega(n) \frac{aC(n)}{aK_i} f(n) dn = 0 \quad i \neq n \quad (7-35)
\]

\[
\int_S \left[ \frac{aC(n)}{aK_i} + \lambda(n) \frac{aC}{aK_n} \right] f(n) dn = 0 \quad (7-36)
\]

where \( \omega(n) = 1 + \lambda(n) \)
\[
\frac{\Delta SC}{\Delta K_n} = - \alpha [1 - H(K_n, P, n)]
\]

The same results shown in the simple case for uncertain demand, uncertain operating costs and uncertain capital costs continue to hold in the case of electricity. In most cases it will be desirable to build less capacity than the level which minimizes expected costs. The presence of multiple generating technologies and time varying demand amplifies the significance of the function \( \omega(n) \). The function \( \omega(n) \) now reflects not only how much total demand changes but also how those changes are distributed throughout the year; not only whether operating or fixed cost are uncertain, but which technologies display the greatest uncertainty. More specific statements cannot be made without assuming more specific demand and cost specifications.

In this section we have extended our model to the case where prices are not set equal to marginal costs but rather average costs. All of the results have been "second best" results given this constraint. In a simple model without time varying demand or any shortages it was demonstrated that less capacity should be chosen than the level which minimizes expected costs under either demand uncertainty, capital cost uncertainty, or certain types of operating cost uncertainty. Only when marginal costs increase less rapidly than average costs along a given demand curve should more capacity be built than the expected-cost-minimizing level of capacity. These results carry over to the case of multiple generating technologies and time varying demand in the case of electric utilities, although the first order conditions for desired
capacity choices now reflect a more complex pattern of cost and demand covariation. The externality caused by shortage costs was shown to cause a deviation in the choice of peaking capacity from cost minimizing principles, but not in the choice of other technologies.

VII.D Risk Aversion

All of the preceding analysis has been based on the assumption of risk neutrality. We relax that assumption here and examine cases involving risk aversion by both producers and consumers. Attention is focused on two cases—the case where consumers and producers are equally risk averse, and the case where producers are less risk averse than consumers. External insurance is assumed not to be available, however as we shall see when consumers are more risk averse than producers, a sort of "internal insurance" is optimal.

The effects of risk aversion on decision making are fairly predictable in most instances: risk aversion biases decisions toward choices which reduce the variance of the objective function, even at a loss in its expected value. This reduction in the expected value is the risk premium agents are willing to pay to reduce the variance of the objective function. The insights gained here arise principally from the fact that investment and pricing decisions may not be made simultaneously.

Previously, we have assumed that the marginal utility of income was both constant over different realizations of the random vector \( \eta \) and equal for all consumers and producers. We assume here that the
marginal utility of income is a function of $n$, $\omega^C(n)$, for the aggregate of all Consumers and $\omega^P(n)$ for the one producer, the electric utility. If we weight consumer surplus minus shortage costs by $\omega^C(n)$ and profits by $\omega^P(n)$ then our expected welfare measure becomes

$$EW = \int_S \omega^C(n) \left[ \int_p^\infty D(p, n) dp - SC(P, K, n) \right]$$

$$+ \omega^P(n)[PD(p, n) - \tilde{C}(K, P, n)] f(n) dn$$

(7-37)

where $SC(P, K, n)$ are shortage costs, $\tilde{D}$ is served demand, and $\tilde{C}$ is fixed plus operating costs. The fact that shortage costs are an externality to the firm and are borne by consumers, is incorporated in this welfare measure. To simplify the analysis rewrite (7-37) in the form

$$EW = \int_S \omega^C(n) \left[ \int_p^\infty D(p, n) dp \right.$$

$$+ \omega^P(n)[PD(p, n) - \tilde{C}(K, P, n)]$$

$$- \alpha(n) \int_{K_n}^\infty (X - K_n) g(X, P, n)] f(n) dn$$

(7-38)

where

$$\alpha(n) = \frac{\omega^C(n)}{\omega^P(n)}(\alpha - P(n)) + P(n)$$

By treating $\alpha(n)$ as simply a random shortage cost we can pretend that
it has in fact been internalized by the firm, while weighting it at its theoretically correct level. The results in Chapter VII, Section A will help us with the interpretation of this random cost. Bearing in mind the fact that \( C(K, P, \eta) \) now incorporates random shortage costs, we can rewrite equation (7-38) as

\[
EW = \int_S \omega^C(\eta)[ \int_P^\infty D(p, \eta)dp

+ \omega^P(\eta)[PD(p, \eta) - C(K, P, \eta)]f(\eta)d\eta
\]  

(7-39)

Suppose that prices are set after \( \eta \) is realized. Then the first order condition for choosing the optimal price becomes

\[
- \omega^C(\eta) D(P) + \omega^P(\eta)[D(P) + P \frac{\partial D}{\partial P} - \frac{\partial C}{\partial P}] = 0
\]  

(7-40)

which we can rewrite as

\[
P = \frac{\frac{\partial C}{\partial P}}{\frac{\partial D}{\partial P} + \frac{\omega^C(\eta)}{\omega^P(\eta)}} D(P)
\]

(7-41)

From equation (7-41) we can see how risk aversion affects prices conditional on any set of investments. In the case where consumers and producers are equally risk averse it is desirable to price at marginal cost, as shown in Chapter IV. (Note also that \( \alpha(\eta) = \alpha \) if \( \omega^C(\eta) = \omega^P(\eta) \).) In the case where consumers are more risk averse than producers, then conditional on a given set of investments it is optimal to price below marginal cost in scenarios where the marginal utility of income for consumers is above that of producers (\( \omega^C(\eta) > \omega^P(\eta) \))
and above marginal cost where the converse is true. As with average cost pricing, this is a second best result, arising from the assumed unavailability of insurance. When consumers are more risk averse than producers then it is desirable for producers to partially insure consumers by systematically deviating from marginal cost pricing principles.

If (as appears most likely) \( \omega^c(\eta) > \omega^p(\eta) \) for realizations of \( \eta \) corresponding to low demand, or high capital costs, then this pricing result is the opposite of the one resulting from average cost pricing examined in the previous section. This would further accentuate the undesirability of average cost pricing. Remembering that operating cost uncertainty may result in prices which vary either too much or too little, risk aversion may move prices in either the same or the opposite direction as average cost pricing in this case.

If \( \omega^c(\eta) > \omega^p(\eta) \) for unfavorable realizations of \( \eta \), then these pricing results imply that consumption should be allowed to vary less with risk aversion than with risk neutrality for any given level of capacity. In sum it appears likely that risk aversion mitigates consumption uncertainty at the cost of increasing price uncertainty in the case where consumers are more risk averse than producers.

What do these price and consumption results imply about desired investments. Differentiating (7-34) with respect to \( K_i \) \( i = 1, \ldots, n \):

\[
\frac{\partial \text{EW}}{\partial K_i} = \int_S \omega^p(\eta) \frac{\partial C(K, P, \eta)}{\partial K_i} f(\eta) d\eta
\]  
(7-42)

Using our specification of \( C(K, P, \eta) \) we can write
\[ \frac{\partial E W}{\partial K_i} = \int_S \omega^P(n) \left( [c_i - c_{i+1}] + (m_i - m_{i+1})[1 - H(K_i | P, n)] \right) f(n) \, dn \quad (7-43) \]

for \( i = 1, \ldots, n-1 \)

\[ \frac{\partial E W}{\partial K_i} = \int_S \omega^P(n) \left( c_n + (m_n - \alpha(n))[1 - H(K_n | P, n)] \right) f(n) \, dn \quad (7-44) \]

From these two equations it can be seen that the only effect which consumer risk aversion has on desired investments is through its effect on prices and \( \tilde{\alpha}(n) \). Producer risk aversion affects investments directly. Consider first the case where consumers and producers are equally risk averse. Then as argued above price should equal marginal costs and \( \tilde{\alpha}(n) = \alpha \). Equations (7-43) and (7-44) imply simply that greater weight should be attached to scenarios where the producer has a higher marginal utility of income. If these correspond to low demand scenarios and costs are not random, then less of all types of capacity should be constructed, since

\[ \int_S \frac{\omega^P}{\bar{w}} [1 - H(K_i | P, n)] f(n) \, dn < \int_S [1 - H(K_i | P, n)] f(n) \, dn \quad (7-45) \]

where \( \bar{w} = \int_S \omega^P f(n) \, dn \)

What this means is that the firm should sacrifice some profits (raise costs) during high demand, high profit periods in order to raise profits during low demand, low profit periods when they are most needed.

Consider next the case of uncertain fuel costs. To keep the analysis simple, suppose that only \( m_1(n) \) is uncertain and that high values for \( \omega^P(n) \) correspond to high values of \( m_1(n) \). Using the results of Chapter VII, Section A, we can write equation (7-43) in the
\[ c_i + \bar{m}_1 \int_S \left[ (1 - H(K_1 \mid P, n)) \frac{\omega^p(n)}{\omega^p} \right] f(n) d\eta \]

\[ + \int_S (m_1(n) - \bar{m})(1 - H(K_1 \mid P, n)) \frac{\omega^p(n)}{\omega^p} f(n) d\eta \]  

(7-46)

If producers were risk neutral, then we know from our results in Section A of this chapter above that with fuel cost uncertainty more of this first technology should be built than is suggested by using expected fuel costs alone since the covariance term (on the middle line) will be negative. With risk aversion, greater weight should be attached to high fuel cost scenarios, reducing and eventually reversing this effect. The overall conclusion that can be reached is that risk aversion by producers reduces desired investment in technologies subject to fuel cost uncertainty.

The case of capital cost uncertainty when consumers and producers are both equally risk averse is straightforward. Since prices and demand are unaffected by this type of uncertainty under marginal cost pricing we need only note that for any uncertain capital cost, \( c_i(n) \),

\[ \int_S \frac{\omega^p(n)}{\omega^p} c_i(n) f(n) d\eta > \int_S c_i(n) f(n) d\eta \]  

(7-47)

Risk aversion reduces desired investment in technologies subject to capital cost uncertainty.

The implications of consumers being more risk averse than producers are difficult to demonstrate analytically but fairly simple to see heuristically. We noted above that in this case there is likely to be
greater price variation but less variation in consumption for any given set of capacity choices, relative to pure marginal cost pricing. Consider the set of capacity choices corresponding to minimizing expected costs with marginal cost pricing. A simple application of the results in Chapter IV, Section E suggests that by reducing the variance of demand, increased use of base load capacity and decreased use of less capital intensive technologies will be possible. Changes in the mean level of demand may modify this result somewhat, but the change in relative usages should be as stated. This may be summarized in the statement: holding producer risk aversion constant while increasing consumer risk aversion should increase desired capital intensity of generating capacity. 18

The results that we have demonstrated in this section on risk aversion are somewhat more tentative than those of previous sections, reflecting the somewhat ethereal nature of the marginal utility of income and risk aversion. There have been few real surprises: risk aversion in each case reduces dependence on those inputs most responsible for the uncertainty in consumer or producer surplus. The most interesting results are those arising when consumers are more highly risk averse than producers. In this case risk aversion reduces consumption uncertainty at the cost of greater price uncertainty. As a result, desired capital intensity increases as consumer risk aversion increases.

This chapter has examined four interesting extensions of the model which we have developed in previous chapters: uncertain fuel and
capital costs, interfuel substitutability, average cost pricing and risk aversion. The results that we have shown are too varied to summarize here. One similarity between all four sections, however, is that desired investments can be found by applying weights to all or a few of the components of the derivatives of our cost function. In most cases such weights should not be difficult to implement on a practical basis in applied investment planning and forecasting models.
Footnotes to Chapter VII

1. Note that this result is separate from the effect of risk aversion, which also increases the relative use of the less risky inputs. Risk aversion is examined in Section VII.D.

2. Capital cost uncertainty is relevant under average cost pricing or with risk aversion. These cases are considered below.

3. Throughout this thesis we have ignored the fact that the heat efficiency of individual plants is affected by their rate of output (higher efficiencies are obtained with output rates: see EGEAS (1980)). For consistency we continue to ignore this issue here. The function of eff(ρ) can be increasing, decreasing, convex or concave depending on the specific fuel and generating technologies. We will not specify it any further here.

4. Lumps sum transfers, particularly through use of declining block rates, could and to some extent already are being used to achieve certain profitability objectives while pricing electricity efficiently at marginal costs, however, regulators have not thoroughly adopted this course of action.

5. See Baumol and Klevorick (1970) for a discussion of the developments in this area.

6. Much of this section could be worked out by using inverse demand functions, as is more conventional, however, we continue our somewhat less convenient formulation in order to maintain consistency with the rest of the thesis, and more importantly, to reintroduce shortages and time varying demand toward the end of the section.

7. This can also be seen directly by noting that MR > MC \[ \Rightarrow \frac{aP}{aD} \Rightarrow \frac{aC}{aD} > \frac{C}{D^2} \Rightarrow \frac{aP}{aD} > \frac{aAC}{aD}. \] This would imply that the demand curve P(D) is less steeply sloped than the average cost curve, a condition for a local minimum rather than a local maximum.

8. Stewart (1979) and Boyd and Thompson (1980) make the point that lower reliability and longer construction periods may wipe out much of the apparent economies of scale of the largest units.

9. Case 1 will hold for electric utilities whenever there is surplus generating capacity, which there is at present throughout much of the U.S. Hence in a static framework there are strong welfare
arguments in favor of pricing below average costs at this time. Dynamic considerations of the cost of capital and risk taking may mitigate this static result.

10. In Chapter IV multiple generating technologies were considered, so that capital intensity was reduced even though total generating capacity might actually increase under demand uncertainty. In the simplified, one fixed input model considered here capital intensity can only be reduced by reducing total capital.

11. The author has tried unsuccessfully to derive formal conditions under which the average cost pricing investments should be lower than marginal cost pricing investments, however, the result apparently depends on second derivatives with respect to η, which are very opaque.

12. We are unable to construct a simple example, however with strong increasing returns there will be a region with ω(η) < 1 and yet aC/αK > 0. If most of the probability mass of η should be concentrated in this region, then we speculate that inequality (7-25) will not hold.

13. Note that the converse result also is true: holding the schedule of prices resulting from optimal AC pricing capacity choices constant, expected welfare could be improved by increasing K and allowing positive expected profits. This follows from

\[ 0 = \int_S \omega(\eta) \frac{\partial C}{\partial K} f(\eta) d\eta > \int_S \frac{\partial C}{\partial K} f(\eta) d\eta \]

which implies \( aW/\alpha K > 0 \).

14. This third case is not necessarily unrealistic. Note that short run demand curves are highly inelastic \( (\mu = 2) \). Rewrite the condition that AC = MC in the form \( D = (FC + VC(D))/MC(D) \) where FC and VC(D) are fixed and variable costs respectively. If fixed costs are large relative to variable costs, then D may decrease very rapidly as VC and MC both increase.

15. If \( \omega^C(\eta) > \omega^P(\eta) \) then \( \alpha(\eta) > \alpha \) which would tend to offset this pricing effect. If these realizations correspond to unfavorable realizations of η, as we posit, then there may be few if any shortages and \( \alpha(\eta) \) will be irrelevant. In any event, we speculate that the direct effect on prices mentioned in the main text will dominate.

16. The case where \( \omega^C(\eta) > \omega^P(\eta) \) for all realizations of η is uninteresting since then unconditional lump sum transfers would be desirable. We ignore this possibility.
17. The opposite directions in which uncertainty and risk aversion move desired investments in the presence of input cost uncertainty is demonstrated by Perrakis (1980) in a more general framework.

18. We noted that the fact that shortage costs are an externality to the firm can be treated via an uncertain shortage cost function

\[ \alpha(n) = \frac{\omega^C(n)}{\omega^P(n)}(\alpha - P(n)) + P(n) \]

If \( \omega^C(n) \) and \( \omega^P(n) \) are strongly positively correlated, as we expect them to be, and consumers more risk averse than producers, then \( \alpha(n) \) and \( \omega^P(n) \) should be positively correlated. If in addition \( [1 - H(K|P, n)] \) is not negatively correlated with \( \omega^P(n) \), then there should be less reliance on the shortage technology as consumer risk aversion increases.
VIII. Simulation Results Incorporating Pricing Effects

How much generating capacity should be built for 1990? What will be the impact of building less capacity? In this chapter a simulation model which implements many of the theoretical results derived above is used to address these questions. It is shown that the social costs of building too little generating capacity can be substantially reduced by choosing the appropriate set of prices. The ability to wait until there is less uncertainty about the level of future demand is also shown to result in significant social benefits. Average cost pricing is shown to result in prices which differ markedly from marginal cost prices in many instances. The results suggest that it is important to incorporate pricing effects in investment planning models. Considering long term demand and input price uncertainty may or may not be critical, depending on the degree of uncertainty, the long term growth rate, fuel costs, and other cost parameters.

State of the art investment planning models used by electric utilities have become highly sophisticated in certain dimensions yet remain naive in others. There is a widespread tendency for supply phenomena to be modeled in great detail while ignoring or giving only superficial treatment to demand effects. Many investment models examine how capacity choices are influenced by capital costs, fuel costs, forced outage rates and reliability constraints, however these same models frequently ignore the impact of these same variables on prices and hence the levels of demand for which this capacity is being built.¹ In a similar manner, uncertainties about the availability of individual units or the flow of water through hydroelectric generating units are often modeled explicitly, while
uncertainty about whether demand will grow 10 percent or 50 percent over
the next ten years is treated only superficially in the choice of
generating capacity.

In this chapter we ignore a great deal of supply complexity in order
to focus on prices and uncertainty and their effect on optimal investments.
Our simulation model differs from previous efforts in that: 1) we use a
closed loop model which explicitly takes into account the effect of
capacity additions on prices and hence the levels of demand for which the
capacity is being chosen; 2) we incorporate uncertainty about both the
level of future electricity demand and fuel prices by choosing capacity
which is optimal for a probability weighted average of many possible
"scenarios"; 3) we explicitly model the sequential nature of investment
and pricing decisions, whereby different decisions are made with differing
amounts of information; and 4) we examine in a few simple situations the
impact of the "irreversibility effect", i.e. the fact that generating
capacity is durable and hence one should build less of it than one would
choose if one could simply rent capacity for one year at a time.

Our model is fundamentally a normative model rather than a descriptive
model: we focus on how prices and investments should be made rather than
how they are in fact chosen under present regulatory regimes. Most of our
discussion focuses on short run marginal cost pricing, by which we mean
prices that are set in the last period after all investments have been
made, and chosen in the socially optimal manner. In choosing our marginal
cost prices, we use the pricing rules derived in chapters IV and V.

Recognizing that prices are currently not set equal to short run
marginal costs, we also examine two alternative pricing rules: exogenous
pricing and average cost pricing. By exogenous pricing we mean any pricing
rule which does not respond to generating capacity choices. One such pricing rule which we use is long run marginal cost pricing: such prices do not depend on the actual choice of generating capacity but rather are based on the cost minimizing choices of capacity.\(^2\) As in chapter VII section C, by average cost pricing we mean prices which are set so as to exactly recover all fixed plus operating costs. In doing so we ignore the serious problems caused by regulatory lags and inflation, which may result in revenues that are above or below true average costs.\(^3\) Nonetheless average cost pricing is perhaps a reasonable approximation of current day pricing practices.

VIII. A. Ten Modeling Frameworks

Table 8-1 classifies ten different investment "modeling frameworks", or sets of assumptions, in terms of their treatment of prices, demand uncertainty, the timing of investment and pricing decisions, and the durability of capital. At the simplest level are one period capital models which ignore uncertainty and treat prices as if they are not affected by the choice of generating capacity. Models using framework I, (such as those described in EPRI(1978) and EGEAS(1980)) employ essentially a two step procedure. First demand is projected into the future, either implicitly or explicitly using some exogenous set of prices. Next, investments are chosen so as to minimize the cost of meeting this projected demand. This "open loop" modeling framework ignores the feedback effect of investments on prices. Since revenues, profits and rates of return are often quite sensitive to prices, this procedure may introduce serious distortions in pictures of future scenarios.

Modeling framework II uses a "closed loop" procedure to solve for
Table 8-1
Alternative Investment Modeling Frameworks

<table>
<thead>
<tr>
<th>capital treated as lasting one period</th>
<th>capital treated as durable and irreversible</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>VI</td>
</tr>
<tr>
<td>II</td>
<td>VII</td>
</tr>
<tr>
<td>III</td>
<td>VIII</td>
</tr>
<tr>
<td>IV</td>
<td>IX</td>
</tr>
<tr>
<td>V</td>
<td>X</td>
</tr>
</tbody>
</table>

Single Scenario Approaches
Uncertainty Ignored
When Choosing Capacity

- prices exogenous when investments are made
- prices and investments chosen simultaneously

Multiple Scenario Approaches:
Uncertainties Incorporated in Objective Function

- prices exogenous when investments are made
- prices and investments chosen simultaneously
- prices and investments chosen sequentially
prices and investments which are mutually consistent. The price feedback loop is explicitly taken into account so that investments minimize the cost of serving the demand arising not from some exogenously forecast set of prices, but rather from the prices which will in fact arise if the planned investments are made. This second modeling framework underlies the models of Rowse (1980), Finger (1980) and Baughman, Joskow, and Kamat (1979). The latter is in a sense a hybrid of the first two modeling frameworks: investments are chosen without considering their impact on prices whereas prices and hence realized demand are based on the chosen set of investments. They use this partially closed loop procedure as a means of mimicking the naive forecasting techniques of many existing utilities.

Modeling frameworks III, IV and V explicitly take into account uncertainty about future demand while continuing to assume that capital lasts only one period. Modeling framework III is an "open loop" framework analogous to the first, in which prices are chosen exogenously. Framework IV is again a "closed loop" framework in which the effects of investments on prices are taken into account, however all pricing and investment decisions are assumed to take place simultaneously. Modeling framework V, a "closed loop" framework, explicitly takes into account the sequential nature of investment and pricing decisions. The sequential nature of decisions takes on particular importance in the presence of uncertainty: decisions which are deferred can be based on better information than decisions which must be made with long lead times.

We are not aware of any empirical models which explicitly take into account uncertainty in the sense that we do in our simulation model. When we say uncertainty is taken into account, we mean that the existence of uncertainty is explicitly incorporated in the objective function which is
being maximized. The conventional procedure for considering the impact of uncertainties is to do sensitivity analysis. After choosing investments (and perhaps prices) for some median or "most likely" future scenario, alternative future scenarios are examined to see how sensitive the investment (and other) choices are to the assumptions underlying the particular scenario selected. Unless weights (probabilistic and/or derived from a social welfare function) are attached to these alternative futures in the choice of investments, these choices will not be made in a socially optimal manner. Fuss and McFadden (1971) have shown, and chapters IV, V and VI above elaborated upon the fact that uncertainty has a systematic effect on desired investments. Specifically, under demand uncertainty, choices which allow greater production flexibility even at a loss of some production efficiency should be preferred. One of the issues we examine in this chapter is the extent to which ignoring uncertainty and basing decisions on a single median demand growth scenario distorts decisions seriously. We find that in many cases it does not, although in some cases base load capacity choices are distorted by 5 percent and more.

The first five modeling frameworks are all based on the assumption that generating capacity lasts only one period. Five alternative modeling frameworks exactly analogous to the first five can be used incorporating the reality that generating capacity is in fact durable. In this chapter we discuss only the last of these durable capital models, framework X. We take it as given that in the real world capital is durable; prices are influenced by investment choices; investments must be chosen sequentially; and there is uncertainty about what future demand will be. Under these circumstances the tenth modeling framework is clearly the most realistic; other modeling frameworks are justifiable only on the grounds of
computational complexity. The simulation model used in this paper, and described in section B below, chooses optimal investments and prices using a set of assumptions corresponding to modeling framework X. Other modeling frameworks can be thought of as special cases of this more complex and realistic set of assumptions.

As a means of organizing our discussion, we start with the simplest modeling framework and add successive degrees of complexity. In section C uncertainty and the durability of capital are ignored while comparing modeling frameworks I and II, and the importance of considering pricing effects is demonstrated. It is shown that the social costs of constructing too little base load generating capacity can be substantially reduced by adopting marginal cost pricing, and raising prices accordingly. Regulatory constraints are unlikely to allow prices to increase in such a manner however. When we use the somewhat more realistic assumption that prices are adjusted only so as to recover total generating costs, i.e. average cost pricing, we get the perverse result that prices are lower when too little base load capacity is constructed, resulting in massive shortages if no other action is taken.

Uncertainty about future demand is added to the model in section D while still maintaining the assumptions that prices and investments are made simultaneously, and that capital lasts only one period. The optimal choice of base load and peaking capacity is compared under sets of assumptions corresponding to the frameworks II and IV in table 8-1, and it is shown that as uncertainty increases, desired base load capacity declines while desired peaking capacity increases.

Section E adds a new complication to the decision making process by incorporating the fact that investment and pricing decisions are made
sequentially rather than all at once. Modeling framework V is used to examine alternative sequences of decisions over time, and it is shown that the ability to choose prices and peaking capacity after most of the uncertainty is resolved confers significant social benefits. Section F contrasts modeling frameworks V and X, and thus incorporates the assumption that generating capacity is durable and irreversible. The "irreversibility effect" is found to be small, but sensitive to both the rate of demand growth and the degree of uncertainty about that future demand growth. Section G examines the sensitivity of our results to the underlying cost and demand parameters. It is speculated that our uncertainty results may be due in part to the simplifying assumptions we have made about the way in which uncertainty is resolved over time and in part to the fact that we allow choices to be made between only two generating technologies.  

In section H we implement two of the extensions of chapter VII, namely the extension to the cases of random operating costs and average cost pricing. As our theoretical model predicts, with risk neutral consumers and producers, increasing operating cost uncertainty which is uncorrelated with demand uncertainty increases the desired use of the technology with uncertain operating costs. Use of this uncertain fuel cost technology decreases as the correlation between operating costs and the level of demand increases. Under average cost pricing the theoretical results that average cost pricing exacerbates consumption uncertainty and that it is not optimal to minimize expected total costs is demonstrated. Using the approach outlined in chapter VII we apply weights to each scenario in choosing capacity optimally and illustrate how this improves welfare. The effects of both operating cost uncertainty and average cost pricing on base load capacity choices remain small in all cases.
Section I returns to the questions posed in the opening paragraph of this chapter. The impact of a one third reduction in base load generating capacity on prices, demand, shortages, profits and investments is examined using the tenth modeling framework. The picture of the future which is painted using this tenth modeling framework is seen to be substantially less frightening than the picture which emerges when prices, uncertainty and sequential decision making are ignored. Generalizations and a discussion of the policy implications of our simulation model are contained in section J.

VIII.B. The Simulation Model

Only two types of generating capacity are presumed to be available: base load capacity which must be chosen with a ten year lead time, and peaking capacity which we examine for lead times of ten years and one year. Base load costs and peaking costs are based on hypothetical nuclear and gas turbine costs, respectively, for 1990, although costs corresponding to other types of generating capacity could obviously be used instead. Capital costs, operating costs, and other plant and demand characteristics are given in Table 8-2. All costs are in 1980 dollars. Capacity of each type is treated as a continuous variable, i.e. there is no lumpiness problem. Forced outages of existing capacity are dealt with by derating capacity. Hence 100 megawatts (1 MW = 10^6 kilowatts) of generating capacity which will be available with only an 80 percent probability is treated as 80 MW of generating capacity which will be available continuously. All transmission and distribution issues are ignored. For most of the analysis we will be dealing with investment
Table 8-2

Data Assumptions Used in Simulations

<table>
<thead>
<tr>
<th></th>
<th>Nuclear (base load)</th>
<th>Gas Turbine (peaking)</th>
<th>Shortages</th>
</tr>
</thead>
<tbody>
<tr>
<td>All costs in $1980</td>
<td>units</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capital cost (a)</td>
<td>$/kw</td>
<td>1200</td>
<td>200</td>
</tr>
<tr>
<td>Availability factor (b)</td>
<td>.65</td>
<td>.75</td>
<td>1.00</td>
</tr>
<tr>
<td>Fixed charge rate (c)</td>
<td>.15</td>
<td>.15</td>
<td>0</td>
</tr>
<tr>
<td>Effective capital cost</td>
<td>cents/kwhr</td>
<td>3.16</td>
<td>.46</td>
</tr>
<tr>
<td>Operating and maintenace</td>
<td>cents/kwhr</td>
<td>.20</td>
<td>.40</td>
</tr>
<tr>
<td>Fuel cost</td>
<td>cents/kwhr</td>
<td>.80</td>
<td>4.00</td>
</tr>
<tr>
<td>Shortage costs</td>
<td>cents/kwhr</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Total operating cost</td>
<td>cents/kwhr</td>
<td>1.00</td>
<td>4.40</td>
</tr>
<tr>
<td>Lead times</td>
<td>years</td>
<td>10</td>
<td>10 or 1</td>
</tr>
</tbody>
</table>

System load factor = .65

Demand scaled to yield a 1980 demand of 37.20 billion kwhrs

Short run price elasticity = -.5

Most likely annual growth rate = 3 percent

Most likely 10 year cumulative growth 34.392%
  with a 90% confidence interval of +11.59% to 61.86%
  if prices are held constant

Source: Baughman et al (1979) except nuclear capital cost
and gas turbine operating costs revised upward to reflect recent phenomena
decisions and pricing decisions for a single period (year) and hence we will ignore the time subscripts which could be used to indicate the year for which the decisions are being made. For simulations in which capital is treated as durable, and hence it is necessary to consider more than one year, we hold all capital costs and operating costs constant in real terms. This assumption substantially simplifies the calculation of optimal investments.11

Every generation planning model must cope in some way with the question of system reliability. System reliability is frequently treated by maximizing an objective function subject to a fixed loss of load probability constraint.12 We choose to use instead the simpler and exactly equivalent technique of introducing a "shortage technology" with a zero capital cost and a constant shortage cost ("operating cost") per kilowatt-hour (kw-hr).13 This procedure is equivalent to substituting the loss of load probability constraint directly into the objective function, and has been widely used in French investment planning models.14 A frequent "rule of thumb" used by electric utilities is a loss of load probability of one day in ten years (.027 percent). Given our peaking unit cost parameters this level of reliability would only be justified if shortages imposed an average burden of $16.83 per kw-hr of unserved demand.15 This implicit shortage cost should be contrasted with the estimate of $.53 (1980$) by Anderson and Perl (1980), and an upper bound of $1.22 (1970$) for New York calculated by Telson (1975). In the face of such a wide discrepancy, we used a compromise shortage cost of $2.00 per kw-hr, which implies a loss of load probability of .88 days per year (.24 percent).16

The quantity of electricity demanded in period t, D_t, is modeled as
a function of a single price, $P_t$, and a random variable, $\eta_t$, which we discuss below. Specifically we assume

$$D_t(P_t, \eta_t) = P_t^{Y} \eta_t$$  \hspace{1cm} (8-1)$$

This demand equation should be thought of as a reduced form forecasting equation. We use a "medium run" price elasticity of $-.5$ for the simulations presented here. This price elasticity was chosen as reflecting a compromise between the low short run price elasticities ($-.1$ to $-.4$) and high long run price elasticities ($-.8$ to $-2$) which have been reported.$^{17}$ Since we are using a reduced form forecasting equation which excludes previous period prices, it is reasonable to use an intermediate value. We also present the results of simulations using elasticities of $-.2$ and $-.8$ below.

One of the distinguishing features of electricity is that the rate at which it is used throughout the year varies. The resulting pattern of demand is usefully summarized in a load duration curve. As in chapter III, the load duration curve should be thought of as the probability distribution of demand conditional on prices, $P_t$, and the realization of the random variable $\eta_t$. Here the load duration curve is assumed to be linear, that is, for any given year the rate at which electricity is demanded is uniformly distributed between two constants which are chosen such that total demand is $D_t(P_t, \eta_t)$. Although this linear load duration curve is frequently used as a representation of the true load duration curve for purposes of approximating production costs,$^{18}$ in retrospect this may not have been the best representation to use for illustrating the impact of uncertainties. This is because the only
convexity or nonconvexity in \( 1-G(K|\eta) \) arises from the truncation of  
\( 1-G(K|\eta) \) at 1 and 0. Uncertainty has a significant impact only in the  
tails of the distribution, near these truncation points with a linear load  
duration curve. For this reason and another discussed below, our  
simulations probably underestimate the impact of uncertainties.  

The load factor (i.e. the ratio of average demand to peak demand) is  
exogenously specified, and hence for each level of demand there corresponds  
a unique load duration curve.\(^{19}\) The load duration curve should be thought  
of as reflecting the short run fluctuations which are resolved after prices  
and investments are chosen. An example of a linear load duration curve is  
shown in Figure 8-1. A linear load duration curves speeds the calculation  
of optimal prices, but this linearity assumption is not critical for our  
model. We have simulated the model for nonlinear curves without the  
computational burdens becoming overwhelming. See Appendix E for a  
discussion of this issue.  

In order to introduce long term uncertainty into our model, we assume  
\( \eta_t \) is stochastic and hence can be forecast only with error. The random  
variable \( \eta_t \) should be thought of as a one dimensional proxy for all of  
the uncertain variables affecting electricity demand, such as income and  
the stock of energy using capital. Specifically we model \( \eta_t \) as a first  
order Markov process,  

\[
\eta_t = \eta_{t-1} \varepsilon_t \tag{8-2}
\]

where the \( \varepsilon_t \) are independently and identically distributed lognormal  
random variables, and hence demand is itself lognormally distributed.  
Uncertainty and demand growth in our model are parameterized through
Figure 8-1

Linear Load Duration Curve

Fraction of the year the demand is below level $x$
choices of the mean and variance of the normal distribution underlying the
lognormal distribution of $\eta_t$. This lognormal distribution is then
approximated using 100 Monte Carlo drawings.

For our "base case" we use a set of assumptions corresponding to a
median demand growth scenario of 34.39 percent over ten years, with a 90
percent confidence interval of 11.59 to 61.86 percent when prices are held
fixed. This corresponds to an average annual demand growth of 3.0 percent
per year with a 90 percent confidence interval of 1.1 to 4.9 percent.
Since, as we will see, average cost pricing results in prices which do not vary
much with the realization of $\eta$, this fixed price confidence interval
is very similar to the confidence intervals resulting from average cost
pricing. Demand was scaled to be 37.20 billion kilowatt hours in 1980.
This corresponds fairly closely to actual use in Massachusetts in that
year, and hence all demand, cost and capacity figures can be thought as
applying to an imaginary utility serving all of Massachusetts in 1990.20

One of the objectives our simulation model is to choose investments and
prices which are mutually consistent. Hence we desire prices which are
optimal given the level of investments, and investments which minimize the
total social cost of meeting the demand which arises from these same
prices. Our model does this by using the pricing and investment rules
derived in chapters IV and V above, which amounts to solving the first
order conditions for a social optimum. We continue to take our social
objective function to be expected consumer plus producer surplus. Since
these investment and marginal cost pricing rules change as we adopt
different modeling assumptions, we do not discuss them here, but rather
discuss them as we add each new level of complexity. The algorithms we use
are described in Appendices E and F.
One of the advantages of our simple demand specification is that we are able to derive a simple expression for Marshallian consumer plus producer surplus. This provides a convenient check that we have in fact found a social optimum, and we report such surplus measures here. It is worth noting however that the ability to compute consumer plus producer surplus is not a necessity for the use of our optimization algorithm. This is important since for more general demand specifications (e.g. nonlinear load duration curves) it will frequently not be feasible to derive such measures.

VIII.C. Price Effects: Frameworks I and II

We begin by considering the simplest investment modeling framework, framework I in table 8-1, in which prices are exogenous, uncertainty about future demand is ignored, and generating capacity is treated as lasting only one period. This framework is the most prevalent modeling framework. For instance, a survey of six major regional generation planning models by Lee, Stoughton and Badertscher (1978) does not discuss price feedback loops as being a part of any of these six models. These models can be used for "sensitivity analysis", i.e. the analysis of scenarios other than the scenario for which capacity is explicitly chosen. Such a procedure does not mean that the choices are optimal (in the sense of this paper) given the uncertainty about which demand scenario will eventually occur. Furthermore, since prices are not considered explicitly, little can be said about revenues and profitability, two variables of considerable interest to today's utilities.

With exogenous prices and no uncertainty, the investment planning
problem is straightforward. Letting $G(x|P, \eta)$ be the probability that demand is below level $x$ given prices $P$ and the realization of random variable $\eta$, we have shown in chapter IV above that the cost minimizing choices of base and peaking capacity, $k_1$ and $k_2$, respectively, can be found by solving

$$c_1 + m_1[1 - G(k_1|P, \eta)] = c_2 + m_2[1 - G(k_1|P, \eta)] \quad (8-3)$$

$$c_2 + m_2[1 - G(k_1+k_2|P, \eta)] = \alpha [1 - G(k_1+k_2|P, \eta)] \quad (8-4)$$

where

$c_1, c_2 =$ annual capital cost per kw of base and peak capacity, respectively
$m_1, m_2 =$ operating cost per kw-year of base and peak capacity, respectively
$\alpha =$ shortage cost per kw-year

The first and second columns of Table 8-3 were derived using modeling framework I: capital is assumed to last only one period, prices are held constant at long run marginal costs, and uncertainty about the future is ignored. The first column presents our base case, where investments are chosen optimally using equations (8-3) and (8-4) for the median realization of $\eta$. Since capacity is chosen optimally and there is no long term uncertainty, long run and short run marginal costs are equal. At a price of 4.7 cents per kilowatt-hour (1980$\$), 1990 demand is 50.0 billion kw-hrs. Even without any uncertainty about future demand it is still not optimal to eliminate power shortages completely. To do so would involve purchasing some peaking capacity which would be used only a few hours each year. Our shortage, operating, and fixed cost assumptions result in unserved energy of 100 thousand kw-hrs (0.002 percent of total energy) and a loss of load.
Table 8-3  
Simple Pricing Effects With No Uncertainty

<table>
<thead>
<tr>
<th>variables</th>
<th>units</th>
<th>(1) base case</th>
<th>(2) 1/3 less nuclear capacity</th>
<th>(3) marginal cost pricing</th>
<th>(4) average cost pricing</th>
</tr>
</thead>
<tbody>
<tr>
<td>price</td>
<td>cts/kwhr</td>
<td>4.71</td>
<td>4.71</td>
<td>6.44</td>
<td>4.69</td>
</tr>
<tr>
<td>demand</td>
<td>$10^9$kwhrs</td>
<td>50.00</td>
<td>50.00</td>
<td>42.75</td>
<td>50.09</td>
</tr>
<tr>
<td>shortages</td>
<td>$10^9$kwhrs</td>
<td>.0001</td>
<td>1.23</td>
<td>.001</td>
<td>1.26</td>
</tr>
<tr>
<td>total surplus</td>
<td>$10^6$</td>
<td>25974</td>
<td>23577</td>
<td>25899</td>
<td>23526</td>
</tr>
<tr>
<td>change in total</td>
<td>$10^6$</td>
<td>(base)</td>
<td>-2397</td>
<td>-75</td>
<td>-2448</td>
</tr>
<tr>
<td>surplus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>profits</td>
<td>$10^6$</td>
<td>.30</td>
<td>7.99</td>
<td>728.94</td>
<td>0.</td>
</tr>
<tr>
<td>shortage costs</td>
<td>$10^6$</td>
<td>.30</td>
<td>2462.</td>
<td>2.9</td>
<td>2515.</td>
</tr>
<tr>
<td>capital costs</td>
<td>$10^6$</td>
<td>1276</td>
<td>916.</td>
<td>916.</td>
<td>916.</td>
</tr>
<tr>
<td>operating costs</td>
<td>$10^6$</td>
<td>1077.</td>
<td>1371.</td>
<td>1110.</td>
<td>1374.</td>
</tr>
<tr>
<td>base capacity</td>
<td>$10^6$ kw</td>
<td>6.00</td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
</tr>
<tr>
<td>peak capacity</td>
<td>$10^6$ kw</td>
<td>7.49</td>
<td>7.49</td>
<td>7.49</td>
<td>7.49</td>
</tr>
<tr>
<td>total capacity</td>
<td>$10^6$ kw</td>
<td>13.49</td>
<td>13.49</td>
<td>13.49</td>
<td>13.49</td>
</tr>
</tbody>
</table>

Column (1), the base case, was derived by choosing investments and prices optimally for a single demand scenario.

Column (2) prices and peaking capacity the same as in the case, only base load (nuclear) capacity reduced by one third.

Column (3) base and peaking capacity is the same as in column (2), except prices are set equal to marginal cost given the level of capacity

Column (4) same as column (3) only average instead of marginal cost pricing used.
probability of .88 days per year.\textsuperscript{25}

Column 2 of table 8-3 holds prices and the quantity of peaking capacity fixed while addressing the question: what happens if the quantity of base load capacity constructed is one third less than the socially optimal level? In some parts of the country this is already occurring due to regulatory constraints on the construction of coal and nuclear generating capacity. Relative to our hypothetical utility serving Massachusetts the reduction of base load capacity from 6 to 4 million kw is roughly equivalent to cancelling two new nuclear or coal plants.\textsuperscript{26} The picture which emerges if neither prices nor peaking capacity is adjusted in response to this reduction in base load capacity is indeed quite grim. Shortages have grown to 1.23 billion kw-hrs (2.5 \% of total demand) and are occurring a horrendous 78 days a year. The loss of total consumer plus producer surplus amounts to $2.4 billion per year. Capital costs have been reduced by almost one fourth, but this is offset by a rise in operating costs since the utility is forced to rely more heavily on fuel intensive peaking capacity. Revenues are down slightly since the utility does not collect revenue for unserved demand. For the cost and demand parameters which we have chosen, profits have gone up slightly to $8 million. These positive profits simply reflect the fact that marginal costs are greater than average costs when there is a shortage of generating capacity.

The dire picture painted by column 2 emerges largely because we have ignored prices. We have also ignored the fact that the shorter lead time needed to build peaking capacity means that the shortfall of base load capacity can (and barring serious regulatory constraints, will) be made up for by peaking capacity. We focus here on price effects, deferring until section I the consideration of how peaking capacity should be reoptimized.
In order to consider pricing effects, some procedure is needed to choose prices. Each of our ten modeling frameworks is consistent with either marginal cost pricing or average cost pricing. Since the pricing rule for average cost pricing is straightforward, we focus our attention on socially optimal marginal cost pricing.\textsuperscript{27} For the case where all electricity is sold at a single price, it is shown in chapter IV that the socially optimal price, \( P^* \), can be found by solving\textsuperscript{28}

\[
P^* = \frac{\frac{\partial C(k_1, k_2, P^*)}{\partial P}}{\frac{\partial D(P^*)}{\partial P}}
\]

where

\( C(k_1, k_2, P^*) = \) total generating plus shortage costs,
given \( k_1, k_2, \) and \( P^* \).

\( D(P^*) = \) total demand given price \( P^* \).

Except for in very special cases the numerator is not readily calculated analytically, hence we use the following discrete approximation:

\[
P^* = \frac{C(k_1, k_2, P^*) - C(k_1, k_2, P^* + \xi)}{D(P^*) - D(P^* + \xi)} \]  \text{[\( \xi \) small]}

This expression of course looks very much like a marginal cost. The difference is that since operating costs vary with the level of demand, we need to take a weighted average of the marginal costs of serving that demand throughout the year. As shown in Appendix C and discussed in chapter IV, section B, the appropriate weights on the marginal cost of serving demand are the product of 1) the level of demand, 2) the price elasticity of demand at that level, and 3) the probability that demand is at that level. This formula automatically weights these three variables
appropriately.

In modeling frameworks in which prices and investments are chosen at the same time, equations (8-3),(8-4) and (8-6) can be solved simultaneously to find the socially optimal set of prices and investments. We used an iterative procedure to solve these first order conditions which continued to work well once uncertainty was introduced even though we were solving for 100 different prices corresponding to 100 different scenarios. The algorithm used is described in Appendix E.

Column 3 of table 8-3 is based on modeling framework II, and incorporates the possibility that prices can be adjusted upward to reflect the higher operating costs that would be experienced without the desired base load capacity. The price was chosen to be equal to short run marginal costs using equation (8-6) while holding peaking and base load capacity at the same levels as shown in column 2. Column 3 therefore corresponds to building too little base load capacity, and then adjusting prices optimally so as to mitigate the impact of this capacity shortfall. A substantially different picture emerges. By allowing the price to rise 36.8 percent, from 4.71 to 6.44 cents per kw-hr, our price elasticity of -.5 yields a 14.5 percent reduction in demand. Shortages and shortage costs have both been reduced by a factor of almost 1000, and capacity induced shortages are now occurring only 1.68 days per year. Since our demand curve is inelastic, the new higher price results in greater revenue as well as lower operating costs: annual profits have grown to $729 million. The loss of consumer plus producer surplus has been reduced from $2.4 billion to $75 million. The social cost of a base load capacity shortage has been substantially lessened. The important conclusion to be drawn from this discussion is simply that capacity choices affect the level of prices which
can have a major impact on demand, system reliability, operating costs and profits.

Given the present regulatory climate it is unlikely that a utility would in fact be allowed to charge the high and very profitable price that we have calculated here, even though given the assumed levels of capacity this would be the socially optimal price to charge. In our simple one price model a more realistic pricing assumption might be that the electric utility would be forced to adopt average cost pricing. This leads to substantially different pricing behavior when too little base load capacity is built.

Simulations corresponding to columns 1, 2 and 3 were made under the assumption of average cost pricing instead of marginal cost pricing using the optimal investment rules developed in chapter VII, section C. Given our uniform price elasticity of demand assumption it came as no surprise that the average cost price when base and peak capacity are chosen optimally is slightly below the socially optimal marginal cost price.\(^{29}\) This is because, as we demonstrated in sections IV.B and VII.C, the socially optimal price reflects the cost of shortages whereas average costs do not. The average cost simulation analogous to simulation 1 yielded a price of 4.7058 versus 4.7064 from simulation 1. (When investments are made optimally shortage costs are small and we expect the two prices to diverge only slightly!) Since other variables also differ only in the fourth and fifth significant digits, we do not present them here. Similarly, since prices are held fixed between simulations 1 and 2, there is no need to present the naive impact of a one third reduction in base load capacity with the (average cost) price being held fixed, for it is the same as shown in column 2.
Marginal and average cost pricing differ markedly, however, when capacity is not chosen optimally. Column 4 of table 8-3 indicates what the impact would be of a one third reduction in base load capacity, holding peaking capacity fixed, when average cost pricing is in effect. Instead of raising prices to reflect higher operating and shortage costs, as occurred with marginal cost pricing, the price has been lowered to reflect reduced capital costs, which slightly outweigh increased operating costs. As a result of this perverse pricing behavior, demand and shortages have both increased slightly, and there has been a further $51 million reduction in consumer plus producer surplus on top of the already horrendous losses caused when prices were held fixed.

This average cost pricing result is very unlikely to occur; in the face of potentially massive power shortages it is much more likely that additional peaking capacity would be added to the system, almost regardless of the cost.\textsuperscript{30} This is facilitated by the shorter lead times needed for constructing peaking capacity. Before considering how the ability to rely on peaking capacity at the last minute affects prices, costs and shortages, it is useful to first consider the effects of uncertainty. To a great extent it is the existence of uncertainty which makes the varying lead times for generating capacity an important variable: since different types of generating capacity have different lead times, decisions about investments must be made with differing amounts of information.\textsuperscript{31} We reexamine our "one third less base load capacity" picture of the future below in section I.

In this section we have done a simple comparative statics analysis of the impact of a shortfall of base load capacity. The social impact of "one third less base load capacity" which is not made up for by peaking capacity
was seen to be drastic if prices are held fixed a la framework I, but
substantially lessened if prices are allowed to rise so as to equal short
run marginal costs, as would be indicated by framework II. Framework II
was also used to analyze one third less base load capacity under average
cost pricing. The average cost price was found to be lower with the
capacity shortfall rather than higher, a perverse pricing result which
exacerbates the capacity shortage. Although it is likely that responses
other than price adjustments could be taken to the capacity shortfall, the
results of this section strongly suggest that pricing effects may have
dramatic effects on demand and desired investments.

VIII.D. Uncertainty and Simultaneous Decision Making: Framework IV

In the previous section we examined prices and investments in models
without any uncertainty: investments and prices were chosen for a single
demand scenario, which in our model corresponds to a single realization of
the random variable $\eta$. In this section we introduce uncertainty about
future demand into our model by using the fourth modeling framework above.
We continue to make two strong assumptions: that all pricing and investment
decisions are made simultaneously, and that capital lasts only one period.
Although these two assumptions are not widely used in empirical planning
models, they are frequently used in theoretical models. For instance,
Dansby (1975), Crew and Kleindorfer (1976, 1978), and Turvey and Anderson
(1977) all focus on models in which prices and investments are chosen
simultaneously. We relax these two assumptions in the next two
sections.

With uncertainty about future demand, it is no longer optimal to
choose capacity for any single demand scenario. Instead, as we have shown above, it is desirable to use a probability weighted average of all of the possible scenarios. Here we assume that base capacity, peaking capacity and a single price must all be chosen ten years before demand is realized. Decisions are made by giving equal weight to 100 drawings of the lognormally distributed random variable $\eta$, which as we have noted previously is assumed here to affect demand multiplicatively. The important simplification which we use to keep the problem tractable is that we collapse all of these 100 different demand scenarios into a single probability distribution of demand, our load distribution curve. As we have noted previously, the load distribution curve is the marginal probability distribution of demand, and is derived by taking expectations over all possible load duration curves, the conditional probability distributions of demand. Using this load distribution curve socially optimal prices and investments can be found by solving equations exactly analogous to equations (8-3), (8-4) and (8-6) above, only using the load distribution curve instead of the load duration curve. Multiple dimensions of long term demand uncertainty are collapsed into a one dimensional probability distribution of demand, after which there is no fundamental difference in the way in which investments and prices should be selected with and without long term demand uncertainty. A more complete description of the procedure we used can be found in Appendix E.

Table 8-4 illustrates how prices and investments should optimally be readjusted as uncertainty about the future increases. Column 1, which duplicates column 1 of table 8-3 for convenience, assumes that there is no uncertainty about future demand, and that demand grows 34.4 percent over the next ten years (3.0 percent per year). Column 2 is based on a moderate
Table 8-4
Effects of Increasing Uncertainty With Simultaneous Pricing and Investment Decisions

<table>
<thead>
<tr>
<th>Marginal Cost Pricing</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>variables</td>
<td>units</td>
<td>no</td>
<td>moderate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>uncertainty</td>
<td>uncertainty</td>
</tr>
<tr>
<td>price</td>
<td>cts/kwhr</td>
<td>4.71</td>
<td>4.89</td>
</tr>
<tr>
<td>demand</td>
<td>$10^9$kwhrs</td>
<td>50.00</td>
<td>49.16</td>
</tr>
<tr>
<td>shortages</td>
<td>$10^9$kwhrs</td>
<td>.001</td>
<td>.007</td>
</tr>
<tr>
<td>total surplus</td>
<td>$10^6$</td>
<td>25974</td>
<td>25935</td>
</tr>
<tr>
<td>profits</td>
<td>$10^6$</td>
<td>.30</td>
<td>18.29</td>
</tr>
<tr>
<td>shortage costs</td>
<td>$10^6$</td>
<td>.30</td>
<td>14.99</td>
</tr>
<tr>
<td>capital costs</td>
<td>$10^6$</td>
<td>1276.</td>
<td>1313.</td>
</tr>
<tr>
<td>operating costs</td>
<td>$10^6$</td>
<td>1077.</td>
<td>1087.</td>
</tr>
<tr>
<td>base capacity</td>
<td>$10^6$ kw</td>
<td>6.00</td>
<td>5.83</td>
</tr>
<tr>
<td>peak capacity</td>
<td>$10^6$ kw</td>
<td>7.49</td>
<td>9.98</td>
</tr>
<tr>
<td>total capacity</td>
<td>$10^6$ kw</td>
<td>13.49</td>
<td>15.81</td>
</tr>
</tbody>
</table>

column (1) based on no uncertainty, most likely ten year demand growth of 34.39 percent

column (2) based on a most likely ten year demand growth of 34.39 percent with a 90 percent confidence interval of 11.59 to 61.86 percent

column (3) based on a most likely ten year demand growth of 34.39 percent with a 90 percent confidence interval of -7.4 percent to 94.9 percent
amount of uncertainty about future demand. The median demand growth remains 34.4 percent, but the 90 percent confidence interval for this demand growth is 11.6 percent to 61.9 percent. This confidence interval corresponds to average annual growth rates of 1.1 and 4.9 percent respectively. Column 3 is based on highly uncertain demand growth. The median cumulative demand growth is still 34.4 percent, but the 90 percent confidence interval is now -7.4 percent to 94.9 percent, corresponding to average annual growth rates of -8 to 6.9 percent, respectively. It should be recalled that these are fixed price confidence intervals, representing the underlying probability distribution of $\eta$. Pricing effects change these confidence intervals.

Four changes take place as uncertainty about future demand increases. First, the optimal price goes up. Prices are 3.8 and 9.3 percent higher for our moderate and high uncertainty scenarios, respectively. Second, shortages become more serious. Even though the loss of load probability is being held constant, the expected size of shortages when they do occur increases. Third, desired base load capacity declines. We discuss this result below. Fourth, desired peak capacity increases substantially. Even with moderate uncertainty, desired peaking capacity increases 33.2 percent. For the high uncertainty simulation, desired peaking capacity increases 86.8 percent.

In chapters IV and VI we distinguished three different effects of uncertainty on desired generating capacity: the direct effect, the price effect, and the irreversibility effect. First, if prices are held fixed, then increasing uncertainty raises the probabilities that demand is below low levels and above high levels. As Fuss and McFadden (1971) have pointed out, the desire for greater production flexibility results in less of the
capital intensive and more of the fuel intensive technology being chosen.\textsuperscript{35} We call this the direct effect since it ignores the feedback of investments on prices and on subsequent investments. A second effect of uncertainty on desired capacity occurs because capacity choices which are optimal \textit{ex ante} will in general differ from the choices which would minimize costs \textit{ex post}. As a result marginal generating costs will usually be higher, and prices should be adjusted upward.\textsuperscript{36} The price induced decline in demand reinforces the reduction in desired base load capacity, but partially offsets the increased reliance on peaking capacity. We call this the pricing effect. Both of these first two effects are reflected in table 8-4; base load capacity is reduced by 2.8 and 9.3 percent in the moderately and highly uncertain demand growth simulations, respectively. A third effect which further reduces the desired base load capacity is evident only when we introduce the fact that capacity is durable. We discuss this irreversibility effect in section F below.

In this section we have examined optimal pricing and investment decisions using a model in which all pricing and investment decisions must be made ten years before demand is realized. Uncertainty is seen to have a modest impact on all of the decision variables in the model, in the direction predicted by the theoretical model developed above. We now move on to consider the impact of sequential decision making.
VIII.E. Sequential Decision Making: Framework V

The most serious deficiency of the previous section is our assumption that all pricing and investment decisions are made simultaneously, ten years before all uncertainty about future demand is resolved. In this section we consider the impact which sequential decision making has on optimal choices, and demonstrate that there are significant benefits to being able to choose prices and investments after all or most demand uncertainty is resolved.

Sequential decision making substantially complicates the investment planning problem. Before, when we took a probability weighted average of many possible demand scenarios, investments and prices were being held constant across all such scenarios. With sequential decision making, choices of prices and peaking capacity made after base load capacity has been selected should reflect the improved information about future demand which has become available in the interim. If base load capacity is to be chosen optimally, then it must take into account the way in which future decisions will subsequently be made. Analytically the result is a series of nested optimizations, with each decision taking into account the manner in which all subsequent decisions will be made.

In chapters III through VI we derived a number of results analytically; however, we were forced to retreat to the strong assumption of exogenous prices before showing our results on sequential and irreversible investment decisions. Our simulation model enables us to examine sequential and irreversible investment decisions without assuming prices are exogenous. Once again, we use our technique of collapsing 100 alternative scenarios into a single probability distribution (our load
distribution curve) before choosing optimal base load capacity. Now, however, each scenario may vary not only because of the direct effect of random variable $\eta$, but also because of prices and peaking capacity which may themselves depend on $\eta$.

The following algorithm, described in more detail in Appendix E, was used to find optimal investments when both base and peaking capacity must be selected ten years in advance. An analogous procedure was used where only base load capacity requires a ten year lead time. First a trial set of investments was selected. Next 100 different realizations of $\eta$ were simulated from a lognormal distribution. For each realization of $\eta$ (i.e. each scenario) equation (8-6) above was solved to set price equal to marginal costs. These short run marginal costs depend on both $\eta$ and the previous choice of investments. The load duration curves corresponding to each of these 100 $\eta$ and price pairs were then collapsed into our load distribution curve. Equations (8-3) and (8-4) were used to find a new set of investments which minimizes the expected costs of meeting the levels of demand summarized in this load distribution curve. This new set of investments was then used with the same 100 realizations of $\eta$ to generate a revised set of prices. This iterative process between prices and investments was repeated until convergence, i.e. investments minimize the expected cost of meeting the demand that occurs when prices are set equal to marginal costs conditioned on those investments. The desirable outcome is that investments are made which take into account not only uncertainty about what the future demand will look like, but also the feedback effect of those investments in each future scenario.

Four different sequences of investment and pricing decisions are contrasted in Tables 8-5, 8-6, 8-7 and 8-8. In addition to reporting the
Table 8-5
Sequential Decision Making:
Base, Peak and Prices All Chosen With a Ten Year Lead Time

<table>
<thead>
<tr>
<th>Marginal Cost Pricing</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>variables</td>
<td>units</td>
<td>low growth scenario</td>
<td>median growth scenario</td>
<td>high growth scenario</td>
</tr>
<tr>
<td>price</td>
<td>cts/kwhr</td>
<td>4.89</td>
<td>4.89</td>
<td>4.89</td>
</tr>
<tr>
<td>demand</td>
<td>$10^9$kwhrs</td>
<td>40.73</td>
<td>49.05</td>
<td>59.08</td>
</tr>
<tr>
<td>shortages</td>
<td>$10^9$kwhrs</td>
<td>0.</td>
<td>0.</td>
<td>.005</td>
</tr>
<tr>
<td>total surplus</td>
<td>$10^6</td>
<td>21428</td>
<td>25906</td>
<td>31260</td>
</tr>
<tr>
<td>profits</td>
<td>$10^6</td>
<td>-62.86</td>
<td>22.65</td>
<td>98.40</td>
</tr>
<tr>
<td>shortage costs</td>
<td>$10^6</td>
<td>0.</td>
<td>0.</td>
<td>11.00</td>
</tr>
<tr>
<td>capital costs</td>
<td>$10^6</td>
<td>1312</td>
<td>1312</td>
<td>1312</td>
</tr>
<tr>
<td>operating costs</td>
<td>$10^6</td>
<td>743</td>
<td>1073</td>
<td>1489</td>
</tr>
<tr>
<td>base capacity</td>
<td>$10^6$ kw</td>
<td>5.83</td>
<td>5.83</td>
<td>5.83</td>
</tr>
<tr>
<td>peak capacity</td>
<td>$10^6$ kw</td>
<td>9.98</td>
<td>9.98</td>
<td>9.98</td>
</tr>
<tr>
<td>total capacity</td>
<td>$10^6$ kw</td>
<td>15.81</td>
<td>15.81</td>
<td>15.81</td>
</tr>
</tbody>
</table>

column (1) based on low growth scenario, the fifth percentile of demand

column (2) based on median growth scenario, fiftieth percentile of demand

column (3) based on high growth scenario, the ninety-fifth percentile of demand

column (4) shows expected values for each variable over all scenarios.
Table 8-6
Sequential Decision Making:
Base Load and Peak Capacity Chosen With a Ten Year Lead Time
Prices Chosen In the Last Period

<table>
<thead>
<tr>
<th>Marginal Cost Pricing</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>variables</td>
<td>units</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>price</td>
<td>cts/kwhr</td>
<td>3.85</td>
<td>4.52</td>
<td>6.49</td>
</tr>
<tr>
<td>demand</td>
<td>$10^9\text{kwhrs}$</td>
<td>45.89</td>
<td>51.04</td>
<td>51.28</td>
</tr>
<tr>
<td>shortages</td>
<td>$10^9\text{kwhrs}$</td>
<td>0.</td>
<td>.00007</td>
<td>.002</td>
</tr>
<tr>
<td>total surplus</td>
<td>$10^6$</td>
<td>21508</td>
<td>25973</td>
<td>31206</td>
</tr>
<tr>
<td>profits</td>
<td>$10^6$</td>
<td>-431.43</td>
<td>-97.30</td>
<td>916.10</td>
</tr>
<tr>
<td>shortage costs</td>
<td>$10^6$</td>
<td>0.</td>
<td>.15</td>
<td>3.70</td>
</tr>
<tr>
<td>capital costs</td>
<td>$10^6$</td>
<td>1288</td>
<td>1288</td>
<td>1288</td>
</tr>
<tr>
<td>operating costs</td>
<td>$10^6$</td>
<td>911</td>
<td>1114</td>
<td>1127</td>
</tr>
<tr>
<td>base capacity</td>
<td>$10^6\text{kw}$</td>
<td>6.03</td>
<td>6.03</td>
<td>6.03</td>
</tr>
<tr>
<td>peak capacity</td>
<td>$10^6\text{kw}$</td>
<td>7.74</td>
<td>7.74</td>
<td>7.74</td>
</tr>
<tr>
<td>total capacity</td>
<td>$10^6\text{kw}$</td>
<td>13.78</td>
<td>13.78</td>
<td>13.78</td>
</tr>
</tbody>
</table>

column (1) based on low growth scenario, the fifth percentile of demand.
column (2) based on median growth scenario, fiftieth percentile of demand.
column (3) based on high growth scenario, the ninety-fifth percentile of demand.
column (4) shows expected values for each variable over all scenarios.
Table 8-7
Sequential Decision Making:
Base Load Capacity Chosen With a Ten Year Lead Time
Peaking Capacity and Prices Chosen In The Last Period

<table>
<thead>
<tr>
<th>Marginal Cost Pricing</th>
<th>(1) low growth scenario</th>
<th>(2) median growth scenario</th>
<th>(3) high growth scenario</th>
<th>(4) average over all scenarios</th>
</tr>
</thead>
<tbody>
<tr>
<td>variables</td>
<td>units</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>price</td>
<td>cts/kwhr</td>
<td>4.45</td>
<td>4.71</td>
<td>4.92</td>
</tr>
<tr>
<td>demand</td>
<td>10^9 kwhrs</td>
<td>42.71</td>
<td>49.96</td>
<td>58.90</td>
</tr>
<tr>
<td>shortages</td>
<td>10^9 kwhrs</td>
<td>.0001</td>
<td>.0001</td>
<td>.0002</td>
</tr>
<tr>
<td>total surplus</td>
<td>$10^6</td>
<td>21557</td>
<td>25974</td>
<td>31270</td>
</tr>
<tr>
<td>profits</td>
<td>$10^6</td>
<td>-11.93</td>
<td>4.30</td>
<td>11.49</td>
</tr>
<tr>
<td>shortage costs</td>
<td>$10^6</td>
<td>.25</td>
<td>.30</td>
<td>.35</td>
</tr>
<tr>
<td>capital costs</td>
<td>$10^6</td>
<td>1218</td>
<td>1270</td>
<td>1333</td>
</tr>
<tr>
<td>operating costs</td>
<td>$10^6</td>
<td>800</td>
<td>1082</td>
<td>1450</td>
</tr>
<tr>
<td>base capacity</td>
<td>10^6 kw</td>
<td>5.96</td>
<td>5.96</td>
<td>5.96</td>
</tr>
<tr>
<td>peak capacity</td>
<td>10^6 kw</td>
<td>5.56</td>
<td>7.51</td>
<td>9.93</td>
</tr>
<tr>
<td>total capacity</td>
<td>10^6 kw</td>
<td>11.52</td>
<td>13.48</td>
<td>15.89</td>
</tr>
</tbody>
</table>

column (1) based on low growth scenario, the fifth percentile of demand
column (2) based on median growth scenario, fiftieth percentile of demand
column (3) based on high growth scenario, the ninety-fifth percentile of demand
column (4) shows expected values for each variable over all scenarios.
### Table 8-8

**Sequential Decision Making:**
Base, Peak and Prices All Chosen In The Last Period

<table>
<thead>
<tr>
<th>Marginal Cost Pricing</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>variables</td>
<td>units</td>
<td>low growth scenario</td>
<td>median growth scenario</td>
<td>high growth scenario</td>
</tr>
<tr>
<td>price</td>
<td>cts/kwhr</td>
<td>4.71</td>
<td>4.71</td>
<td>4.71</td>
</tr>
<tr>
<td>demand</td>
<td>$10^9$kwhrs</td>
<td>41.52</td>
<td>50.00</td>
<td>60.22</td>
</tr>
<tr>
<td>shortages</td>
<td>$10^9$kwhrs</td>
<td>.0001</td>
<td>.0001</td>
<td>.0002</td>
</tr>
<tr>
<td>total surplus</td>
<td>$10^6$</td>
<td>21567</td>
<td>25974</td>
<td>31282</td>
</tr>
<tr>
<td>profits</td>
<td>$10^6$</td>
<td>.24</td>
<td>.30</td>
<td>.33</td>
</tr>
<tr>
<td>shortage costs</td>
<td>$10^6$</td>
<td>.25</td>
<td>.30</td>
<td>.36</td>
</tr>
<tr>
<td>capital costs</td>
<td>$10^6$</td>
<td>1059</td>
<td>1276</td>
<td>1536</td>
</tr>
<tr>
<td>operating costs</td>
<td>$10^6$</td>
<td>895</td>
<td>1077</td>
<td>1298</td>
</tr>
<tr>
<td>base capacity</td>
<td>$10^6$ kw</td>
<td>4.98</td>
<td>6.00</td>
<td>7.23</td>
</tr>
<tr>
<td>peak capacity</td>
<td>$10^6$ kw</td>
<td>6.22</td>
<td>7.49</td>
<td>9.02</td>
</tr>
<tr>
<td>total capacity</td>
<td>$10^6$ kw</td>
<td>11.20</td>
<td>13.49</td>
<td>16.24</td>
</tr>
</tbody>
</table>

- **Column (1)** based on low growth scenario, the fifth percentile of demand
- **Column (2)** based on median growth scenario, fiftieth percentile of demand
- **Column (3)** based on high growth scenario, the ninety-fifth percentile of demand
- **Column (4)** shows expected values for each variable over all scenarios.
expected values of a number of interesting variables over all possible scenarios, we also report their values for three specific demand scenarios. These three scenarios correspond to a low demand growth scenario based on the fifth percentile of the probability distribution of $\eta$, a median growth scenario based on the 50th percentile, and a high growth scenario based on the 95th percentile. Together these three scenarios give an impression of how investments and prices vary across different scenarios. For table 8-5 prices, base load capacity, and peaking capacity are all chosen with a ten year lead time; for table 8-6 only prices are chosen in the last period. For table 8-7, base load capacity is chosen with a ten year lead time while both prices and peaking capacity are chosen in the last period. This is the most realistic set of assumptions to make of the four. For purposes of comparison, table 8-8 is based on the decidedly unrealistic assumption that all prices and capacity can simultaneously be chosen once the random variable $\eta$ is known. The four tables thus present four increasingly favorable sequences of investment and pricing decisions.

Several interesting relationships are revealed by a comparison of tables 8-5, 8-6, 8-7, and 8-8. The constancy of prices across different scenarios in table 8-5 stands in sharp contrast with the widely varying prices of table 8-6. Table 8-7 in which peaking capacity is is allowed to respond to the realization of $\eta$ is characterized by smaller price fluctuations than table 8-6, while prices are again constant (and equal to long run marginal costs) in table 8-8 when prices and all types of capacity can be chosen in the last period. This makes a fundamental point: the greater the flexibility in choosing capacity in the last period, the smaller the fluctuations in the socially optimal price. Of course, the greater the price variations the smaller the fluctuations in consumption.
The 90 percent confidence interval for total demand for the decision sequence which displays the greatest price dispersion (table 8-6) is (45.89, 51.28). This should be contrasted with the 90 percent confidence interval of (41.52, 60.22) when peaking capacity can be chosen in the last period (table 8-7). This trade off between demand and price uncertainty is also revealed by a comparison of tables 8-7 and 8-8.

The stabilizing effect of some or all capacity decisions being made in the final period also carries over to shortages and profits. Without any ability to adjust capacity in the last period, shortages are concentrated among high demand growth scenarios. Profits are highest when demand growth is high, and negative when demand turns out to grow slowly. The range over which shortages and profits vary diminishes as investment decisions are made closer to the resolution of uncertainty.

A comparison of consumer plus producer surpluses for different decision sequences reveals that there are decided welfare gains to be had by deferring pricing and investment decisions until long term uncertainties have been resolved. When all investments must be made ten years in advance, the ability to choose prices in the last period rather than ten years in advance improves welfare by $80 million in the low growth scenario, $67 million in the median growth scenario, and $54 million in the high demand growth scenario. When expectation are taken across all possible scenarios the expected gain in welfare is $65 million. The ability to choose peaking capacity as well as prices in the last period results in an additional gain of $18 million, while the incremental gain from choosing base load capacity in the last period, given that prices and peaking capacity are already being chosen at that time, is only $3 million. In all three cases the gains are larger when demand growth does not follow
the median growth scenario. What these comparisons tell us is that there are decided gains to be had from deferring at least some decisions until there is less uncertainty about future demand, however once one or two decision variables (e.g. prices and peaking capacity) are available with which to react to unexpected demand developments, there appears to be only modest gains to be had from further flexibility. These results have a direct interpretation in terms of the value of information: the information gained by deferring pricing and peaking capacity decisions for ten years when all long term uncertainty is resolved is worth $88 million per year.

Since we consider the most realistic set of "one period capital" assumptions to be those in table 8-7, for which base load capacity is chosen ten years in advance while prices and peaking capacity are chosen in the last period, it is interesting to compare the four columns in table 8-7 to column (1) of table 8-3, in which uncertainty is ignored and attention is focused only on the single median demand growth scenario. The comparisons can also be made using the median growth scenario of table 8-8: when all capacity and prices are chosen in the last period, it is the same as if there is no uncertainty, hence column (2) of table 8-8 is identical to column (1) of table 8-3. The striking fact is that for the levels of uncertainty considered here, taking uncertainty into account has only a slight impact on desired base load capacity. The ability to choose prices and peaking capacity in the last period reduces the costs of having too much or too little base load capacity, so that optimal base load capacity is only .67 percent below the level which would be chosen by simply using the median demand scenario while ignoring uncertainty. The theoretically predicted effect of exogenous long term demand uncertainty on
desired base load capacity is present but for the parameter values which we use in our base case it is of minor empirical relevance. We defer considering how sensitive these results are to our parameter assumptions until section I.

In this section we have extended our simulation model to examine the impact of alternative sequences of decision making on prices, demand, investments and other significant variables. It is seen that the ability to alter peaking capacity after much of the uncertainty has been resolved has a stabilizing effect on prices, system reliability and profitability. The improved information available by waiting until long term uncertainty is resolved before setting prices and choosing peaking capacity results in meaningful gains in expected consumer plus producer surplus. There are diminishing returns to this improvement in information, however, so that the level of base load capacity chosen ten years in advance is not significantly different from the level which is chosen if uncertainty is ignored and base load capacity is simply chosen for the median future scenario.

VIII.F. Durable Capital: Framework X

Up until now we have been assuming that generating capacity lasts only one period. This permits capacity choices to be made one period at a time without worrying about the impact of those choices on subsequent investment decisions. In this section we consider the polar opposite case, where capacity is assumed to last forever. The more realistic intermediate case where capacity lasts for thirty to fifty years would be much closer to this perfectly durable case than to the one period case.
Three new complications are introduced once it is recognized that capital is durable. One complication is that it becomes difficult to calculate the appropriate capital charge (rent) to be charged each period for each type of capacity. We have nothing to contribute to this issue here, and will in fact assume that rental charges are based on exogenous replacement costs. A second complication arises because investments should be made in discrete increments ("lumps") to take advantage of economies of scale. Forward and backward dynamic programming procedures can be used to find the optimal timing of these discrete investments. We have nothing to contribute to this issue either and avoid it by our assumption that investments can be made in continuous increments.

We focus here on the third complication introduced by the fact that capital is durable. This third complication arises only in the presence of uncertainty. Simply stated, the problem is that investments must be positive. The decision to build (rent) additional capacity today restricts the choices which will be available next year: *ex post* one has foregone the option of renting less of that capacity in the future. The consideration of this fact *ex ante* means that lower levels of capacity are optimal than those implied by static optimization rules. Henry (1973) and Arrow and Fisher (1974) labeled this the "irreversibility effect." In the more traditional literature the complication has been analyzed in terms of the difference between a putty-putty and a putty-clay technology.

Conceptually it is useful to think about the irreversibility effect in terms of option values. Whenever one rents (builds) an additional unit of generating capacity, one has implicitly foregone the option not to rent that capacity in the future. The value of that option should be taken into account when the original investment decision is made. This option value
is a fixed cost which should be added on to the original capital cost of new capacity when deciding how much of this generating capacity to construct. To state this in a slightly different way, whenever one adds new capacity, one is taking a risk of holding too much generating capacity in the future.\textsuperscript{45} Too much capacity implies costs in excess of the minimum cost capacity choices. The expected value of these excess costs is what a risk neutral investor should be willing to pay to avoid taking this risk. This is also the option value.

Once one begins to think in terms of option values it becomes apparent that there are many option values involved when a given investment is made. Suppose that in period $s$ a utility is already committed in period $t = s + 10$ to hold $s_{k_1,t}^1$ MW of base load capacity of base load and $s_{k_2,t}^2$ MW of peaking capacity. The decision to begin constructing an additional $I$ MW of base load capacity which will take ten years to construct means that one has foregone the option to hold less than $s_{k_1,t}^1 + I$ MW in periods $t + 1$, $t + 2$, $t + 3$, etc. One has also foregone the option to hold less than $s_{k_1,t}^1 + s_{k_2,t}^2 + I$ MW of base plus peaking capacity in periods $t$, $t + 1$, $t + 2$, etc. Unless one has an excess of peaking capacity, the second option has very little value. Given that peaking capacity has a much shorter lead time, it is highly likely that during the ten years from $s$ to $t = s + 10$ additional peaking capacity will be constructed which will be available in year $t$.\textsuperscript{46} If so then it was not the choice of base load capacity in period $s$ which constrained choices in $t$, $t + 1$, $t + 2$, etc., but rather the subsequent additions of peaking capacity. The second option has value only when no additional peaking capacity is added in the intervening ten years.\textsuperscript{47}

Here we model the impact of only the first option value, so that we
address only the question: how significantly does the choice of base load capacity in one period constrain base load choices in subsequent periods? Using the analytical results of chapter VI, section B and Appendix D, we approximated the option value foregone by constructing an additional unit of base load generating capacity using Monte Carlo simulations of future demand growth. We give here an intuitive description of how the option values were computed. A more rigorous exposition is contained in Appendix F.

The first step in finding optimal "durable capital" investments was to find the optimal "one period capital" investments. After doing this, 2500 random walks into the future were taken, starting from period s, the point at which base load investment decisions must be made for period t. An initial step forward was taken corresponding to a random drawing of \( \epsilon_{s+1} \), the lognormally distributed random variable affecting \( \eta_{s+1} \) and hence demand, \( D_{s+1} \). Given knowledge about \( D_{s+1} \), the question was asked: should additional base load capacity be constructed for period \( t+1=s+11 \) ? If the answer was yes then that random walk was discontinued: for that realization of \( \eta_{s+1} \) the previous choice of base load capacity in period s did not constrain any subsequent choices, and the option value foregone has a zero value. If the answer was no, then an approximation was taken of the value of the expected increment in total generating costs incurred because of the previous choice of generating capacity. This excess cost was discounted at 3 percent, averaged into the option value, and another step forward was taken to see if the previous choice of generating capacity had also constrained choices two periods into the future. If so the excess costs were again discounted and averaged into the option value and the random walk forward continued.
In order to simplify the decision as to whether any new capacity should be added, we restricted our analysis to the case where all capital, operating and shortage costs are constant in real terms. This keeps the optimal choice of each type of generating capacity proportional to $n^{49}$. The calculation of the excess costs of holding too much capacity was simplified by two approximations: the fact that prices can be adjusted in response to this excess capacity was ignored$^{50}$, and a linear approximation of the load distribution curve in the neighborhood of the optimum was used.$^{51}$ These latter two assumptions were sufficient to avoid having to regenerate the load distribution curve at each step of each random walk, an enormous savings in computation time.

Upon finishing all 2500 random walks, the total option value was treated as a fixed cost, and its derivative (calculated concurrently with the total option value) was used to find revised base load investments in the initial period, period $s$. The total option value was then recalculated based on these revised set of investments. Iterating between investments and option values conditional on those investments was repeated until convergence was achieved.$^{52}$

Table 8-9 presents the results of incorporating option values into the base case assumptions underlyng table 8-7 (i.e., a 10 year lead time for base load capacity, one year lead time for prices and peaking capacity investments, 3 percent median annual growth with a ninety percent confidence interval of 1.1 to 4.9 percent.) Our model suggests that the option value foregone by building additional base load generating capacity is of trivial importance for the base case assumptions which we have used. The total option value foregone is only $57$ thousand, which should be contrasted with total annual capital charges of $1.3$ billion. The change
Table 8-9
Sequential Decision Making with Durable Capital
Base Load Capacity Chosen With a Ten Year Lead Time
Peaking Capacity and Prices Chosen In the Last Period

<table>
<thead>
<tr>
<th>Marginal Cost Pricing</th>
<th>(1) low growth scenario</th>
<th>(2) median growth scenario</th>
<th>(3) high growth scenario</th>
<th>(4) average over all scenarios</th>
</tr>
</thead>
<tbody>
<tr>
<td>variables</td>
<td>units</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>price</td>
<td>cts/kwhr</td>
<td>4.45</td>
<td>4.72</td>
<td>4.92</td>
</tr>
<tr>
<td>demand</td>
<td>$10^6$ kwhrs</td>
<td>42.67</td>
<td>49.92</td>
<td>58.87</td>
</tr>
<tr>
<td>shortages</td>
<td>$10^6$ kwhrs</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>total surplus</td>
<td>$10^6$</td>
<td>21557</td>
<td>25974</td>
<td>31269</td>
</tr>
<tr>
<td>profits</td>
<td>$10^6$</td>
<td>-115.</td>
<td>7.6</td>
<td>117.</td>
</tr>
<tr>
<td>shortage costs</td>
<td>$10^6$</td>
<td>0.25</td>
<td>0.30</td>
<td>0.35</td>
</tr>
<tr>
<td>capital costs</td>
<td>$10^6$</td>
<td>1213</td>
<td>1264</td>
<td>1328</td>
</tr>
<tr>
<td>operating costs</td>
<td>$10^6$</td>
<td>803</td>
<td>1085</td>
<td>1454</td>
</tr>
<tr>
<td>base capacity</td>
<td>$10^6$ kw</td>
<td>5.93</td>
<td>5.93</td>
<td>5.93</td>
</tr>
<tr>
<td>peak capacity</td>
<td>$10^6$ kw</td>
<td>5.58</td>
<td>7.54</td>
<td>9.95</td>
</tr>
<tr>
<td>total capacity</td>
<td>$10^6$ kw</td>
<td>11.51</td>
<td>13.47</td>
<td>15.88</td>
</tr>
</tbody>
</table>

total option value = $57,000

option value foregone by adding one more kw of base load capacity = $0.87/kw

column (1) based on low growth scenario, the fifth percentile of demand

column (2) based on median growth scenario, fiftieth percentile of demand

column (3) based on high growth scenario, the ninety-fifth percentile of demand

column (4) shows expected values for all variables.
in the total option value from adding one more kilowatt of generating
capacity is only $.86 per kilowatt of capacity.

Given that these option values are small it is not surprising that
taking them into account has a negligible impact on prices, demand, and
generating capacity choices. Prices and demand differ between tables 8-7
and 8-9 by less than .1 percent. Optimal base load capacity has declined
from 5.96 to 5.93 MW. This .50 percent reduction, combined with the
previous .67 percent reduction found in section E means that taking into
account uncertainty properly instead of simply using the median demand
growth scenario reduces desired base load capacity by only 1.20 percent.
This discrepancy is small enough to be ignored for essentially all
practical applications.

In this section we have introduced the reality that generating
capacity is durable, and hence that building additional generating capacity
in one period constrains the choices which are available in subsequent
periods. For the base case assumptions which we have been using throughout,
the "irreversibility effect" was found to be empirically insignificant. We
examine the sensitivity of our results to the underlying parameter values
in the next section.

VIII.G. Sensitivity Analysis

In order to see whether or not these results are sensitive to the base
case parameter values we used, we resimulated our model for a variety of
cost and demand parameter assumptions. Table 8-10 displays the percent
change in base load capacity choices, expected prices, and total option
values respectively for a variety of deviations from our base case
assumptions. Specifically, the own price elasticity of demand, the load
Table 8-10

Sensitivity Analysis - Nonstochastic Demand and Cost Parameters

Percent Deviations From Base Case Using Frameworks II and X

<table>
<thead>
<tr>
<th>Framework =</th>
<th>II</th>
<th>X</th>
<th>II</th>
<th>X</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Base Load</td>
<td>Price Averaged</td>
<td>Total Option</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Capacity</td>
<td>Over 100 Scenarios</td>
<td>Value</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Elasticity of Demand:
(base case: -.5)
low: -.2  
0.0  -0.1  0.0  0.0  0.0
high: -.8 
0.0  +0.0  0.0  0.0  0.0

Load Factor
(base case: .65)
low: .60  
-12.3  -12.3  2.8  2.8  -36.8
high: .75 
20.3  20.4  -3.2  -3.2  124.6

Base Load Operating Cost
(base case: 1.0 cents/kwhr)
low: -20%  
8.5  8.4  -3.0  -3.0  22.8
high: +20% 
-9.2  -9.1  2.8  2.8  -36.8

Peaking Operating Cost
(base case: 4.4 cents/kwhr)
low: -15%  
28.0  31.2  -5.9  -5.9  -84.2
high: +20% 
22.7  22.8  4.9  5.1  94.7

Base Load Capital Charge
(base case: 3.16 cents/kwhr)
low: -20%  
36.7  36.8  -10.6  -10.8  -77.2
high: +20% 
31.2  -34.2  7.9  7.9  77.2

Peaking Capital Charge
(base case: .46 cents/kwhr)
low: -20%  
-3.5  -3.5  -1.7  -1.9  -7.0
high: +20% 
3.3  3.4  1.5  1.7  8.8

Shortage Cost
(base case: 200 cents/kwhr)
low: -50%  
0.0  0.0  0.0  0.0  0.0
high: +50% 
0.0  0.0  0.0  0.0  0.0

Note: Framework II is equivalent to using the 50th percentile realization of \( \eta \) and ignoring uncertainty, while for framework X the full simulation model was used, taking into account uncertainty, prices and the durability of capital and using 100 drawings of \( \eta \) from a lognormal distribution.
factor, operating, capital, and shortage costs were each varied; prices and investments reoptimized; and the effects on prices and investments noted. These deviations were simulated using two different modeling framework: framework II in which uncertainties but not prices are ignored, and framework X, taking into account demand uncertainty, the price feedback and the durability of capital. The magnitudes of these percentages indicates the sensitivity of base load capacity, prices and option values to the given parameters. The difference between the percentages for the two sets of simulations indicates whether or not taking into account demand uncertainty has a meaningful impact.

Several results are noteworthy in table 8-10. The own price elasticity of demand has virtually no impact on the option values and the optimal base load capacity choice. Since the elasticities were altered by rotating the demand curve around the price-quantity combination occurring in the median demand scenario, the prices are by construction constant for different price elasticities in the median scenario and virtually unchanged in the full simulations.

Increasing the system load factor has a substantial impact on desired base load capacity, prices and option values. As demand variability throughout the year decreases, more base load capacity and less peaking capacity can be constructed (the direct effect) and prices can be lowered to reflect the improved operating ratios (the pricing effect). Option values increase to reflect the fact that a given size deviation from the cost minimizing investments is more expensive. Taking into account uncertainty rather than using the median demand growth scenario still causes a negligible change in desired base load investments.

Operating and capital costs for both generating technologies were
adjusted upward and downward by 20 percent. The one exception was the peaking unit operating cost, which was adjusted downward only 15 instead of 20 percent (a 20 percent reduction made it optimal to invest in only peaking capacity). Prices are moderately sensitive to changes in cost parameters. A 20 percent increase in peaking unit operating costs raised expected short run marginal costs, and hence the expected price by 4.9 percent. A twenty percent increase in base load capital costs raised prices 7.9 percent. Base load operating costs and peaking unit capital costs have less of an impact on prices by virtue of being smaller relative to total costs.

Optimal base load investments are highly sensitive to cost parameter changes. A 20 percent increase in base load capital costs causes nearly a one third reduction in desired base load capacity. A 20 percent increase in peaking operating costs, on the other hand causes a 22.8 percent increase in desired base load investments. Option values also vary with cost parameters but remain small in magnitude. The 15 percent reduction in peaking operating costs almost doubles the total option value, but it remains only $111 thousand.

An examination of table 8-10 reveals that the error caused by ignoring uncertainty is somewhat sensitive to cost parameter changes which significantly decrease the attractiveness of base load capacity. A 15 percent decrease in peaking capacity operating costs and a 20 percent increase in base load capital costs cause differences of 5.6 and 4.4 percent respectively, between the capacity choices which are chosen using the median scenario and those chosen using the tenth modeling framework, and hence incorporating the impact of uncertainties. The magnitudes are still small relative to total costs however.
The bottom two rows of table 8-10 consider fairly dramatic changes in the assumed shortage costs: shortage costs were both halved and doubled. Surprisingly, optimal prices, base load capacity and option values are virtually unaffected. Although not shown in the tables, peaking capacity changed by approximately 5 percent in each case. Doubling the shortage cost nearly halves the optimal loss of load probability. As a result the doubled shortage cost receives only about half of the weight it previously received in the determination of the optimal price. The socially optimal price changes insignificantly, and hence desired base load capacity and option values also remain virtually unchanged.

A different set of simulations was run in order to examine the sensitivity of our results to the rate of demand growth and the degrees of uncertainty. Five different growth rates and six different degrees of uncertainty around each of these growth rates were considered. Table 8-11 presents three variables from each of these thirty simulations: the total option values, the marginal change in the option value resulting from a one kw increase in base load capacity, and the percent difference between the base load capacity which is optimal for the median demand scenario, and the base load capacity choice which is chosen optimally based on the degree of uncertainty which is present. (What we have used for our base case is not shown in the table since for table 8-11 we scaled the degree of uncertainty to have increments of ±.5 percent in the ten year 90 percent confidence interval for the average annual growth rate, and our base case corresponds to a 90 percent confidence interval for average annual demand growth of ±1.9. percent.)

The simulations are revealing. Total option values increase significantly as the median growth rate declines and the degree of
Table 8-11

Sensitivity Analysis
Demand Growth and Demand Uncertainty

Total Option Value - Millions of Dollars

<table>
<thead>
<tr>
<th>Growth Rate</th>
<th>Confidence Intervals for Demand Growth Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>±.5%</td>
</tr>
<tr>
<td>0</td>
<td>.95</td>
</tr>
<tr>
<td>1</td>
<td>.00</td>
</tr>
<tr>
<td>2</td>
<td>0.</td>
</tr>
<tr>
<td>3</td>
<td>0.</td>
</tr>
<tr>
<td>4</td>
<td>0.</td>
</tr>
</tbody>
</table>

Change in Option Value per kw of Additional Base Load Capacity

<table>
<thead>
<tr>
<th>Growth Rate</th>
<th>Confidence Intervals for Demand Growth Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>±.5%</td>
</tr>
<tr>
<td>0</td>
<td>$17.06</td>
</tr>
<tr>
<td>1</td>
<td>.13</td>
</tr>
<tr>
<td>2</td>
<td>0.</td>
</tr>
<tr>
<td>3</td>
<td>0.</td>
</tr>
<tr>
<td>4</td>
<td>0.</td>
</tr>
</tbody>
</table>

Percent Change in Base Load Capacity From "Most Likely" Scenario

<table>
<thead>
<tr>
<th>Growth Rate</th>
<th>Confidence Intervals for Demand Growth Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>±.5%</td>
</tr>
<tr>
<td>0</td>
<td>-10.2</td>
</tr>
<tr>
<td>1</td>
<td>-.2</td>
</tr>
<tr>
<td>2</td>
<td>-.1</td>
</tr>
<tr>
<td>3</td>
<td>-.1</td>
</tr>
<tr>
<td>4</td>
<td>-.1</td>
</tr>
</tbody>
</table>

Note: Confidence intervals are only approximate: since they are based on the lognormal distribution they are not actually symmetric.
uncertainty increases. With a 1 percent median growth rate and a confidence interval of (approximately) ±2 percentage points, the total option foregone when building the optimal amount of base load capacity is $2.7 million. The change in the value of this option is $17.56 per kw of additional capacity. Option values are large for higher expected growth rates (e.g. 4 percent per year) only if there is very substantial uncertainty surrounding this expected growth rate (e.g. ±3 percent per year). Slow and highly uncertain demand growth also affect desired base load capacity in the predicted manner. Optimal base load capacity when the median demand growth is one percent with a 90 percent confidence interval of ±2 percentage points is 11.3 percent below the level which would be chosen using the median scenario. Other more extreme cases result in even larger discrepancies. Such differences are large enough to be empirically relevant for investment policy decisions, although the exaggerated parameter values which are needed to get such large differences should be recognized.

At this point we would like to highlight two simplifying features of our model which undoubtedly influence our results. The first is that we are using a reduced form forecasting equation which is conditional on independently and identically lognormally distributed random variables, \( \eta_t \). The only new information arriving in each period is a realization of \( \eta_t \), and this information is used merely to update the forecast of future levels of demand. The information is not used to update estimates of any structural parameters such as the growth rate or its variance.\(^{58}\) This is a particularly simple information structure. We suspect that the existence of serial correlation plus the updating of estimated structural parameters would tend to increase the value of information gained by deferring any decisions. The dependence of demand in any given period on
previous period prices (our model assumes they are independent) would further increase serial correlation of demand over time, and improve the value of the information gained by deferring investment decisions. We speculate that our model understates the true option values foregone when additional generating capacity is constructed. Whether or not this speculation is correct can be verified only by using a more realistic structural model of demand, an undertaking beyond the scope of this paper.

A second and equally important assumption which we have used is that only two generating technologies are available: base load and peaking (gas turbine) capacity. The absence of any intermediate load generating capacity means that we have exaggerated the fraction of the year that base load capacity can optimally be used. This is compounded by our treatment of forced outages by derating capacity. Our operating and capital cost assumptions imply that once base load capital cost is inflated to reflect its unavailability, it is desirable to build base load capacity as long as it will be used at least 51.6 percent of the year. This may be reasonable if the only alternative is extremely fuel intensive gas turbine capacity, however the existence of at least one intermediate form of generating capacity would raise the fraction of the year that base load capacity must be used in order to be economic. This fact is illustrated in figure 8-2, where the "screening curve" for any economic intermediate load technology must increase the fraction of the year that base load capacity must be operated to be economic. This is relevant because it is in the tails of the probability distributions that our load distribution curve differs most from the load duration curve conditional on the expected value of \( \eta \). The presence of at least one intermediate form of generating capacity would move the optimal choice of base load capacity further out.
Figure 8-2

Effect of Intermediate Capacity on Optimal Base Load Use

-cost per kw per year-

Shortages
Base Load
Intermediate Load
Peaking

Load Distribution Curve
Load Duration Curve given \( E_\eta(\eta) \)

level of demand

0 fractions of a year
0' 1

x

1 fractions of a year

1
into the tail of the load distribution curve, and hence make base load capacity choices more sensitive to the degree of uncertainty. This suggests that the sensitivity of prices and base load capacity choices to uncertainty is probably understated by our simulation model.

In this section we examined the sensitivity of our results to the underlying parameter values using modeling framework X. The optimal price and base load capacity were found to be fairly sensitive to the assumed load factor, operating costs and capital costs, but insensitive to the price elasticity of demand and shortage cost. Uncertainty was seen to have a significant impact on base load choices only when demand growth is slow and highly uncertain. It was speculated that these results may depend on the simplified way in which uncertainty is resolved over time in our model, and the fact that only two generating technologies are presumed available.

VIII.H. Random Operating Costs and Average Cost Pricing

In this section two of the extensions of chapter VII are added to the model: operating cost uncertainty and average cost pricing. We return to a one period capital framework where both prices and peaking capacity must be chosen in the last period, i.e. framework V in table 8-1. Base load capacity must still be chosen ten years in advance.

Fuel cost uncertainty changes the first order conditions for optimal capacity choices, however the changes are readily handled by introducing new covariance terms as discussed in chapter VII, section A. We consider here the case where only peak operating costs are uncertain, and we continue to examine 100 possible future scenarios. Each scenario is now fully described by two random variables: a lognormally distributed random variable, \( \eta \), affecting demand multiplicatively, and a lognormally
distributed random variable, \( \eta^2 \), affecting peak operating costs.

Specifically, we model the uncertain variables by the following relations:

\[
D = a p^Y \eta^1 \quad \text{and} \quad m_2 = b \eta^2
\]

where \( \log(\eta^1) \) normally \((0, \sigma_1^2)

\( \log(\eta^2) \) normally \((-\sigma_2^2, \sigma_2^2)\)

and \( \rho \) = correlation between \( \log(\eta^1) \) and \( \log(\eta^2) \)

The mean of \( \log(\eta^2) \) was chosen so as to maintain the same mean of \( m_2 \) as uncertainty about \( m_2 \) increases. As pointed out in chapter VII both the degree of fuel cost variability (parametrized via \( \sigma_2^2 \)) and its correlation with demand variations (\( \rho \)) are relevant for capacity choices. We consider here correlations between \( \log(\eta^1) \) and \( \log(\eta^2) \) ranging from \(-1\) to 1.

In table 8-12 the desired choices of base load and peaking capacity are shown for several different levels of peak operating cost uncertainty and correlations between fuel and demand uncertainties. Although the magnitudes are small, the two theoretical results derived in section VII.A are verified. First, increasing uncertainty about peak operating costs which is uncorrelated with the uncertain random variables affecting demand increases desired reliance on the uncertain fuel cost technology. This occurs by virtue of the price induced negative correlation between fuel costs and consumption: high fuel costs increase prices and hence lower consumption. The second theoretical result is that holding the degree of fuel cost uncertainty constant, desired reliance on the variable operating cost technology decreases as the correlation between operating costs and level of demand increases. An inverse relationship between the two is the
Table 8-12
Effects of Fuel Cost Uncertainty

Increasing Peak Operating Cost Uncertainty

<table>
<thead>
<tr>
<th></th>
<th>90 percent confidence intervals for $m_2(\eta^2)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$(4.40, 4.40)$ $ (3.60, 5.20)$ $ (2.92, 5.88)$ $ (2.33, 6.87)$</td>
</tr>
<tr>
<td>Base Load Capacity $(10^5$ kw)</td>
<td>5.96</td>
</tr>
<tr>
<td>Peaking Capacity $(10^6$ kw)</td>
<td>7.54</td>
</tr>
</tbody>
</table>

Increasing the Correlation Between Demand and Fuel Costs

<table>
<thead>
<tr>
<th></th>
<th>Correlation between $\log(\eta^1)$ and $\log(\eta^2)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-1.0</td>
</tr>
<tr>
<td>Base Load Capacity $(10^5$ kw)</td>
<td>6.05</td>
</tr>
<tr>
<td>Peaking Capacity $(10^6$ kw)</td>
<td>7.48</td>
</tr>
</tbody>
</table>
most favorable for peak capacity, while a positive relationship between the
two is the least favorable. In the example shown, the positive
relationship is sufficiently strong that one should purchase less peak
capacity than is suggested by using expected operating costs.

We also accommodated the average cost pricing results of section VII.C
in our model. Three modifications of the simulation model were needed.
First, instead of marginal cost pricing, the price conditional on \( K \) and \( \eta \)
was found by solving \(^{60}\)

\[
P = \frac{C(K, P, \eta)}{D(P, \eta)} \quad (8-7)
\]

where \( C(K, P, \eta) \) = total fixed plus operating costs, excluding shortages
\( D(P, \eta) \) = total served demand

Second, the externality of shortage costs was dealt with in the last period
(once \( \eta \) is known) by adjusting the shortage cost upward or downward as
shown by equation (7-36). Finally, each scenario was weighted by the
weights \( \omega(\eta) \) derived in chapter VII, section D, when minimizing costs.
We thus did not minimize total expected costs, as is desirable under
marginal cost pricing and risk neutrality, but rather we minimized a
weighted average of costs, where the weights take into account how far
prices deviate from marginal costs.

Table 8-13 displays optimal prices and capacity choices given that
prices are set so as to exactly cover expected costs. The fifth, fiftieth,
and ninety-fifth percentile scenarios for \( \eta \) are each shown, together with
the (unweighted) expected value of each variable over all 100 scenarios.
Note that the price (which equals average costs) is slightly higher for the
fifth and ninety-fifth percentile realizations of \( \eta \) than for the fiftieth
percentile (median) realization. The short run average cost curve is
### Table 8-13

**Average Cost Pricing**

**Sequential Decision Making**

**Base Load Capacity Chosen With a Ten Year Lead Time**

**Peaking Capacity and Prices Chosen In the Last Period**

<table>
<thead>
<tr>
<th><strong>variables</strong></th>
<th><strong>units</strong></th>
<th><strong>(1)</strong></th>
<th><strong>(2)</strong></th>
<th><strong>(3)</strong></th>
<th><strong>(4)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>price</strong></td>
<td>cts/kwhr</td>
<td>4.74</td>
<td>4.71</td>
<td>4.73</td>
<td>4.71</td>
</tr>
<tr>
<td><strong>demand</strong></td>
<td>(10^9)kwhrs</td>
<td>41.40</td>
<td>50.00</td>
<td>60.07</td>
<td>50.13</td>
</tr>
<tr>
<td><strong>shortages</strong></td>
<td>(10^9)kwhrs</td>
<td>.0001</td>
<td>.0001</td>
<td>.0002</td>
<td>.0001</td>
</tr>
<tr>
<td><strong>total surplus</strong></td>
<td>(10^6)</td>
<td>21554</td>
<td>25933</td>
<td>31201</td>
<td>25980</td>
</tr>
<tr>
<td><strong>profits</strong></td>
<td>(10^6)</td>
<td>0.</td>
<td>0.</td>
<td>0.</td>
<td>0.</td>
</tr>
<tr>
<td><strong>shortage costs</strong></td>
<td>(10^6)</td>
<td>.27</td>
<td>.30</td>
<td>.34</td>
<td>.30</td>
</tr>
<tr>
<td><strong>capital costs</strong></td>
<td>(10^6)</td>
<td>1205</td>
<td>1266</td>
<td>1337</td>
<td>1266</td>
</tr>
<tr>
<td><strong>operating costs</strong></td>
<td>(10^6)</td>
<td>755.</td>
<td>1088</td>
<td>1504</td>
<td>1093</td>
</tr>
<tr>
<td><strong>base capacity</strong></td>
<td>(10^6) kw</td>
<td>5.94</td>
<td>5.94</td>
<td>5.94</td>
<td>5.94</td>
</tr>
<tr>
<td><strong>peak capacity</strong></td>
<td>(10^6) kw</td>
<td>5.23</td>
<td>7.55</td>
<td>10.27</td>
<td>7.56</td>
</tr>
<tr>
<td><strong>total capacity</strong></td>
<td>(10^6) kw</td>
<td>11.17</td>
<td>13.39</td>
<td>16.21</td>
<td>13.50</td>
</tr>
<tr>
<td><strong>weights</strong> (w(\eta))</td>
<td></td>
<td>1.038</td>
<td>.998</td>
<td>.977</td>
<td>1.002</td>
</tr>
</tbody>
</table>

- column (1) based on low growth scenario, the fifth percentile of demand
- column (2) based on median growth scenario, fiftieth percentile of demand
- column (3) based on high growth scenario, the ninety-fifth percentile of demand
- column (4) shows expected values for all variables.
U-shaped, as expected. Also note that prices display very little
variability under average cost pricing. This is the justification behind
our using fixed price confidence intervals for scaling the degree of demand
uncertainty, as discussed in section VIII.D above. This relative stability
of prices is reflected in higher consumption variability. Whereas under
marginal cost pricing the ninety percent confidence interval for
consumption was 42.71 to 58.90 billion kw-hrs (see table 8-7) under average
cost pricing it is 41.40 to 60.07 billion kw-hrs. This supports our
theoretical result that average cost pricing exacerbates consumption
uncertainty.

The welfare loss from adopting average cost instead of marginal cost
pricing is on the order of $1 million per year. This is the difference
between expected consumer plus producer surplus in tables 8-7 and 8-13.
(Although not shown, we also simulated the model without weighting costs,
and the difference in expected consumer surplus was identical to 7 digits.)
This small a difference is not particularly surprising given that the
demand for electricity is fairly inelastic (we are using an elasticity of
-.5), and by assuming that demand is isoelastic throughout the year we know
from our theoretical model that marginal costs are very close to average
costs. (see section IV.B) As we show below, the losses can be greater when
capacity is not chosen optimally.

Although the magnitudes once again are miniscule, a comparison of
investments under marginal and average cost pricing using tables 8-7 and
8-13 reveals that with average cost pricing less base load capacity
(5.94 versus 5.96) should be constructed than under marginal cost
pricing. For low realizations of demand less peaking capacity is
desirable under average cost pricing than under marginal cost pricing, with
the reverse pattern holding for high realizations of demand where marginal costs are above average costs.

The bottom row of table 8-13 shows the weights that were calculated as the optimal weights to be attached to different scenarios when minimizing expected costs. The fact that the weights are so near 1 (the ninety percent confidence interval is 1.038 to .977) reflects the fact mentioned above that when capacity is chosen optimally, average cost prices do not differ substantially from marginal cost prices. This tentatively suggests that ignoring the "second best" corrections outlined in chapter VII may be reasonable for most practical applications.

We also ran a few simulations incorporating operating cost uncertainty both with and without demand uncertainty under a regime of average cost pricing. The results of these simulations are not presented here. The only observation which we make is that with both operating cost and demand uncertainty, scenarios in which average cost prices deviate significantly from marginal cost prices are more frequent, and hence the weights which we introduced take on greater importance.

In this section we have introduced two further complications into our simulation model: random operating costs and average cost pricing. Although these complications were found to change prices, consumption and desired investments in the theoretically predicted directions, the magnitudes of the changes were so small as to be inconsequential. The suggestions made in the previous section as to why our simulation model understates the impact of uncertainties continue to be valid for the two extensions considered here.

VIII.I. "One Third Less Base Load Capacity"
How much generating capacity should be built for 1990? What will be the impact of building less generating capacity? Our simulation model suggests that it is important to consider pricing effects when answering these questions, since prices are sensitive to investment decisions, the realization of uncertain variables, and certain cost and demand parameters. Whether or not it is critical to consider long term demand uncertainty depends on the underlying cost and demand parameters, however our model tentatively suggests that ignoring uncertainty and focusing only on the median scenario may not significantly distort capacity choices for most practical purposes.

In section C we demonstrated that the impact of building one third less than the optimal level of base load capacity depended critically on the way in which prices are set. Our subsequent analysis of uncertainty and the effect of sequential decision making suggests that the ability to build additional peaking capacity in order to make up for any shortfall in base load capacity may also be important. Furthermore, although average and marginal cost prices may be similar when capacity is chosen optimally, when capacity is not chosen optimally the two may diverge significantly. We examine these possibilities here.

The impact of building too little base load generating capacity can be analyzed as a constrained optimization problem. In section C we assumed that the only choice variable available was the price of electricity: base load and peaking capacity were treated as fixed. The appropriate action when not enough base and peaking capacity is constructed is to raise prices to reflect higher short run marginal costs. If we impose yet another constraint on the optimization problem, namely that price be set equal to average costs, then there are no choice variables left. We found
previously that average cost pricing resulted in the perverse behavior that prices were lower when too little base load capacity is constructed. When it is recognized that certain types of generating capacity have shorter lead times than others then the ability to rely on these short lead time technologies to meet demand introduces a new choice variable into the optimization problem. In our model the only such short lead time technology is peaking capacity.

Table 8-14 summarizes how prices and peaking capacity should optimally be chosen given the constraint that only 3.97 MW of base load capacity is available, two thirds of the socially optimal level. The results here were derived using the fifth modeling framework, and hence are directly comparable to table 8-7.62

Several contrasts between tables 8-14 and 8-7 are of interest. First, prices are 3 to 11 percent higher when too little base load capacity is constructed. This should be contrasted with the 36.8 percent price increase which was optimal when peaking capacity could not be reoptimized. Second, considerable peaking capacity should be built in order to compensate for the shortfall in base load capacity. Peaking capacity averages 19.9 percent higher than in table 8-7 where there is no base load shortage. Third, the loss of consumer plus producer surplus has been further mitigated by the appropriate choices of prices and peaking capacity. In terms of expected consumer plus producer surplus, welfare remains $36 million lower than that attainable with the optimal choice of base load capacity, however this is far removed from the $2.4 billion burden when neither prices nor peaking capacity is adjusted. The losses of welfare are concentrated heavily in high demand scenarios: a comparison of tables 8-7 and 8-14 reveals that welfare is only $1 million lower in the
Table 8-14

One Third Less Nuclear Capacity  
Sequential Decision Making  
Base Load Capacity Chosen With a Ten Year Lead Time  
Peaking Capacity and prices Chosen In the Last Period

<table>
<thead>
<tr>
<th>Marginal Cost Pricing</th>
<th>(1) low growth scenario</th>
<th>(2) median growth scenario</th>
<th>(3) high growth scenario</th>
<th>(4) average over all scenarios</th>
</tr>
</thead>
<tbody>
<tr>
<td>variables</td>
<td>units</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>price</td>
<td>cts/kwhr</td>
<td>4.95</td>
<td>5.09</td>
<td>5.11</td>
</tr>
<tr>
<td>demand</td>
<td>$10^9$kwhrs</td>
<td>40.48</td>
<td>48.06</td>
<td>57.81</td>
</tr>
<tr>
<td>shortages</td>
<td>$10^9$whrs</td>
<td>.0001</td>
<td>.0001</td>
<td>.0002</td>
</tr>
<tr>
<td>total surplus</td>
<td>$10^6$</td>
<td>21556</td>
<td>25936</td>
<td>31204</td>
</tr>
<tr>
<td>profits</td>
<td>$10^6$</td>
<td>88.97</td>
<td>152.</td>
<td>159.</td>
</tr>
<tr>
<td>shortage costs</td>
<td>$10^6$</td>
<td>.24</td>
<td>.29</td>
<td>.34</td>
</tr>
<tr>
<td>capital costs</td>
<td>$10^6$</td>
<td>897.</td>
<td>951.</td>
<td>1019</td>
</tr>
<tr>
<td>operating costs</td>
<td>$10^6$</td>
<td>1018</td>
<td>1346</td>
<td>1774</td>
</tr>
<tr>
<td>base capacity</td>
<td>$10^6$ kw</td>
<td>3.97</td>
<td>3.97</td>
<td>3.97</td>
</tr>
<tr>
<td>peak capacity</td>
<td>$10^6$ kw</td>
<td>6.95</td>
<td>8.99</td>
<td>11.62</td>
</tr>
<tr>
<td>total capacity</td>
<td>$10^6$ kw</td>
<td>10.92</td>
<td>12.96</td>
<td>15.59</td>
</tr>
</tbody>
</table>

column (1) based on low growth scenario, the fifth percentile of demand  
column (2) based on median growth scenario, fiftieth percentile of demand  
column (3) based on high growth scenario, the ninety-fifth percentile of demand  
column (4) shows expected values for all variables.
low growth scenario but $66 million worse in the high growth scenario. Finally, despite the increased expenditures on peaking capacity, profits are positive at the socially optimal prices.

As noted in section C, it is unrealistic to suppose that regulators will allow prices with significantly positive profits. Table 8-15 presents simulation results analogous to those of table 8-14, only assuming prices are set only so as to recoup operating costs plus capital charges. They are directly comparable to the unconstrained average cost pricing simulation depicted in table 8-13. There is only one unconstrained choice variable, namely the choice of peaking capacity.

From a comparison of tables 8-14 and 8-15 it can be seen that average cost prices are too low on average, but only by about 5 percent. Peaking capacity must be increased somewhat to meet the resulting slight increases in demand. Expected welfare is $2 million lower than when prices are set equal to short run marginal costs. Profits are, by construction, zero. All in all average cost pricing is still not substantially different from marginal cost pricing, and is not nearly as perverse as when capacity choices are all constrained.

The important conclusion to be drawn from this discussion is that the ability to set prices and choose peaking capacity so as to compensate for any shortfall in base load generating capacity can substantially mitigate the potentially drastic impact of such a shortfall. Even if prices can not be chosen optimally because of average cost pricing constraints the ability to rely on some short lead time technology at the last minute results in a substantial improvement.
Table 8-15

**One Third Less Nuclear Capacity**
Sequential Decision Making
Base Load Capacity Chosen With a Ten Year Lead Time
Peaking Capacity and prices Chosen In the Last Period

<table>
<thead>
<tr>
<th>variables</th>
<th>units</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Cost</td>
<td></td>
<td>low</td>
<td>median</td>
<td>high</td>
<td>average</td>
</tr>
<tr>
<td>Pricing</td>
<td></td>
<td>growth</td>
<td>growth</td>
<td>growth</td>
<td>over all</td>
</tr>
<tr>
<td></td>
<td></td>
<td>scenario</td>
<td>scenario</td>
<td>scenario</td>
<td>scenarios</td>
</tr>
<tr>
<td>price</td>
<td>cts/kwhr</td>
<td>4.74</td>
<td>4.79</td>
<td>4.84</td>
<td>4.79</td>
</tr>
<tr>
<td>demand</td>
<td>$10^9$kwhrs</td>
<td>41.39</td>
<td>49.58</td>
<td>59.38</td>
<td>49.66</td>
</tr>
<tr>
<td>shortages</td>
<td>$10^9$whrs</td>
<td>0.0001</td>
<td>0.001</td>
<td>0.002</td>
<td>0.0001</td>
</tr>
<tr>
<td>total surplus</td>
<td>$10^6$</td>
<td>21554</td>
<td>25933</td>
<td>31201</td>
<td>25980</td>
</tr>
<tr>
<td>profits</td>
<td>$10^6$</td>
<td>0.</td>
<td>0.</td>
<td>0.</td>
<td>0.</td>
</tr>
<tr>
<td>shortage costs</td>
<td>$10^6$</td>
<td>.23</td>
<td>.28</td>
<td>.33</td>
<td>.28</td>
</tr>
<tr>
<td>capital costs</td>
<td>$10^6$</td>
<td>903</td>
<td>961</td>
<td>1030</td>
<td>961</td>
</tr>
<tr>
<td>operating costs</td>
<td>$10^6$</td>
<td>1057</td>
<td>1413</td>
<td>1844</td>
<td>1417</td>
</tr>
<tr>
<td>base capacity</td>
<td>$10^6$ kw</td>
<td>3.97</td>
<td>3.97</td>
<td>3.97</td>
<td>3.97</td>
</tr>
<tr>
<td>peak capacity</td>
<td>$10^6$ kw</td>
<td>7.19</td>
<td>9.40</td>
<td>12.05</td>
<td>9.43</td>
</tr>
<tr>
<td>total capacity</td>
<td>$10^6$ kw</td>
<td>11.16</td>
<td>13.37</td>
<td>16.02</td>
<td>13.40</td>
</tr>
</tbody>
</table>

column (1) based on low growth scenario, the fifth percentile of demand

column (2) based on median growth scenario, fiftieth percentile of demand

column (3) based on high growth scenario, the ninety-fifth percentile of demand

column (4) shows expected values for all variables.
VIII.J. Summary of Simulation Results

In this chapter we have developed and used a simple simulation model to analyze optimal pricing and investment decisions by electric utilities in the presence of uncertainty. Our model differs from many previous models in that 1) we incorporate a "closed loop" pricing procedure which explicitly considers the feedback effect of investment decisions on prices, 2) we explicitly incorporate both demand and operating cost uncertainty in our objective function, 3) we model the sequential nature in which decisions are made under uncertainty, and 4) we consider the fact that capital investments are in fact durable and hence irreversible.

Perhaps the single most important finding is that prices can be significantly affected by investments, the realization of uncertain variables, the load factor and the underlying cost parameters. Modeling frameworks which attempt to treat prices as exogenous may seriously misrepresent future scenarios.

As in the theoretical chapters above, we distinguished three different effects of uncertainty on desired investments. The direct effect, for which prices are held fixed, unambiguously reduces desired base load capacity and increases desired peaking capacity. The price effect occurs because these changes in capacity choices affect optimal prices. In our model the price effect reduces desired base load capacity, although in general the pricing effect may be to either increase or decrease desired base load capacity. The third effect, the irreversibility effect, further reduces desired base load capacity by taking into account the fact that capacity is durable, and hence less is desired than is suggested by static investment rules.

While all three effects of uncertainty on optimal capacity choices
were distinguished, their cumulative impact was found to be of negligible practical importance for our base case assumptions. Sensitivity analysis found this result to be fairly robust to a broad range of parameter changes, although cost parameter changes which substantially decreased optimal base load capacity augmented the importance of uncertainty, and for slow and highly uncertain demand growth simulations uncertainty was found to have an impact of 5 percent and more on desired base load capacity. The preliminary indication of our simulations is that using the median demand scenario instead of our probability weighted average of many possible scenarios may not seriously distort capacity choices. We speculate that this result is due to our simplified information structure, our linear load duration curve, our treatment of forced outages by simply derating capacity, and the fact that there are no intermediate load technologies and as a result, base load capacity is being overutilized. We have not run simulations to test these speculations, however.

Two extensions of our basic simulation model were also examined: random operating costs and average cost pricing. We demonstrated that it may be desirable to increase reliance on the more uncertain operating cost technology if its cost is negatively correlated with the level of demand. Average cost pricing was seen to exacerbate consumption uncertainty and alter capacity choices in the theoretically predicted direction. The magnitudes remain inconsequential, however.

Finally, our model suggests that the impact of building too little base load generating capacity can be substantially mitigated through the appropriate choices of prices and peaking capacity. Even when prices are constrained to recover only operating and capital costs, the ability to rely on peaking capacity at the last minute has a substantially beneficial effect.
Footnotes to Chapter VIII

1. Lee, Stoughton and Badertscher (1978) state: "At present, generation planning is an open loop process. The effect of the cost of electricity on the load forecast is not considered."

2. Long run marginal costs may depend on the realization of $\eta$ and hence be stochastic, yet they are still exogenous in the sense that we indicate.

3. Joskow (1976) develops a behavioral model of rate setting which emphasizes threshold behavior in the decision to apply for rate increases. Such behavior will be difficult to model at any level except that of an indivisible utility. Baughman, Joskow and Kamat (1979) model regulatory rate setting in their regionalized electricity model. Their regulatory submodel is described more fully in Kamat (1975). Werth (1978) examined the impact of alternative regulatory regimes on the cost of capital without finding strong evidence of any major impact.

4. Baughman et al state on page 21: "While the model incorporates a set of econometric demand equations to generate actual electric energy demand given a vector of prices of all basic energy inputs (coal, oil, natural gas, and the endogenous electricity price) it does not assume that the electric utilities employ such a sophisticated analysis of the own-price and cross-price elasticities to project demand. Rather their projections of demand are specified by the same technique as the supply decision variables. As a result of this approach actual electricity consumption will generally be different from projected consumption."

5. See Electric Council of New England (1979); and Motlagh (1976,p. 18)

6. See Baughman et al (1979) and Gordon (1975) for models employing sensitivity analysis. Motlagh (1976) uses a model which incorporates lead time uncertainty but does not consider demand uncertainty per se. He notes (p. 21) "Interval estimates are widely accepted in statistical forecasting and have been shown to provide a greater amount of information about the uncertainty of the variables being forecast than do point estimates. It is interesting to note, however, that public utilities rely mainly on these latter in their forecasting activities. Even in those cases where more than one point estimate is provided (such as a high, mean, and low forecast), no probability statements are attached to the intervals created by the various point estimates."
7. Investment models often use "levelized" capital costs, a discounted average of future capital charges, as a means of incorporating the durability of capital. While this is one way of calculating the rental cost of capital, it does not actually take into account the durability and irreversibility of capital.

8. Specifically, the value of the new information learned in each period depends critically on the degree of serial correlation of the random variables over time. This issue is discussed further in section F. below.

9. Our model is capable of handling any number of different types of generating capacity, with computational burdens increasing only modestly. We present a simple model here for simplicity. Originally we intended to contrast an investment sequence in which peaking capacity was chosen with lead times of two to five years, however the difference between peaking capacity choices made with 10 and 1 year lead times is sufficiently small that it appears that no further insights would be gained by these additional simulations.

10. This method of treating unit outages avoids the necessity of using a convolution procedure for incorporating probabilistic unit outages. See EGEAS (1980) page 5-16 for a discussion of this technique, as well as further refinements.

11. This constant cost assumption results in a stationary growth path in which the capacity of each type is growing at the same rate as demand. In fact the less restrictive assumption that all costs are growing at a constant rate can also be used with similar implications. We do not discuss this alternative further here.

12. See e.g. Brown and Johnson (1969), Meyer (1975), and EGEAS (1978)

13. Letting $K$ be total capacity $(k_1 + k_2)$ and using the notation derived in the main text the constraint that the loss of load probability not be greater than some level $\theta$ can be written as

$$\text{prob}(x > K) = 1 - G(K|P) < \theta$$

By choosing $\alpha$ such that $\theta = c_2/(\alpha - m_2)$ and assuming that the constraint is satisfied exactly the constraint can be written as

$$[1 - G(K|P)] = c_2/(\alpha - m_2)$$

or

$$c_2 + m_2 [1 - G(K|P)] = \alpha [1 - G(K|P)]$$

This equation is one of the equations to be solved to find the constrained maximum of the objective function. Since it is also one of the first order conditions derived by using our linear shortage cost function (see equation (4)) it follows that there is a linear loss function equivalent to any loss of load probability.
constraint. In the constrained optimization, \( \alpha \) appears as the lagrange multiplier on the loss of load probability constraint.


15. Transforming the expression in footnote 13 we can write

\[
\text{shortage cost} = \frac{c_2}{6} + m_2 = (\frac{.46}{(1/8760*10)}) + 4.40
\]

\[
= $16.83
\]

16. If brownouts are ignored, a loss of load probability of .88 days a year implies only that some customers will face shortages. If the expected size of these shortages is taken into account, shortages are allocated randomly among all users, and our linear load duration curve is used, this loss of load probability implies that an individual user can expect to go without electricity because of capacity induced shortages for less than one hour every 39 years (shortages = .000149, expected demand 50.00, fraction of demand urmet = 2.9x10e-6. If demand were consumed uniformly throughout the year 1/(8760x2.9x10e-6) = 39.34. Since demand is higher when shortages are most likely, the expected duration will be less.)

17. Baughman et al (1979, p. 75) cite six electricity demand studies suggesting long run own price elasticities of demand are in the range -.9 to -1.5, while their own and other short run elasticities are in the range -.1 to -.4

18. See, for instance, EGEAS (1980), page 5-17, and Hicks and Lee (1976).

19. Changes in the shape of the load duration curve would most reasonably be modeled by incorporating multiple prices and more sophisticated pricing effects. Recent work on peak load pricing and spot pricing addresses this issue. See Dubin and McFadden (1980), Bohn (1980) and Veall (1981) for three different approaches to this issue.

20. In recent years electricity demand has been growing more slowly than this in Massachusetts, however we use a 3 percent growth rate in order to reflect the growth which is typical for much of the U.S.

21. Since we use an isoelastic demand curve with an elasticity less than one, consumer surplus is infinite for any finite price. Since we are interested only in changes in welfare, we truncate the demand curve at a price of $2.00 for our calculations and look at changes in areas under the demand curve below this price.

22. See especially table 1 on page 5 of Lee et al, and the quote cited in footnote 1 above.
23. Sensitivity analysis does not synthesize uncertainty into the objective function in any systematic manner, while our approach does so explicitly.

24. Since we are using a lognormal distribution, the mode of \( \eta \) (the "most likely" realization) is below the median, which is in turn below the mean of \( \eta \).

25. As previously noted, our constant shortage cost per kw/hr of unmet demand is exactly equivalent to a constant loss of load probability.

26. An alternate future which could be examined would be one in which one third of the base load capacity is completely or partially built but cannot be used. This would of course be much more expensive from a social point of view.

27. In addition to the normative interest in marginal cost pricing, two part tariffs make such pricing feasible for electric utilities even when strict marginal cost pricing would yield profits which are either too large or too small (negative). Current regulatory imperfections make the efficient use of such prices very difficult in practice.

28. Since there is no uncertainty in the model at this point the only distortion from average cost pricing arises from the fact that shortage costs are an externality to the firm, and are not included in the costs used for calculating prices. These shortage costs are small relative to total costs; taking them into account properly when choosing peaking capacity increased desired peaking capacity by only .6 percent, and raised prices only in the fifth decimal place. A slightly stronger result is obtained under uncertainty, which is considered below in section VIII.H.

29. See Appendix C and the more intuitive discussion in chapter IV, section C.

30. Another alternative is to purchase electricity from neighboring utilities. We do not consider such possibilities here since it would substantially complicate the analysis. Lee, Stoughton, and Badertscher (1978) note: "No generation planning methodology is available to address the regional planning problem with a multi-dimensional objective function, as would be the case with multiple decision makers with different objectives." p. 3.

31. Lengthy lead times result in high interest costs which we have subsumed in our fixed capital cost. Uncertain lead times have an impact similar to that of uncertain demand, although we do not examine them here. We focus here on the implications of the fact that different types of generating capacity have differing lead times.
32. In some sense all of these theoretical works have examined only short term uncertainty, e.g. weather and seasonal demand fluctuations, which are resolved after all prices and investment decisions are made.

33. As noted in chapter IV this immediately implies that if capacity is chosen optimally for a given load duration curve, then the optimal price will be slightly above average costs. We have also examined simple cases where peak demand is less elastic than demand at lower levels, with the result that optimal prices are below average costs.

34. As shown below, this result disappears when we allow prices and peaking capacity to be chosen with a one year lead time.

35. p.378 Fuss and McFadden do not consider pricing effects and use a one period capital model to derive their results.

36. This result does not hold for all demand specifications, since it is possible to find perverse examples in which price elasticities at different levels of demand change with \( \eta \) in such a way that optimal prices decline as uncertainty increases.

37. These assumptions correspond to those of frameworks III and IV in table 1.

38. We discuss in section VIII.G reasons why our assumptions may be leading us to this result.

39. Since for a lognormal distribution the median is below the mean, the demand scenario we analyze is somewhat more pessimistic than the level of demand corresponding to the mean realization of \( \eta \). The difference in the optimal choice of capital is negligible for the degrees of uncertainty which we consider.

40. As demonstrated in chapter VI, section A, the relevant rental charge for period \( t \) if there is no physical depreciation of capital would be the difference in real capital costs per kilowatt between period \( t \) and period \( t+1 \). To avoid having to calculate real cost escalation, we use a constant annual fixed charge rate, which when multiplied by the total capital cost gives the annual rental charge.

41. See EGEAS (1980)

42. If desired one could go back to find the optimal discrete approximation to this continuous investment path, however we have not done so here.

43. With physical depreciation net investments can be negative, although still not substantially so.

44. See Fuss and McFadden (1971)
45. Whenever one does not build new capacity one is also foregoing the option to rent that new capacity in period \( t \). This decision is not irreversible for period \( t+1 \) however, and is already reflected in the choice of capacity for period \( t \). We are ignoring the fact that initiating new capacity buys the option to defer finishing that capacity at the scheduled time. We speculate that this option is not as valuable as the option value we consider, since deferrals made after construction has already been started are quite expensive.

46. Although we do not examine it here, the possibility of holding excess peaking capacity is not without interest. Due to dramatic changes in relative oil and natural gas prices, many utilities which historically relied on these fuels for base load generation currently have excess peaking capacity.

47. With more than two generation technologies similar statements can be made about all such types of capacity.

48. All random walks were truncated after 30 years to guarantee that they were finite and to approximate the 30 year life span of real generating capacity.

49. Examining equations (3), and (4), it can be seen that if all costs are constant in real terms, \( G(k_t|P, n) \) will also be constant over time. Our constant elasticity specification further implies that real prices should be constant over time. The multiplicative error specification means that all load duration curves shift up and down equiproportionately, and hence optimal capacity choices are proportional to \( n \).

50. This is a particularly bothersome assumption since our pricing results suggest that the excess costs of holding too much capacity can be partially alleviated by lowering prices. We see no simple alternative.

51. This linear approximation means that our option value tends to underestimate the true option value, while the previous approximation tends to overestimate it. See Appendix F.

52. Our approximation of the option value foregone is only correct at the optimum and hence we cannot approximate its value when there is too much capacity chosen. If there is grossly too little capacity, then we can say that the option value is zero, however, simply because future decisions are not constrained in this case.

53. With high load factors, small changes in demand growth cause dramatic changes in the fraction of the year that the marginal unit of base load generating capacity is used. As a limiting case, a load factor of 100 percent (i.e. electricity consumed at a constant rate throughout the year) means that for sufficiently slow demand growth some capacity may not be used at all during the year. The excess costs will be enormous in such cases. This is why the option values increase so dramatically.
54. This result is largely due to the fact that we have only two generating technologies, and hence the tradeoff is between base and peaking capacity rather than peaking and "intermediate" capacity. With multiple generation technologies the transitions between different types of capacities will be less sensitive to operating costs.

55. Changes which decrease the attractiveness of base load generating capacity move the desired level of base load capacity further out into the tail of the probability distribution for demand, and hence make choices more sensitive to the degree of uncertainty. This is discussed below.

56. As long as shortage costs (\( \alpha \)) are large relative to peaking operating costs, (\( m_a \)), the expression for the loss of load probability (\( \Theta \)) in footnote 13 implies that a doubling of \( \alpha \) nearly halves \( \Theta \).

57. The confidence intervals shown in table 13 are only approximations since they use the approximation \( \log(1+r) = r \).

58. In our model forecasts of future growth rates are not revised downward as a result of slow demand growth, although this type of revision of structural parameter estimates appears very likely in real life.

59. Using "derated" capital costs, nuclear capacity is desirable as long as it will be used when available 79.41 percent of the year. Given that it will only be available 65 percent of the year due to forced outages, then capacity need only be used 79.4 \times 0.65 = 51.6 percent of the year.

60. It may be of interest to some readers that average cost prices are much easier to solve for than marginal cost prices, and hence the more realistic pricing assumption is also computationally faster.

61. All capacity choices were estimated until convergence at six significant digits, hence this small of a difference is not an aberration.

62. Since far too little base load is being built, the option to rent less capacity in the future has no value, and table 14 can also be thought of as using the tenth modeling framework. We are hesitant to do this, however, since here we are explicitly building too much peaking capacity. With too much peaking capacity relative to base load capacity, the choice of peaking capacity should take into account its effect of reducing the attractiveness of future base load investments. Hence we have returned to the fifth modeling framework for our analysis of the "one third less base load capacity" future.
IX. Summary and Conclusions

In this thesis we have developed a consistent theoretical framework for incorporating prices and uncertainties into an investment planning model. Multiple demand periods, multiple user classes, intra-period time varying demand and long term demand uncertainty are all collapsed into a single probability distribution which we call the load distribution curve. In the same way that the load duration curve characterizes demand in a deterministic world, the load distribution curve, the expectation of these curves, provides a useful summary of demand patterns in an uncertain setting.

Our simple treatment of demand uncertainty permits a number of new pricing and investment issues to be considered. We allow price elasticities to vary within each demand period, we do not impose any restrictions on cross price elasticities of demand between different demand periods, and we consider optimal pricing when there are multiple classes of users (e.g., residential and industrial) or multiple dimensions of product service (e.g., peak and total use) as well as multiple demand periods (e.g., daytime and nighttime). As a result our pricing rules show that optimal pricing should take into account not only the pattern of change within any demand period in response to a price change, but also between different demand periods. One interesting implication is that if there are restrictions on the number of different demand periods during which different prices can be charged, and if cross price elasticities of demand are non-zero, then welfare optimal prices are no longer bounded to lie between the operating costs of the least expensive (base load) and most expensive (shortages) technologies. A second interesting result is that if own and cross price elasticities are
constant, although not necessarily identical, within each demand period
then optimal prices will generate expected revenues which are slightly
greater than fixed costs plus expected operating costs.

Our load distribution curve provides a simple and intuitive picture
of the effect of increasing uncertainty and lengthening lead times on
investments, costs and prices. If the effect of increasing uncertainty
is to raise the probabilities that demand is below low levels and above
high levels of capacity, then by ignoring uncertainty and lead times
there will be systematic tendencies to invest in too much base load
capacity relative to total capacity. As a result costs will tend to be
higher than expected cost minimizing investments would imply:

One interesting set of issues which we examine using our
representation of demand uncertainty is the impact of sequential decision
making in an environment where uncertainty is being resolved slowly over
time. This slow resolution of uncertainty over time is important for
three reasons. First, prices are generally set after investments are
made, when most or all of the uncertainty is resolved. If these prices
are themselves dependent on capacity levels through either marginal cost
pricing or average cost pricing, then \textit{ex ante} investment decisions are
complicated substantially. These pricing effects are more readily
examined using a simulation model, which was done and is discussed below.
After considering this complication, we assumed long run marginal cost
pricing, which amounts to assuming prices are exogenous, in order to
derive further theoretical results with sequential decision making.

The slow resolution of uncertainty over time is also important
because different types of generating capacity require different lead
times, so that investment decisions must be made with differing amounts
of information for different types of generating capacity. In a one period capital model where investments are made in order of decreasing capital intensity, we have shown that the only modification to the static investment rules which must be made is that the load distribution curve on which investments are based must be updated each period to reflect the improved information which is available. Although the simplifying assumptions must be noted, this result is of great practical use since it vastly simplifies the ex ante choices of capacity: "screening curves" can simply be projected onto the appropriate load distribution curve.

The third reason why the slow resolution of uncertainty over time is important is that capital is durable, so that investment commitments made in one period may constrain capacity in subsequent periods to be higher than would be chosen optimally. If demand is growing so rapidly that capacity of each type is added in each period, then the durability of capital is not critical, and investments may be chosen using the same investment rules derived in our static model, with the one period rental cost of capital used in place of its capital cost. In the more realistic case where demand growth is slow and uncertain enough that desired capital stocks may decrease, then this fact should induce investments to be lower than static investment models would suggest.

We extended our model in four new directions using a one period capital model: uncertain input prices, interfuel substitution, average cost pricing, and risk aversion were each examined. Several interesting results were demonstrated. Fuel cost uncertainty was shown to be actually desirable if fuel prices were negatively correlated with the level of demand (or more precisely, the fractions of the year that demand is above various stated levels). The additional loading flexibility
allowed by interfuel substitutability may make a technology desirable
even if it has a higher expected fuel cost than a technology which does
not permit such flexibility.

Average cost pricing was seen to exacerbate consumption uncertainty,
and under average cost pricing it is no longer optimal to minimize
expected total costs. Several "second best" type effects were shown to
alter desired investments so as to differ from cost minimizing levels.
We dealt with optimal pricing and investments in the presence of risk
averse consumers and producers in a manner very similar to our treatment
of average cost pricing. When consumers and producers are both equally
risk averse, prices should still be set equal to marginal costs, but
again it is no longer optimal to minimize expected total costs. Not too
surprisingly, investments should be made not only so as to lower costs,
but also so as to reduce uncertainties. In the presence of demand
uncertainty this takes the form of increased capital intensity, while
with operating or capital cost uncertainty, relatively less reliance
should be placed on the more highly uncertain generating technologies.
Increasing consumer risk aversion while holding producer risk aversion
constant increases price uncertainty but reduces consumption uncertainty,
and hence further increases desired capital intensity.

The simulation model, developed here with three objectives in mind,
was less instructive than originally hoped. Our first objective was to
demonstrate that our theoretical model can be implemented on a practical
basis. We feel that we were successful in this endeavor: our basic model
incorporated a price feedback loop, uncertain demand, sequential decision
making and durable capital. Further extensions considered in a one
period capital model both uncertain operating costs and average cost
pricing. This was all feasible with relatively low computational burdens. The most difficult feature of our model was the modeling of the durability of capital. The simplifying assumptions that were used were strong, and it is not clear that generalizations to cases involving multiple dimensions of uncertainty (e.g., uncertain demand and fuel costs) or multiple option values are feasible.

In a one period capital model numerous extensions of our simulation model are possible. Although we considered only a single price, two generating technologies, and one or two uncertain variables; extending our one period capital model to cases involving more prices, generating technologies and dimensions of uncertainty appears very feasible. This is particularly true if the number of uncertain future scenarios that are analyzed is reduced from 100 to perhaps 5 or 10; this could be done without a great loss of realism.

A second objective of our simulation model was to verify certain theoretical results which depend on the demand and cost parameters. We verified that when prices and investments must be made simultaneously, increasing uncertainty reduces desired base load capacity, and this result holds even when prices and peaking capacity can be chosen after long term uncertainties have been realized. The correlation between the level of demand and operating costs was found to influence desired investments in the predicted direction, and average cost pricing was shown to increase consumption uncertainty and reduce desired capital intensity. Our simulation model displayed all of our theoretically predicted effects.

The third objective of our simulations was to establish that the effects of uncertainty were not only in the theoretically predicted
direction, but also that they are large enough to be of concern to applications. Our simulation model did not meet this objective. In fact the somewhat surprising finding of our model was the insignificance of incorporating uncertainty in long term investment planning. In almost every case ignoring uncertainties and basing decisions on the median future scenario resulted in base load capacity choices that differed less than 3 percent from those incorporating the full effect of uncertainties. The only exceptions were simulations involving slowly growing and highly uncertain demand, in which case the irreversibility effect becomes important, and simulations involving high base load capital costs or low peaking operating costs, which decrease the desired reliance on base load capacity and hence make it more sensitive to uncertainty. The suggestion of our simulation model is that uncertainties do not substantially influence base load capacity choices, and can be ignored for most practical purposes.

Here, however the simplified features of our simulation model should be reemphasized, for they undoubtably influence our results. First as we noted in chapter VIII, section B, our linear load duration curve provides a stringent test of the impact of uncertainty in that it lacks convexities. Second, the assumption of zero serial correlation between the random variables influencing demand over time reduces the value of the information gained by waiting until some or all uncertainties have been resolved. Third, we treat forced outages by simply derating capacity a fixed percentage. We discuss an extension of the model which could improve upon this assumption below. Fourth and finally, we consider only two generating technologies. The absence of some intermediate form of generating capacity means that base load capacity is
being overutilized, and hence is less influenced by changes in the tails of our load distribution curves. It is these tails which are the most sensitive to the degree of uncertainty. In light of these simplifications we believe that our simulation model is still inconclusive as to whether or not it is critical to consider uncertainties when choosing base load capacity.

Perhaps the most serious shortcoming of both our theoretical and simulation model is that we have not incorporated any supply side uncertainties about the availability of generating capacity. Forced outages, uncertain fuel availability, and uncertain lead times all have the property of reducing available capacity. One means of incorporating these "outage" demand uncertainties in our modeling framework is to treat them as a type of demand uncertainty, adding their uncertain demand (negative supply) onto the usual load duration curve. This technique introduces new theoretical and practical complications. The theoretical complication is that this outage demand does not depend directly on any price, but rather it is dependent on actual capacity choices and the level of non-outage demand. The practical difficulty is that it is computationally burdensome to repeatedly calculate outage demand for each of numerous different scenarios and price levels while converging to the optimal price and capacity configuration.

An alternative procedure which we have explored only superficially, is to treat unavailable capacity as a type of fuel cost uncertainty which is uncorrelated with other random variables affecting the level of demand. An advantage of this procedure is that it preserves the supply/demand dichotomy. Each generating unit of type i can be viewed as operating only a fraction of the year at an operating cost m_i, and the
remainder of the year at an operating cost $\alpha$, the shortage cost. Optimal rearrangement of the loading order mimics the effect of a unit being not available, and can be treated as a type of interfuel substitutability. If prices are not expected to adapt instantaneously to the availability or nonavailability of individual generating units (this rules out spot pricing of the type considered in Bohn et al., 1981) then it appears feasible by this technique to avoid having to recalculate outage demand for each price and each scenario. This would be a major simplification for computational purposes. These comments are readily seen to be very preliminary, but point to an area that merits further investigation.

This thesis has developed a theoretical framework for analyzing demand and input price uncertainty which we believe may be of major use to empirical planning and forecasting models. Our model differs fundamentally from other models. Conventional investment planning models typically consider three (sometimes more) different future scenarios, corresponding perhaps to high, medium and low demand growth. The question which these models answer is: if demand and fuel costs were to grow according to Scenario A, then what would be the ex post optimal mix of generating capacity to have chosen? The answer to this question is then derived using a deterministic load duration curve which may or may not incorporate pricing effects. The ex post optimal investment strategy corresponding to the medium scenario is then often adopted as the basis of policy recommendations.

The question which we have addressed is a very different one. We have asked: given today's uncertainty about future demand growth and input prices, what is our ex ante optimal investment strategy? We answer
this question by using load distribution curves which summarize possible levels of demand over all of the various demand scenarios. If we update the load distribution curves over time to reflect the decreases in uncertainty over time, then we could also answer a question similar to the conventional question: if information about future demand were to evolve according to Scenario A, then what would be the \textit{ex ante} optimal mix of generating capacity to have chosen? For policy recommendations, the answer to this question would appear to be of much greater interest than the answer to the conventional question.

Our theoretical framework suggests that the answers to these two sets of questions will differ systematically, however our simulation model tentatively suggests that the differences may not be all that important in practice. Whether or not this is the correct assessment awaits the development of a more realistic empirical model incorporating uncertainties in a manner consistent with the theory developed here.
APPENDICES A-F
Appendix A - Derivation of Consumer Surplus Measures

It is shown here that the area under the load distribution curve, integrated over prices is an appropriate approximation of expected consumer surplus. It is also shown that in the case of two prices, whether they be prices charged for different demand periods, different classes of users, or different product dimensions, the appropriate measure of consumers' surplus can always be put in one simple form.

Assume that there exists an indirect welfare function, $V(P, t, \eta)$ which is a twice continuously differentiable, quasi convex function of a vector of prices, P, time, t, a random variable, $\eta$, and other variables which are presumed held constant and hence which are suppressed. Using Roy's identity, $\frac{\partial V}{\partial P_i} = X_i \frac{\partial V}{\partial \text{income}}$, and assuming risk neutrality with the marginal utility of income equal to one, the Marshallian consumers' surplus from the consumption of electricity at a given instant of time, t, when all electricity is sold at a single price $P^0$ and other prices $\bar{P}$ are held constant can be written

$$\text{Consumers' surplus at } t = V(P^0, \bar{P}, t, \eta) - V(\infty, \bar{P}, t, \eta)$$

$$= \int_{P_0}^{\infty} \frac{\partial V}{\partial P} dp = \int_{P}^{\infty} X(p, \bar{P}, t, \eta) dp \quad (A-1)$$

We will suppress the prices $\bar{P}$ which are held constant. Integrating over the entire year $(0,1)$ while discounting we can derive the total discounted consumers' surplus for the year conditional on $\eta$: 
Consumers surplus over entire year = \( CS(p^0, n) \)

\[
= \int_0^1 \int_{p_0}^{\infty} e^{-rt} x(p, t, n) \, dp \, dt
\]

\[
= \int_{p_0}^{\infty} \int_0^1 e^{-rt} x(p, t, n) \, dt \, dp
\]

Since we are only considering demand during a single year, total discounted consumer surplus will not be substantially different from undiscounted consumer surplus. We will therefore follow the traditional approach and ignore intra-period discounting. For any possible price and realization of \( n \) we can always rearrange the levels of demand, \( x(p, t, n) \) in ascending order to form a load duration curve. Call the new functional representation of demand \( x(\theta, p, n) \) where \( \theta \) lies in the region \([0,1]\) and \( \frac{\partial x}{\partial \theta} > 0 \). Using our inverse representation of the load duration curve, \( x(\theta, p, n) = G^{-1}(\theta, p, n) \), defined in Section III, Part C:

\[
CS(p^0, n) = \int_{p_0}^{\infty} \left[ \int_0^1 G^{-1}(\theta|p, n) \, d\theta \right] dp \quad (A-2)
\]

The term in brackets is simply the integral of the area under the load duration curve. Using the change of variables \( \theta = G(x|p, n) \),

\( d\theta = g(x|p, n) \) and taking expectations over \( n \), expected consumer surplus can be written

\[
CS(p^0) = \int_S \int_{p_0}^{\infty} \int_0^{\infty} x \, g(x|p, n) \, dx \, dp \, f(n) \, dn = \int_{p_0}^{\infty} D(p) \, dp \quad (A-3)
\]
This is the form in which consumer surplus is most frequently analyzed in the main text, for example, equation (3-6).

An analogous derivation can be used to find expressions for consumer surplus when there are multiple prices to consider. Marshallian consumers' surplus from two different commodities, \( x_1(p_1, p_2, t, n) \) and \( x_2(p_1, p_2, t, n) \), can be calculated using any of the infinite number of line integral measures of consumer surplus. Assuming \( \frac{\partial^2 V}{\partial p_i \partial p_j} = \frac{\partial^2 V}{\partial p_j \partial p_i} \) for all \( i \) and \( j \) (which implies \( \frac{\partial^2 V}{\partial p_j \partial p_i} = \frac{\partial^2 V}{\partial p_i \partial p_j} \)) then all of these line integrals will obtain the same measure of consumer surplus. A particularly simple line integral is (ignoring \( t \) and \( n \))

\[
CS(p^1, p^2) = V(p^1, p^2) - V(\infty, \infty)
\]

\[
= [V(p^2, p^2) - V(\infty, p^2)] + [V(\infty, p^2) - V(\infty, \infty)]
\]

\[
= \int_{p^1}^{\infty} x_1(p, p^2)dp + \int_{p^2}^{\infty} x_2(\infty, p)dp \quad (A-4)
\]

\( x_1 \) and \( x_2 \) can each be ordered into ascending order, our inverse notation and change of variables used, and expectations taken over \( n \).

Consider these examples.

Suppose there is more than one demand period. Let \( p^1 \) = daytime price, \( p^2 \) = nighttime price. We can represent demand by separating the load duration curve into daytime \( (G^1) \) and nighttime \( (G^2) \) segments

\[
G(x|p, n) = G^1(x|p^1, p^2, n) + G^2(x|p^1, p^2, n)
\]

The appropriate measure of consumer surplus in this case is
CS(p_1, p_2, n) = \int_S \left[ \int_{p_1}^{\infty} \int_0^{\infty} xg_1(x \mid p, p_2, n) \, dx \, dp \right. \\
+ \left. \int_{p_2}^{\infty} \int_0^{\infty} [xg_2(x \mid \infty, p, n) \, dx \, dp] \right] f(n) \, dn \\
= \int_{p_1}^{\infty} D_1(p, p_2) \, dp + \int_{p_2}^{\infty} D_2(\infty, p) \, dp \quad (A-5)

Suppose there is more than one class of users. Let \( x_1 = \) residential demand, \( x_2 = \) industrial demand then:

\[ x = x_1 + x_2 \]

\( x_1, x_2 \) with joint density function \( g(x_1, x_2 \mid p_1, p_2, n) \)

\[ g_1(x_1 \mid p_1, n) = \int_0^{\infty} g(x_1, x \mid p_1, p_2, n) \, dx \]

\[ g_2(x_2 \mid p_2, n) = \int_0^{\infty} g(x_1, x_2 \mid p_1, p_2, n) \, dx \]

\[ CS(p_1, p_2) = \int_S \int_{p_1}^{\infty} \int_0^{\infty} xg_1(x \mid p, n) \, dx \, dp \]

\[ + \int_{p_2}^{\infty} \int_0^{\infty} xg_2(x \mid p, n) \, dx \, dp \, f(n) \, dn \]

\[ = \int_{p_1}^{\infty} D_1(p) \, dp + \int_{p_2}^{\infty} D_2(p) \, dp \quad (A-6) \]

Except for the fact that in this case it appears reasonable to assume
zero cross-price elasticities, the expressions are of the same form as above, and can be derived via integrals under a load distribution curve if desired.

Finally, suppose the prices correspond to two different dimensions of demand service. For example, suppose \( \bar{p} \) = price per Kw-hr., \( \bar{p} \) = price per Kw for peak Kw use, \( Q(\bar{p}, \bar{p}, \eta) \) = peak Kw use. Here the appropriate consumers' surplus measure is

\[
CS(\bar{p}, \bar{p}) = \int_{S} \left\{ \int_{0}^{\infty} \int_{0}^{\infty} xg(x \mid \bar{p}, p, \eta) \, dx \, dp \\
+ \int_{\bar{p}}^{\infty} Q(\infty, p, \eta) \right\} f(\eta) \, d\eta
\]

\[
= \int_{\bar{p}}^{\infty} D(p, \bar{p}) \, dp + \int_{\bar{p}}^{\infty} Q(\infty, p) \, dp
\]

Here the result is somewhat different in that only the left hand term can be derived via integrating under the load duration curve. The level of peak kilowatt use by individual users is well defined, however, and we expect it to change as \( \bar{p} \) changes. Whether these changes at the micro level affect costs at the aggregate level is an empirical question. If they do not then they should have a zero \textit{generation} charge. (Transmission costs may be non-negligible, justifying a fee, however).
Appendix B – Inverse Representation of Total Generating Costs

It is shown here that if investments are always made optimally, then total expected generating plus shortage costs can be expressed as the sum of areas under the load distribution curve weighted by the appropriate operating or shortage cost.

Equation (4-2) of the main text gave expected total fixed, operating and shortage costs as

\[
C(k, p) = \int_S \sum_{i=1}^{n} \left\{ c_i k_i + \int_{K_{i-1}}^{K_i} m_i (x - K_i) \cdot g(x | p, n) \, dx \right\} \\
+ \int_{K_i}^{\infty} m_i k_i \cdot g(x | p, n) \, dx \\
+ \int_{K_n}^{\infty} \alpha (x - K_n) \cdot g(x | p, n) \, dx \cdot f(n) \, dn
\]  

(B-1)

Omitting the price argument, using \(\int_S g(x | n) f(n) \, dn = h(x)\), \(k_i = K_i - K_{i-1}\) and \(\int_{K_i}^{\infty} h(x) \, dx = 1 - H(K_i)\), and adopting the conventions \(K_0 = 0\), \(m_{n+1} = \alpha\), \(c_{n+1} = 0\) we may rewrite (B-1) as

\[
C(K^*) = \sum_{i=1}^{n} \left( (c_i - c_{i+1}) K_i + m_i \int_{K_{i-1}}^{K_i} x h(x) \, dx \right) \\
- (m_{i+1} - m_i) K_i [1 - H(K_i)] \\
+ \alpha \int_{K_n}^{\infty} x h(x) \, dx
\]

(B-2)
The first order conditions (4-4) and (4-5) in the general case imply

\[ c_i - c_{i+1} = (m_{i+1} - m_i) [1 - H(K_i)] \]  \hspace{1cm} (B-3)

substituting this in B-2 the first and third terms cancel, leaving

\[ C(K^*) = \sum_{i=1}^{n} \int_{K_{i-1}}^{K_i} x h(x) dx + \alpha \int_{K_1}^{\infty} x h(x) dx \]  \hspace{1cm} (B-4)

Using the change of variables: \( x = H^{-1}(\theta) \), \( h(x)dx = d\theta \), and noting \( H(K_1^*) = \theta_1^* \), \( H(\infty) = 1 \) we can rewrite this as

\[ C(\theta^*) = \sum_{i=1}^{n} \int_{\theta_{i-1}}^{\theta_i^*} m_i H^{-1}(\theta)d\theta + \int_{\theta_n^*}^{1} \alpha H^{-1}(\theta)d\theta \]  \hspace{1cm} (B-5)

This expression indicates that total expected generating costs may be derived by considering the areas of vertical bands under the load distribution curve, and weighting these areas by the appropriate operating cost \( m_i \) or shortage cost, \( \alpha \). Since the \( \theta_i^* \) are independent of prices, this is a particularly convenient expression for evaluating price changes. A diagrammatic representation of these total costs in the simple case of only two generating technologies is presented in Figure (B-1).
Figure B-1

Diagrammatic Representation of Total Generating Costs

When Capacity is Always Chosen Optimally

\[ \theta = H(x|P) \]
\[ x = H^{-1}(\theta|P) \]

\( \theta = \text{fraction of period that demand is below level } x \)
Appendix C - Profitability Results of Alternative Pricing Specifications

We prove here that in the case where prices and investments are all chosen simultaneously, the expected revenues generated by welfare optimal prices will be above, equal to or less than expected fixed costs plus operating cost plus shortage costs according to whether price elasticities of demand increase, stay constant or decrease with the level of demand. The intuitive content of the results derived here are discussed more fully in section IV, parts C and D of the main text.

In deriving our results we will rely heavily on expressions (B-4) and (B-5) derived in Appendix B for representing total expected generating plus shortage costs. Using the conventions $K_0^* = 0$, $K_{n+1}^* = \infty$, $m_{n+1} = \alpha$, and $\theta_i^* = H(K_i^*)$, $i = 0, \ldots, n+1$ (hence $\theta_0^* = 0$ and $\theta_{n+1}^* = 1$) we can write equations (B-4) and (B-5) in the simplified form

$$C(K^*, P) = \sum_{i=1}^{n+1} \int_{K_{i-1}^*}^{K_i^*} m_i x h(x, P) \, dx$$

(C-1)

$$= \sum_{i=1}^{n+1} \int_{\theta_{i-1}^*}^{\theta_i^*} m_i X(\theta, P) \, d\theta$$

(C-2)

$$= C(\theta^*, P)$$

Recall that this expression includes all capital costs, $\sum_{i=1}^{n} c_i (K_i^* - K_{i-1}^*)$, which are reflected in the choice of the $K_i^*$ and hence the $\theta_i^*$.

Using (C-2) we can conveniently express the long run marginal costs of a price change as
\[
\frac{\partial C(\theta^*, p)}{\partial p} = \sum_{i=1}^{n+1} m_i \int_{\theta_{i-1}^*}^{\theta_i^*} \frac{\partial X(\theta, p)}{\partial p} \, d\theta \\
= \sum_{i=1}^{n+1} m_i \int_{\theta_{i-1}^*}^{\theta_i^*} \frac{\partial X}{\partial p} \frac{p}{X} \frac{X}{p} \, d\theta \\
= \sum_{i=1}^{n+1} m_i \int_{\theta_{i-1}^*}^{\theta_i^*} -\mu(\theta, p) \frac{X(\theta, p)}{p} \, d\theta \tag{C-3}
\]

where \( \mu(\theta, p) \) is the elasticity of total demand with respect to price \( p \).

The long run marginal cost of a price change is thus a weighted average of all operating and shortage costs, where the weights are the integrals of total demand, \( X(\theta, p) \), weighted by the price elasticities of demand \( \mu(\theta, p) \), divided by the price \( p \), for all times during which each technology is the marginal generating technology.

Using the optimal pricing rule in the case where a single price is charged for all electricity, as derived in the main text, equation (4-7), we can write the welfare optimal price, \( p^* \), as

\[
p^* = \frac{\frac{\partial C}{\partial p}}{\frac{\partial D}{\partial p}} = \sum_{i=1}^{n+1} m_i \int_{\theta_{i-1}^*}^{\theta_i^*} \mu(\theta, p) X(\theta, p) \, d\theta = \frac{\int_{0}^{1} \mu X(\theta, p) (\theta, p) \, d\theta}{\int_{0}^{1} X(\theta, p) (\theta, p) \, d\theta} \tag{C-4}
\]

Here it is readily seen that the optimal price, \( p^* \), is an elasticity and demand weighted average of all operating and shortage costs.

In comparing welfare optimal expected revenue to expected costs we will use the following lemma.
Lemma C-1: If for all \( \theta \in [\theta, \overline{\theta}] \) \( X(\theta) > 0, \mu(\theta) > 0, \frac{\partial \mu(\theta)}{\partial \theta} > 0 \) and the

\[
\int_{a}^{\theta} \mu(\theta) X(\theta) \, d\theta \quad \text{and} \quad \int_{a}^{\theta} X(\theta) \, d\theta
\]

are both finite for all \( a \in [\theta, \overline{\theta}] \), then

\[
\frac{\int_{a}^{\theta} \mu(\theta) X(\theta) \, d\theta}{\int_{a}^{\theta} \mu(\theta) X(\theta) \, d\theta} > \frac{\int_{a}^{\theta} X(\theta) \, d\theta}{\int_{a}^{\theta} X(\theta) \, d\theta}
\]

The intuition behind lemma C-1 is that if we attach positive and strictly increasing weights \( \mu(\theta) \) to any strictly positive function \( X(\theta) \) while integrating over \( \theta \) then the weighted integral over high values of \( \theta \) will contain a greater proportion of the total integral than the unweighted integral.

Proof:

\[
\frac{\partial \mu(\theta)}{\partial \theta} > 0 \Rightarrow \mu(\theta) > \mu(a) \quad \text{for all} \quad \theta > a
\]

\[
\Rightarrow \mu(\theta) X(\theta) > \mu(a) X(\theta) \quad \text{for all} \quad \theta > a
\]

\[
\Rightarrow \int_{a}^{\theta} \mu(\theta) Z(\theta) \, d\theta > \int_{a}^{\theta} \mu(a) X(\theta) \, d\theta
\]

\[
\Rightarrow \int_{a}^{\theta} \mu(\theta) X(\theta) \, d\theta > \mu(a)
\]

By an opposite argument
\[ \mu(a) > \frac{\int_{\theta}^{a} \mu(\theta) X(\theta) \, d\theta}{\int_{\theta}^{a} \mu(\theta) X(\theta) \, d\theta} \]

Combining the two and rearranging yields

\[ \frac{\int_{\theta}^{a} X(\theta) \, d\theta}{\int_{a}^{\theta} X(\theta) \, d\theta} \quad > \quad \frac{\int_{\theta}^{a} \mu(\theta) X(\theta) \, d\theta}{\int_{a}^{\theta} \mu(\theta) X(\theta) \, d\theta} \]

Adding one to both sides and taking inverses yields the desired result.

Q.E.D.

Equation (C-4) representing the optimal price, \( p^* \), can conveniently be rewritten as

\[ p^* = \sum_{i=1}^{n+1} (m_i - m_{i-1}) \left( \int_{\theta_{i-1}}^{1} \mu(\theta, P) X(\theta, P) \, d\theta \right) \frac{1}{\int_{0}^{1} \mu(\theta, P) X(\theta, P) \, d\theta} \quad (C-5) \]

This expression can be contrasted with the expected total of fixed, operating and shortage costs, \( C(\theta^*, P) \), divided by expected demand, \( D(P) \), which can be written

\[ \frac{C(\theta^*, P)}{D(P)} = \sum_{i=1}^{n+1} (m_i - m_{i-1}) \left( \int_{\theta_{i-1}}^{1} X(\theta, P) \, d\theta \right) \frac{1}{\int_{0}^{1} X(\theta, P) \, d\theta} \quad (C-6) \]
Note first that if the price elasticity of demand with respect to \( p \), \( \mu(\phi, p) \) is constant over different values of \( \phi \) \( \frac{\partial \mu}{\partial \phi} = 0 \) then the elasticities cancel out of equation (C-5) and expressions (C-5) and (C-6) are equivalent. This implies that welfare optimal prices are equal to total expected costs divided by total expected demand.

A direct application of Lemma C-1 to the corresponding terms in (C-5) and (C-6) for \( i=2,3,..., n+1 \) yields the result that if \( \mu(\phi, p) \) is increasing in \( \phi \), then \( p^* > \frac{C(\phi, p)}{D(p)} \). By a symmetric argument it is readily seen that if \( \mu(\phi, p) \) is decreasing in \( \phi \) then the optimal price will be below expected costs divided by expected demand.

The above results have important implications about revenues and profits. Since shortages may occur when demand exceeds generating capacity, expected sales will in general be less than expected demand, and the firm will neither collect revenue for nor bear the costs of the electricity which is demanded but not served. If the firm did collect revenue for this unmet demand, then the above results imply that expected total revenue would be above, equal to or less than expected total fixed, operating and shortage costs according to whether the price elasticity of demand was increasing, constant or decreasing with \( \phi \). Since the firm does not in fact collect this revenue, the firm's revenue is decreased by \( p \int_{K_{n+1}}^{\infty} (x - K_{n+1}) h(x|p) \, dx = pZ \) while the shortage costs not born by the firm are \( \alpha \int_{K_{n+1}}^{\infty} (x-K_{n+1}) h(x|p) \, dx = \alpha Z \). With \( p<\alpha \) this implies that profits will be higher than they would be if these costs and revenues were internalized by the firm. In other words, we need to modify the above profitability result to say that expected
profits will be above, equal to, or below \((a-p)Z\) according to whether
the price elasticity of demand is increasing, constant or decreasing,
respectively with \(\theta\).

These pricing and profitability results can be extended to cases
where there are several prices. In the case of two disjoint demand
periods, the distribution of demand can be represented by separating the
load distribution curve \(H(x|P)\) into two probability distributions,
\(H^1(x|P)\) and \(H^2(x|P)\) and total expected costs can be written

\[
C(K,P) = \sum_{i=1}^{n+1} \left[ c_i(K_i - K_{i-1}) \right. \\
+ (m_i - m_{i-1}) \int_{K_{i-1}}^{\infty} (x-K_{i-1}) \left[ h^1(x|P) + h^2(x|P) \right] dx \right]
\]

In the case where own and cross price elasticities of demand within each
of the two demand periods are invariant to the level of \(X\) (and hence \(\theta\))
there will exist equivalent representations of demand in each period:

\(xh^1(x|P) = a^1(P)z\tilde{h}^1(z)\) and \(xh^2(x|P) = a^2(P)z\tilde{h}^2(z)\)

so that total costs can be written as

\[
C(K,P) = \sum_{i=1}^{n+1} \left[ c_i(K_i - K_{i-1}) \right. \\
+ (m_i - m_{i-1}) \int_{K_{i-1}}^{\infty} \frac{a^1(P)z - K_{i-1}}{a^1(P)} \tilde{h}^1(z) \, dz \\
+ (m_i - m_{i-1}) \int_{K_{i-1}}^{\infty} \frac{a^2(P)z - K_{i-1}}{a^2(P)} \tilde{h}^2(z) \, dz \right]
\]  
\(\text{(C-8)}\)
Differentiating with respect to \( p_j \), \( j = 1, 2 \) and simplifying yields

\[
\frac{\partial C}{\partial p_j} = \sum_{i=1}^{n+1} [(m_i - m_{i-1}) \frac{\partial a}{\partial p_j} \frac{a^1}{p_j} \int_{K_{i-1}}^{\infty} z \tilde{h}^1(z) \, dz] + [(m_i - m_{i-1}) \frac{\partial a^2}{\partial p_j} \frac{a^2}{p_j} \int_{K_{i-1}}^{\infty} z \tilde{h}^2(z) \, dz]
\]

\[
= \sum_{i=1}^{n+1} [(m_i - m_{i-1}) \frac{a^1}{p_j} \frac{1}{a^1(p)} \int_{K_{i-1}}^{\infty} \frac{a^1(p)z}{p_j} \tilde{h}^1(z) \, dz] + [(m_i - m_{i-1}) \frac{a^2}{p_j} \frac{1}{a^2(p)} \int_{K_{i-1}}^{\infty} \frac{a^2(p)z}{p_j} \tilde{h}^2(z) \, dz]
\]

\[
= -\frac{\nu_{1j}}{p_j} \sum_{i=1}^{n} (m_i - m_{i-1}) \int_{K_{i-1}}^{\infty} x h^1(x, p) \, dx + \frac{\nu_{2j}}{p_j} \sum_{i=1}^{n} (m_i - m_{i-1}) \int_{K_{i-1}}^{\infty} x h^2(x, p) \, dx
\]

\[
= -\frac{\nu_{1j}}{p_j} C_1^*(p) + \frac{\nu_{2j}}{p_j} C_2^*(p) \quad (C-9)
\]

where \( \nu_{1j} \) and \( \nu_{2j} \) are the \( p_j \)th price elasticity of demand for periods 1 and 2, respectively and \( C_1^* \) and \( C_2^* \) are as defined in (C-9). By an argument similar to that in appendix B, \( C_1^* \) and \( C_2^* \) can be shown to add up to the expected total of fixed, operating and shortage costs, \( C(\theta^*, P) \). Using (C-9) and the optimal pricing rules (4-10) and (4-11) in
the main text it is readily seen that the welfare optimal prices $p_1$ and $p_2$ must satisfy the equations.

\[ p_1 \mu_{11}D^1 + p_2 \mu_{21}D^2 = \mu_{11}C^1 + \mu_{21}C^2 \]  \hspace{1cm} (C-10)

\[ p_1 \mu_{12}D^1 + p_2 \mu_{22}D^2 = \mu_{12}C^1 + \mu_{22}C^2 \]  \hspace{1cm} (C-11)

Whether or not the cross price elasticities are zero, the only solution to these equations is

\[ P_1 = \frac{C^1}{D^1} \quad \text{and} \quad P_2 = \frac{C^2}{D^2} \]

The total revenue collected if all of $D^1$ and $D^2$ were sold at prices $p_1$ and $p_2$ respectively would be

\[ \text{Revenue} = p_1 D^1 + p_2 D^2 = C^1 + C^2 = C(\theta^*, P) \]

We have therefore shown that in the case where optimal prices and investments are both chosen simultaneously, and own and cross price elasticities of demand within each demand period are constant, then welfare optimal prices will generate revenues sufficient to cover costs.

A similar argument can be used to analyze the case of two different classes of concurrent users whose individual demands, $X_1$ and $X_2$, sum up to be total demand $X = X_1 + X_2$. As in appendix A, their pattern of consumption throughout the year can be represented by a joint probability density function $h(x_1, x_2 | p_1, p_2)$ and our usual load distribution curve can be derived as
\[ h(x|p_1, p_2) = \int_0^\infty h(x - x_2, x_2|p_1, p_2) \, dx_2 \]

In this case we can write total costs as

\[ C(K, P) = \sum_{i=1}^{n+1} c_i (K_i - K_{i-1}) \]

\[ + (m_i - m_{i-1}) \int_{K_{i-1}}^\infty (x - K_{i-1}) h(x|P) \, dx \]

\[ = \sum_{i=1}^{n+1} c_i (K_i - K_{i-1}) \]

\[ + (m_i - m_{i-1}) \int_{K_{i-1}}^\infty \int_0^\infty (x - K_{i-1}) h(x - x_2, x_2|P) \, dx_2 \, dx \]

\[ = \sum_{i=1}^{n+1} c_i (K_i - K_{i-1}) \]

\[ + (m_i - m_{i-1}) \int_0^\infty \int_{K_{i-1} - x_2}^\infty (x_1 + x_2 - K_{i-1}) h(x_1, x_2|P) \, dx_1 \, dx_2 \]

\[ = \sum_{i=1}^{n+1} c_i (K_i - K_{i-1}) \]

\[ + (m_i - m_{i-1}) \int_0^\infty \int_{K_{i-1} - x_2}^\infty (x_1 + x_2 - K_{i-1}) h(x_1, x_2|P) \, dx_1 \, h(x_2|P) \, dx_2 \]

Once again, if price elasticities of demand are independent of the level of demand, \( x \), then there will exist an equivalent representation of demand
\[ x_1 h(x_1 \mid x_2, p) = a_1(p_1) z_1 h^1(z_1 \mid z_2) \quad \text{and} \quad x_2 h(x_2 \mid p) = a_2^2(p_2) z_2 h^2(z_2) \]

Here we have also used the fact that with different classes of uses it is reasonable to assume that the cross price elasticities of demand are zero. Hence an equivalent representation of costs will be

\[
C(K, P) = \sum_{i=1}^{n+1} c_i (K_i - K_{i-1})
\]

\[
+ (m_i - m_{i-1}) \int_0^{\infty} \left\{ \begin{array}{l}
\frac{a_1^2(p_1) z_1 + a_2^2(p_2)(z_2 - K_{i-1})}{K_{i-1} - a_2^2(p_2) z_2} \\
\frac{a_1^2(p_1)}{a_1^1(p_1)}
\end{array} \right\} dz_1 h^1(z_1) \, dz_2 h^2(z_2)
\]

Differentiating with respect to \( p_1 \) and simplifying yields

\[
\frac{\partial C}{\partial p_1} = -\frac{\mu_{11}}{p_1} \int_0^{\infty} \int_{K_{i-1} - x_2}^{\infty} x_1 h^1(x_1 \mid x_2, p) \, dx_1 h^2(x_2 \mid p_2) \, dx_2
\]

\[
= -\frac{\mu_{11}}{p_1} C_1^* + C_1^*
\]

By an analogous argument it can be shown that

\[
\frac{\partial C}{\partial p_2} = -\frac{\mu_{22}}{p_2} C_2^* + C_2^*
\]

Using these results in our expression for optimal pricing (equation 4-12 in the main text) immediately gives us
\[ p_1 = \frac{C_1^*}{D_1} \quad \text{and} \quad p_2 = \frac{C_2^*}{D_2} \]

As in the previous case, it can be shown that \( C_1^* + C_2^* = C(d^*, P) \) if investments are made optimally, and that if revenue for all demand were to be collected, then expected revenue would cover expected shortage, operating and fixed costs. Since expected sales will be less than expected demand and shortage costs are not born by the firm, actual profits will be slightly positive.

Three comments about our pricing results are worth emphasizing. First, there is a common misperception that simply because electric utilities have high fixed costs, then welfare optimal prices will not generate revenues sufficient to cover total operating plus fixed costs. We have shown here that once uncertainty and shortage costs are properly taken into consideration, whether or not expected revenues from optimal prices will cover expected generating costs depends largely on the empirical properties of price elasticities of demand. There should be no general presumption about expected profits being negative at the welfare optimum without reference to these elasticities.

Second, although there are many important complications to the electric rate setting process, at the heart of most procedures is the precept that the average rate should be set equal to the average cost of providing electricity. The analysis here suggests that average cost pricing would be a very close (slightly too low) approximation to welfare optimal pricing if all electricity were sold at a uniform price and if price elasticities of demand were constant throughout the year and over
different realizations of any random variables. Of course inflation, multiple rate classes and demand periods, declining block rates, and historic cost pricing all cloud this result. It is interesting however that at its heart the common precept of average cost pricing is not without some merit.

A third comment is that here we have focused on the importance of price elasticities of demand and their relationship to profitability. Others (e.g. Crew and Kleindorfer, 1979) have focused on whether the (scalar) uncertain variable \( \eta \) affects demand multiplicatively (i.e. \( x^1 = a_1(P)\eta \)) or additively (i.e. \( x^2 = a_2(P) + \eta \)). It is readily seen that price elasticities of demand \( (\mu_1 = \frac{\partial a_1}{\partial p} \frac{p}{a_1} \text{ and } \mu_2 = \frac{\partial a_2}{\partial p} \frac{p}{(a_2 + \eta)} \) respectively) in these two cases are constant and decreasing, respectively, with respect to the realization of the uncertain variable. In restricting their analysis to these special cases, others have restricted price elasticities of demand to behave in a given manner and imposed certain profitability results even before considering how this demand will be supplied. It would appear more fruitful not to impose these restrictions \textit{ex ante}. 
Appendix D – Costs of Holding the Wrong Capacity Mix

In this appendix we derive an expression indicating the cost of holding the wrong capacity mix which should prove useful in choosing investments in the presence of indivisibilities.

In the case with two generating technologies, suppose a utility holds \((K_1, K_2^*)\) instead of \((K_1^*, K_2^*)\). The additional costs borne for any given set of prices \(P\) are then:

\[
C(K_1, K_2^*, P) - C(K_1^*, K_2^*, P)
\]

\[
= c_1 K_1 + c_2 (K_2^* - K_1) + \int_0^\infty m_1 x h(x) \, dx
+ \int_{K_1}^{\infty} (m_2 - m_1)(x - K_1) \, h(x) \, dx
+ \int_{K_2}^{\infty} (\alpha - m_2)(x - K_2^*) \, h(x) \, dx
\]

\[
- [c_1 K_1^* + c_2 (K_2^* - K_1^*) + \int_0^\infty m_1 x h(x) \, dx
+ \int_{K_1^*}^{\infty} (m_2 - m_1)(x - K_1^*) \, h(x) \, dx
+ \int_{K_2^*}^{\infty} (\alpha - m_2)(x - K_2^*) \, h(x) \, dx]
\]

\[
= (c_1 - c_2)(K_1 - K_1^*) + (m_2 - m_1)(K_1^* - K_1)[1 - H(K_1^*)]
\]
\[
+ (m_2 - m_1) \int_{K_1^*}^{K_1} (x-K_1) \, h(x) \, dx
\]

Using the first order conditions for optimal investment derived above, the first two terms cancel, yielding

\[
C(K_1, K_2^*) - C(K_1^*, K_2^*) = (m_2 - m_1) \int_{K_1^*}^{K_1} (x - K_1) \, h(x) \, dx
\]

If \( K_1 > K_1^* \) so that too much baseload capacity is held relative to peaking capacity, then the increment to total generating costs is simply the fuel cost differential between peak and base capacity multiplied by the unused generating capability of the excess capacity when it is the marginal capacity, i.e. when the excess is being used, but at less than full capacity. If \( K_1 > K_1^* \) so that there is too little base load capacity, the incremental generating costs are the fuel cost difference multiplied by the generation by peaking capacity as it is being used to make up the shortfall in base capacity, but before demand reaches \( K_1 \). Figure D-1 illustrates diagrammatically what these excess costs are. The costs of too much or too little capacity are regions A and B, respectively, in the diagram multiplied by the fuel cost differential, \((m_2 - m_1)\).

Three aspects of this representation of excess costs are noteworthy. First, excess costs are small relative to total generating costs if the capacity excess or shortfall is not too large, however they increase substantially as the discrepancy grows in magnitude. In fact, to a first order approximation, the excess cost grows in proportion to the square of the size of the capacity excess or shortfall.
Cost of the Wrong Generating Mix

Cost of too much base load capacity \((K_1^* > K_1)\) is \((m_2 - m_1)\) times area A

Cost of too little base load capacity \((K_1'' < K_1^*)\) is \((m_2 - m_1)\) times area B
Second, the excess cost of a given deviation from the optimal capacity holdings decreases as the load duration curve becomes steeper, i.e. demand is more dispersed. This result has important ramifications with uncertainty, for increased uncertainty has precisely this effect on the load distribution curve.

Finally, as an empirical observation, the load duration curve, \( G(x|P) \), is typically convex in \( x \) in the neighborhood of the optimal change over point between base and peaking capacity, \( \theta^*_1 \). For a given deviation from the optimum \( K^*_1 \) this implies that the cost of excess capacity is greater than the cost of a capacity shortfall. Cost minimizing behavior in the presence of increased uncertainty motivates a utility to move towards a situation where base load shortages are more likely. It turns out empirically that the reverse result holds for peaking capacity: since the load duration curve is curving upward at \( (K^*_2, \theta^*_2) \), capacity shortages of a given magnitude are more expensive than excess capacity. Increased uncertainty motivates a firm towards more peaking and total capacity than in the certain world.
Appendix E - Detailed Description of Simulation Model

In this appendix we describe more fully the simulation model used in chapter VII and discuss the algorithms used to solve for investments and prices with and without taking into account uncertainty. The calculation of option values and their incorporation into the investment choice process is considered in appendix F.

As stated in the main text, we model demand, $D_t$, as:

$$D_t(P_t, \eta_t) = a P_t^\gamma \eta_t \quad \text{(E-1)}$$

where
- $a =$ scaling factor
- $P_t =$ price in period $t$
- $\gamma =$ price elasticity of demand
- $\eta_t =$ random variable affecting demand

The load factor (LF) is specified exogenously. The linear load duration curve is represented by a uniform probability density function, $g(x)$, with endpoints $X$ and $\overline{X}$, where

$$\overline{X} = D_t/LF$$
$$X = D_t(2-1/LF)$$

Figures E-1a and E-1b illustrate the probability density and distribution functions, respectively, corresponding to a given realization of $\eta$, i.e. a given load duration curve.

We used a 30 to 80 point approximation (depending on the level of demand) of this probability distribution which facilitates the collapsing of multiple load duration curves into a single load distribution curve. This approximation is complicated by the fact that $X$ and $\overline{X}$ need not be
integer values, and hence the endpoints must be dealt with in some manner. Figures E-1c and E-1d illustrate a step function approximation which we tried and rejected, since the discontinuities at each point make the analysis of small capacity changes very difficult, and transform the investment process into an integer programming problem. Instead we used the approximation depicted in figures E-1e and E-1f. This piecewise-linear approximation of the load duration curve requires exactly the same memory as the more conventional step function approximation, but it is continuous and there is a one-to-one correspondence between levels of capacity and fractions of the year over the relevant intervals. Its only deficiency is that the probability density function is still discontinuous near the two end points. This is inconsequential for $\bar{x}$ but significant for $\bar{X}$, since peaking capacity is chosen to be very near $\bar{X}$, and hence is affected by this discontinuity. It has the further effect of causing discontinuities in the price derivatives of demand and costs, and hence complicating the calculation of marginal costs. Fortunately the problem largely disappears once uncertainties are introduced and both end points become somewhat "fuzzy".

We have described our method of representing the load duration curve at some length in part because we have not seen this convenient approximation used elsewhere, and in part because we wish to emphasize that our linearity approximation is not critical. Since we already use a discrete representation of the load duration curve, a discrete approximation of an alternative probability distribution could be used (e.g. a normal or higher moment approximation) without an overwhelming increase in computational burdens.

One of the central features of our model is our technique of
collapsing multiple future scenarios into a single probability distribution of demand. We use a notation here in which each future scenario is represented by a single dimensional random variable \( \eta^i \), although it should be clear from our notation that multiple dimensions of demand uncertainty can readily be handled. A second dimension of uncertainty is introduced below in the form of peak operating cost uncertainty. If \( f(\eta^i) \) is the probability of scenario \( \eta^i \) occurring and \( g(x|P^i, \eta^i) \) is the conditional probability that demand is at level \( x \) given \( P^i \) and \( \eta^i \) then \( h(x) \), the unconditional (or marginal) probability that demand is at level \( x \) is simply

\[
h(x) = \sum_{i} g(x|P^i, \eta^i) f(\eta^i) \tag{E-2}
\]

This adding up procedure must of course be completed for each possible level of demand. Upon completion of this procedure one has a piecewise linear approximation of the load distribution curve, which can then be used to find optimal capacity choices and expected costs. Figure E-2 depicts a simple 10 point representation of a load distribution curve derived from three equally likely load duration curves.

For any given load duration or load distribution curve, finding the cost minimizing capacity choices is straightforward. Solving equations (E-3) and (E-4) for the optimal choices of base and peaking capacity

\[
c_1 + m_1[1 - G(k_1|P, \eta)] = c_2 + m_2[1 - G(k_1|P, \eta)] \tag{E-3}
\]

\[
c_2 + m_2[1 - G(k_1+k_2|P, \eta)] = \alpha [1 - G(k_1+k_2|P, \eta)] \tag{E-4}
\]
Figure E-2

Example of Deriving a Load Distribution Curve From Three Equally Likely Load Duration Curves

- **Low demand**
  \[ g(x|\eta^1) = 0.33 \]

- **Medium demand**
  \[ g(x|\eta^2) = 0.25 \]

- **High demand**
  \[ g(x|\eta^3) = 0.20 \]

- **Density function representation of load distribution curve**
  \[ h(x) \]

- **Load Distribution Curve**
  \[ H(x) \]
where
\[ c_1, c_2 = \text{annual capital cost per kw of base and peak capacity,} \]
\[ m_1, m_2 = \text{operating cost per kw-year of base and peak capacity,} \]
\[ \alpha = \text{shortage cost per kw-year} \]

amounts to finding two cut off points, \( k_1 \) and \( k_1 + k_2 \), such that the cumulative probabilities that demand is above those levels are equal to certain fractions. For our piecewise linear load distribution and load duration curves the cut off points are unique. In the one period capital case these fractions are simple ratios of capital and operating cost differences. In the engineering literature, this solution is equivalent to projecting screening curves down on to the load duration curve. In our model we use the load duration curves for capacity decisions where capacity can be chosen without any uncertainty about \( n \), but our load distribution curve where decisions must be made before uncertainties about future demand are resolved.

Solving for the desired price conditioned on a given set of investments is also straightforward. For short run marginal cost pricing, i.e. when prices are chosen optimally in the last period based on existing capacity choices, we mentioned in the main text that the socially optimal price can be approximated by solving

\[ p^* = \frac{C(k_1, k_2, P^*) - C(k_1, k_2, P^* + \xi)}{D(P^*) - D(P^* + \xi)} \]  
\[ (E-5) \]

[\( \xi \) small]

where
\[ C(k_1, k_2, P^*) = \text{total generating plus shortage costs,} \]
given \( k_1, k_2, \) and \( P^*. \)
\[ D(P^*) = \text{total demand given price } P^*. \]
For long run marginal cost pricing, an equation analogous to equation (E-5) can be used, only instead of taking capacity choices to be fixed, levels of capacity are chosen so as to minimize expected generation costs. When average cost pricing is employed, the desired price is found by solving

\[ p^* = \frac{C(k_1, k_2; p^*)}{D(p^*)} \]  \hspace{1cm} (E-6)

where

\[ C = \text{total capital plus operating costs excluding shortage costs.} \]
\[ D = \text{total demand minus unserved demand} \]

Unfortunately neither equation (E-5) nor equation (E-6) are in closed form, and hence numerical methods must be used. We used an IMSL subroutine which employed Brown's Method, a modified Newton-Raphson type procedure, and found optimal prices in an average of about 20 iterations.

One of the primary goals of our model is to make sure the prices and investments are mutually consistent. In modeling frameworks in which prices and investments are chosen simultaneously, equations (E-3), (E-4) and (E-5) can in principle be solved simultaneously to find the socially optimal set of prices and investments. Unfortunately this procedure breaks down in cases where decisions must be made sequentially based on differing amounts of information. If for instance prices are set after investments are made then investments need to take into account not just a single price, but rather a whole array of different prices which are contingent on the way in which (uncertain) demand and cost variables grows. Solving for all of these prices and levels of capacity simultaneously runs into problems of dimensionality. To cope with this dimensionality problem we used an iterative procedure which takes advantage of the fact that the pricing subproblems corresponding to each demand growth scenario are
mutually exclusive, and hence can be solved independently. By iterating between investments and prices which are conditioned on those prices a set of mutually consistent investments and prices can be reached.

The procedure is most easily explained for the case without any uncertainty. First an initial set of capacities, \( k^0 \), are chosen. The optimal price based on those levels of capacity are then found by solving equation (E-5). Based on this price, \( p^1(k^0) \), we can generate a load duration curve \( G(x|p^1) \) which can in turn be used to find the cost minimizing set of capacities, \( k^1(p^1) \) for serving \( G(x|p^1) \). This set of capacities can then be used to generate a new set of prices, \( p^2(k^1) \) and the process can be repeated until convergence. Schematically we can represent this process by

\[
\begin{align*}
k^0 & \rightarrow p^1(k^0) \rightarrow k^1(p^1) \rightarrow p^2(k^1) \rightarrow \ldots \rightarrow p(k^*(p^*)) = p^* \\
\end{align*}
\]  

(E-7)

It is instructive to consider what this procedure implies in the simple case where there is no uncertainty and the demand curve is constant throughout the year. A sequence of iterations in this simple case is depicted in figure E-3. An initial choice of capacity, \( k^0 \), yields average and marginal cost curves of \( AC(x|k^0) \) and \( MC(x|k^0) \) respectively. The optimal price is as pictured where \( p^1 = MC(x|k^0) \). At the demand corresponding to \( p^1 \), the cost minimizing choice of capacity will be \( k^1 \), which yields a new marginal cost curve and hence \( p^2 \). This process can be repeated until one arrives at \( p^* \) where \( p^* = MC(x|k^*) \), which is the social optimum to the investment and pricing problem.

An analogous procedure is used when there is uncertainty and prices are chosen conditional on both \( k \) and the realization of the random variable
Illustration of a Simple Algorithm for Finding $P^*$ and $K^*$

\[ K^0 \rightarrow P^1(K^0) \rightarrow K^1(P^1) \rightarrow P^2(K^1) \rightarrow \ldots (P^*,K^*) \]
\( \eta \). In this case we can write \( P(k, \eta) \). In the case where only two scenarios are possible, \( \eta^1 \) and \( \eta^2 \) we can represent the process by

\[
k^0 \rightarrow \begin{cases} p^1(k^0, \eta^1) \\ p^2(k^1, \eta^2) \end{cases} \rightarrow \begin{cases} p^1(k^1, \eta^1) \\ p^2(k^1, \eta^2) \end{cases} \rightarrow \begin{cases} p^1(k^1, \eta^2) \\ p^2(k^1, \eta^2) \end{cases} \rightarrow \ldots \text{ (E-8)}
\]

Each scenario is now based on a different price, with the capacity at each iteration chosen by minimizing the expected cost of meeting a probability weighted average of all of these scenarios. We employed 100 scenarios rather than 2, but the principle is the same: for each realization of equations (E-5) or (E-6) were solved for the desired price, and the resulting load duration curves accumulated in our load distribution curve.

The programming and computational burdens of our simulations may be of interest to some readers. Our model was programmed in FORTRAN and consists of a series of subroutines which could be called in various sequences, and hence gave us the versatility to consider a number of different modeling frameworks. The major programs and subroutines, together with brief descriptions are listed in Table E-1. We emphasized programming flexibility rather than computational speed, and hence our subroutines and optimization algorithms are undoubtedly not the most efficient. The one important simplification which we did take advantage of for some of our runs was to use our linear load duration curve specification to simplify the calculation of prices. A fast cost simulation (SIMPLECOST) used areas of trapezoids to calculate total generating and shortage costs, which
### Table E-1

Programs and Subroutines

<table>
<thead>
<tr>
<th>name</th>
<th>purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAIN</td>
<td>main program - prompts user for key parameters, executes subroutines in prescribed order and writes out basic output</td>
</tr>
<tr>
<td>INITIATE</td>
<td>initiates all variables</td>
</tr>
<tr>
<td>LOADUR</td>
<td>generates a load duration curve conditional on a price, $P$, and a realization of the random vector $\mathbf{n}$</td>
</tr>
<tr>
<td>LOADIST</td>
<td>accumulates a given load duration curve into the load distribution curve</td>
</tr>
<tr>
<td>SRC</td>
<td>calculates short run costs; that is, the expected total costs of serving the given load duration or load distribution curve with the given capacity levels</td>
</tr>
<tr>
<td>LRC</td>
<td>calculates long run marginal costs; the same as SRC, only capacity is chosen so as to minimize total generating plus shortage costs for the given load duration or load distribution curve</td>
</tr>
<tr>
<td>MRC</td>
<td>calculates &quot;medium&quot; run costs; the same as LRC except only base plus peaking capacity is optimized. Base load capacity is taken to be at the prescribed level</td>
</tr>
<tr>
<td>SIMPLECOST</td>
<td>calculates short, medium or long run costs while taking advantage of the simplifications of a linear load duration curve</td>
</tr>
<tr>
<td>POPTIMIZE</td>
<td>calls an IMSL subroutine in order to find the optimal price. The three options for price setting are short run marginal costs, long run marginal costs, and average costs</td>
</tr>
<tr>
<td>SIMNORMAL</td>
<td>generates Monte Carlo drawings to be used for uncertainty simulations</td>
</tr>
<tr>
<td>OPTIONVALUE</td>
<td>approximates option values for a given load distribution curve and base load capacity choice. The approximation is only correct upon convergence</td>
</tr>
<tr>
<td>SURPLUS</td>
<td>calculates consumer surplus for a given price and realization of $\mathbf{n}$ using the linear load duration curve</td>
</tr>
<tr>
<td>WRITERESULTS</td>
<td>writes out simulation results</td>
</tr>
</tbody>
</table>
speeded computation by as much as a factor of five. It also avoided the endpoint problem mentioned above, which caused serious problems finding the optimal price in some cases where the discontinuity in derivatives was important.

Choosing optimal base load capacity for a probability weighted average of 100 future scenarios, with peaking capacity and prices chosen optimally in each scenario (i.e. the assumptions of Table 7) required 46 CPU seconds on a Honeywell 6880. Incorporating option values, which necessitated 2,500 random walks as well as revising capacity and pricing choices iteratively increased the computational burdens to 150 CPU seconds. These times are not excessive, and suggest that substantially greater complexity and realism can be added before the computational burdens become excessive.

We have discussed simulations based on 100 equally likely future scenarios, however similar results can be achieved using a smaller number of probability weighted scenarios. This would also facilitate the consideration of other dimensions of uncertainty (e.g. load factors, peak demand levels, forced outages, price elasticities, etc.).
Appendix F - Derivation of Excess Generation Cost Approximation of Option Values.

This appendix builds upon the results of chapter V, section B and appendix D in deriving approximations of 1) the option value foregone by building new capacity when capacity is durable, 2) the change in option values resulting from an increment to base load capacity, which is critical for 3) the optimal choice of base load capacity taking into account the fact that it is durable and irreversible.

As discussed in the main text, we are concerned here only with a single option value, the option to hold less base load capacity foregone when any additional capacity is added. We derive our approximations for this option value and the capacity choices dependent on it while ignoring the possibility that prices can be adjusted in response to suboptimal capacity choices. Throughout most of this appendix we will speak of excess generation costs, or more simply excess costs rather than option values. We discuss pricing effects and how our excess costs are related to option values towards the end of the appendix.

Suppose decisions must be made in period \( s \) about base load capacity for period \( t \), i.e. \( k_t \). Using the same notation as in the main text, let \( s^C_t(k_t, k_{t+1}, s) \) be expected capital, operating and shortage costs in period \( t \), as expected from the perspective of period \( s \), given capacity \( k_t \) for period \( t \), capacity \( k_{t-1} \) for period \( t-1 \), and the information available in period \( s \), \( n_s \). The discounted value of all costs in periods \( t \) and thereafter, \( s^V_t \), can be written as
\[ s_{V_t} = \sum_{i=0}^{\beta^i E} s_{t+1}C_{t+1}(k_{t+1}, k_{t+1-1}, \eta_{s+1}) \quad (F-1) \]

where

- \( k_{t-1} \) = the (exogenous) capacity already chosen for period \( t-1 \)
- \( \beta \) = constant discount factor \( (1/(1+r)) \)

Expanding this to allow us to look more closely at the first two terms, we can write

\[ s_{V_t} = s_{C_t}(k_t, k_{t-1}, \eta_s) \]
\[ + \beta E \left[ s_{t+1}C_{t+1}(k_{t+1}, k_t, \eta_{s+1}) \right] \quad (F-2) \]
\[ + \sum_{i=2}^{\beta^i E} \left[ s_{t+1}C_{t+1}(k_{t+1}, k_t, \eta_{s+1}) \right] \]

When capacity is chosen for period \( t \) the choices explicitly take into account their impact on period \( t \) costs, which are fully reflected by the first term in equation (F-2). When capacity is durable, choices also affect costs in subsequent periods through the constraint \( k_{t+1} > k_t \).

Let \( k_{i^{**}} \) be the optimal choice of capacity for period \( i \) when the choice is unconstrained by the previous period choice. Our goal is to derive an expression for the optimal choice of \( k_{i^{**}} \).

Focusing on the second term, which corresponds to expected costs in period \( t+1 \), there are only two possible partitions of states of the world which are of interest: either the constraint that \( k_{t+1} > k_t \) is binding, in which case we will choose \( k_{t+1} = k_t \), or else the constraint is not binding, in which case we will choose \( k_{t+1^{**}} \) in the optimal manner, with \( k_{t+1} > k_t \). Since we use a one dimensional representation of uncertainty, we can define a cut off function, \( \delta^{s+1}(k_t) \), such that if...
\( \eta_{s+1} > \delta^{s+1}(k_t) \) then additional capacity will be constructed, 
\( (k_{t+1}^{**} > k_t) \) while if \( \eta_{s+1} > \delta^{s+1}(k_t) \) then no new capacity will 
be constructed \( (k_{t+1} = k_t < k_{t+1}^{**}) \). Notice that the cut off 
function for \( \eta_{s+1} \) depends on \( k_t \); as more capacity is added the 
probability increases that the utility will find itself with excess 
capacity in the following period.

Taking expectations over the partitions of \( \eta_{s+1} \), we can expand 
equation (F-2) into

\[
V_t^{s'} = C_t^{s'}(k_t, k_{t-1}'^{s'}, \eta_s) \\
+ \beta E_{s+1}^{C_{t+1}}(k_t, k_t, \eta_{s+1}) \\
+ \beta E_{s+1}^{C_{t+1}}(k_{t+1}^{**}, k_t, \eta_{s+1}) \\
+ \sum_{i=2}^{\infty} \beta^i E_{s+1}^{C_{t+1}}(k_{t+i}, k_{t+i-1}, \eta_{s+1})
\]
which can be written as

\[ sV_t = sC_t(k_t, k_{t-1}, s) \]

\[ + \beta \mathbb{E} \left[ \sum_{s+1}^{s+1} \right] \left( k_t, s+1 \right) \]

\[ - s+1 C_{t+1}(k_{t+1}, s+1) \]

\[ + \beta \mathbb{E} \left( s+1 C_{t+1}(k_{t+1}, s+1) \right) \]

\[ \eta_{s+1} \delta_{s+1}(k_t) \]

\[ + \sum_{i=2}^{\infty} \beta^i \mathbb{E} \left( s+1 C_{t+i}(k_{t+i}, s+1) \right) \]

The first term remains the expected cost incurred in period \( t \). The second term is the expected increment in total expected costs caused by having too much capacity in period \( t+1 \). In appendix D we derived an exact expression for the term within square brackets which we develop further below. The third expression is for expected total costs where capacity is chosen optimally for all realizations of \( \eta_{s+1} \), regardless of whether this involves a positive or negative net investment. The final term is a summation of expected costs over all periods beyond period \( t+1 \).

So far we have focused on period \( s+1 \) choices for period \( t+1 \). If period \( s \) choices constrain period \( s+1 \) choices, then it is possible that they may also constrain period \( s+2 \) choices for period \( t+2 \). By expanding another term inside the summation sign, a new term similar to the second term in equation (F-3) can be derived reflecting the expected increment in period \( t+2 \) costs resulting from holding excess capacity:
\[ E \left[ s+2C_{t+2}(k_t, k_{t+1}^{**}, n_{s+1}) - s+2C_{t+2}(k_t^{**}, k_t, n_{s+1}) \right] \quad (F-4) \]

\[ (n_{s+1}^{<S+1}(k_t)) \cap (n_{s+2}^{<S+2}(k_t)) \]

Note that this increment to costs is only relevant when no additional capacity is added for period \( t+1 \), i.e. when \( n_{s+1}^{<S+1}(k_t) \). Whenever additional capacity is later added it is that subsequent investment which incurs future excess costs, not the previous period choices.

There are additional excess cost terms corresponding to more distant time periods. The discounted sum of these excess costs is similar to what we discuss in the main text in terms of option values. The option value foregone by building a given level of capacity \( k_t \) reflects the increased costs incurred in subsequent periods as a result of that \( k_t \) choice, and should be incorporated in the initial choice of how much generating capacity to construct.

Equation (F-3) can be used to derive an explicit expression for the optimal choice of capacity, \( k_t^{**} \), taking into account expected excess generating costs. If all of the terms such as the second term of equation (F-3) and expression (F-4) are gathered together into a single expression which we will abbreviate EC for excess costs. As we mentioned above, we call this "excess costs" rather than the option value because we have ignored pricing effects in their calculation. The first derivative of EC with respect to \( k_t \), \( \frac{\partial \text{EC}}{\partial k_t} \), can be taken, yielding

\[ c_1 + E\left[ \frac{\partial \text{EC}}{\partial k_t} \right] + m_1[1 - G(k_1; P, \eta)] \]

\[ = c_2 + m_2[1 - G(k_1; P, \eta)] \quad (F-5) \]
Restrictions on C are needed to guarantee the second order conditions are satisfied, however we ignore those restrictions here. The challenge is to make the solution of this equation feasible.

In appendix D we derived the result that excess generating costs from holding base load capacity \( k_1 \) instead of \( k_1^{**} \) can be written as

\[
EC = C(k_1, \eta) - C(k_1^{*}, \eta) = (m_2 - m_1) \int_{k_1^*}^{k_1} (x-k_1) h(x|\eta) \, dx \tag{F-6}
\]

where

\[
\begin{align*}
  m_1 &= \text{base load operating cost} \\
  m_2 &= \text{peaking operating cost} \\
  h(x|\eta) &= \text{probability density function for demand, conditional on } \eta
\end{align*}
\]

This expression reflects the fact that there are operating cost savings from having the excess capacity which partially offset the excess capital costs. The cost is represented diagrammatically in figure F-1, where the cost of holding too much base load capacity is \((m_2 - m_1)\) times area A.

In order to simplify the calculation of optimal investment choices using equations (F-5) and (F-6) several assumptions and approximations were used. As we have previously noted, all costs were held fixed in real terms. This implies that the switch over points (expressed as fractions of the year) between different generating technologies remain constant over time. Our multiplicative growth and uncertainty specification means that the load distribution curve is shifting up and down equiproportionately (and randomly) over time. Together with the constant cost assumption, this means that the optimal choice of generating capacity is simply proportional to \( \eta \). This result is shown analytically in figure (F-2), where \( k_t^{**} \) and \( k_{t+1}^{**} \) are proportional to demand.
Figure F-1

Costs of Excess Capacity and Approximation Used for Simulations

**Exact Expression**

\[ EC = (m_2 - m_1) \int_{K_*}^{K_1} (x - K_1) h(x) dx \]

\[ \frac{\partial EC}{\partial K_1} = (m_2 - m_1) \{H(K_1) - H(K_*)\} \]

**Approximation Used**

\[ EC = (m_2 - m_1) \frac{(K_1 - K_*)^2}{2 h(K_*)} \]

\[ \frac{\partial EC}{\partial K_1} = (m_2 - m_1) \frac{(K_1 - K_*)}{h(K_*)} \]
Figure F-2

Proportionality of Optimal Base Load Capacity to $\eta$

Fraction of the year that demand is below level $x$
Taking derivatives of (F-5) with respect to \( k_t \), it is easy to see that the derivative of the excess costs incurred as a result of too much capacity being held can be written as

\[
\frac{\partial EC}{\partial k_t} = (m_2 - m_1) \left( H(k_1 | \eta) - H(k_1^{**} | \eta) \right) \tag{F-7}
\]

In figure F-1 this derivative can be seen to be the horizontal, top side of the nearly triangular region.

In order to simplify the calculation of both (F-6) and (F-7) we used a linear approximation of \( h(k_1 | \eta) \) around \( k_1^{**} \) and approximated \( EC \) and \( \frac{\partial EC}{\partial k_t} \) via

\[
EC = (m_2 - m_1) \frac{(k_1 - k_1^{**})^2}{2 \ h(k_1^{**} | \eta)} \tag{F-8}
\]

and

\[
\frac{\partial EC}{\partial k_t} = (m_2 - m_1) \frac{(k_1 - k_1^{**})}{h(k_1^{**} | \eta)} \tag{F-9}
\]

Approximation (F-8) is depicted diagrammatically by the area of the triangular region, bounded from below by the dashed line in figure F-1, while approximation (F-9) is the top side of that triangular region. This approximation will tend to understate the true excess costs, and its derivatives if the load distribution curve \( H(x) \) is convex in the neighborhood of \( k_t^{**} \), as we expect to be true empirically.

In order to choose \( k_t \) optimally we need to know how \( k_{t+1}^{**} \) will subsequently be chosen. In line with our constant cost assumption which
imposes a sort of stationarity over time, we use the approximation that the change in the excess costs in period s due to a capacity addition will be close to the change in excess costs in period s+1. This is actually a slight overestimate of future period excess costs, and hence this biases our approximation of option values upward slightly. Using the previously stated property that $k_t^{**}$ and $h(k_t^{**})$ will be proportional and inversely proportional, respectively to $n$, and normalizing $n_s$ to be one, we can write:

$$k_{t+1}^{**} = k_t^{**} n_{s+1} \quad \text{(F-10)}$$

and

$$h(k_{t+1}^{**}) = h(k_t^{**})/n_{s+1} \quad \text{(F-11)}$$

whenever $n_{s+1} < n_s$. Substituting these relationships into (F-8) yields

$$EC_{t+1} = \frac{k_t^{**} (1 - n_{s+1})^2 n_{s+1}}{2 h(k_t^{**})} \quad \text{(F-12)}$$

$$\frac{\partial EC}{\partial k_t} = \frac{k_t^{**} (1 - n_{s+1}) n_{s+1}}{h(k_t^{**})} \quad \text{(F-13)}$$

These are the actual formulas used to calculate the excess costs for our simulations.

In order to implement these results some method of taking expectations is needed while determining the cut off points for $n$ below which no additional capacity will be added. With our simple demand specification and our constant cost assumptions this problem is vastly simplified: at the optimum additional generating capacity will be added whenever demand has grown between s and s+1, i.e. whenever $n_{s+1} > n_s$. Note that this is an equilibrium condition only: if there is too much capacity, then no new
capacity should be added, even if \( \eta_{s+1} > \eta_s \). The opposite will be true when there is too little. Since we use this equilibrium condition to calculate our excess costs, our approximation will only be correct at the optimum.

The first step we used to calculate the expected excess costs was to find the base load capacity choice \( k_t^* \) based on static investment rules using the fifth framework. 2,500 random walks into the future were then made. For each random walk a drawing of \( \varepsilon_{s+1} \) was made from the lognormal density function corresponding to the exogenously specified growth rate and degree of uncertainty (these were parameterized in terms of the mean and variance of the normal distribution underlying the lognormal distribution.)

If \( \varepsilon_{s+1} = \eta_{s+1} < 1 \), so that expected future demand has declined, then EC and \( \frac{\partial \text{EC}}{\partial k_t} \) were calculated using (F-12) and (F-13), discounted by a constant discount factor \( \beta \) and added on to the total and marginal option values, respectively. Another random drawing, \( \varepsilon_{s+2} \), was made and if \( \eta_{s+2} = \varepsilon_{s+2} \eta_{s+1} < 1 \)

(i.e. \( \eta_{s+2} < \eta_s \) \( \land \) \( \eta_{s+1} < \eta_s \)) then EC and \( \frac{\partial \text{EC}}{\partial k_t} \) were again added on to the total and marginal excess cost calculations.

This random walk was continued until either \( \eta_{s+1} > 1 \), or \( i=30 \). Truncating the walk after 30 years was intended to approximate the 30 year life time of real world generating capacity. After completion of 2,500 random walks, the derivative of EC was used to solve for a new capacity choice using equation (F-5). Since our approximation of the excess costs and its derivative are only correct at the optimal choice of \( k_t^{**} \), we iterated between the choice of \( k_t \) and the determination of the excess cost until convergence, which took five iterations or less in most cases.

Up until now we have been ignoring prices in our discussion. If
prices change in response to capacity choices which are not optimally chosen, then in addition to the excess generation costs (which will be reflected in profits) there will also be changes in consumer surplus. Since prices would not be adjusted unless social welfare could be improved we know that our excess cost calculations are greater than the true option values or burdens on society as a whole from having excess generating capacity.

We did not adjust prices in response to capacity choices which are suboptimal because of unexpectedly slow demand growth. This avoided having to recalculate prices and hence the load distribution curve whenever capacity choices were constrained. Instead prices were based on the optimal choices of capacity, $k_t^{**}$, even for realizations of $\eta$ for which choices of generating capacity are in fact constrained, i.e.

$\eta_{s+i} < \eta_{s+i-1}$. The excess costs correspond to these suboptimal prices rather than the socially optimal prices.

An alternative means of calculating option values would have been to take random walks forward while simulating the model with different investment rules (e.g. different marginal option values) and then choosing the investment rule which resulted in the largest expected consumer plus producer surplus. We rejected this approach because dimensionality constraints would soon rule out this as a possible approach for any more realistic and complex model. The great advantage of our approximating technique is that it does not require resimulating the model over multiple random futures. By using a local approximation of the load distribution curve and developing a simple cut off function $\delta^{s+1}(k_t)$ we vastly simplify the procedure. Even with changing operating and capital cost over time, fairly simple "rule of thumb" type cut off functions may be available
which eliminate the necessity of excessive forward simulations of the model. It remains to be seen whether this procedure will be useful in more complex modeling frameworks.
Bibliography


Biographical Note

Randall Poor Ellis was born on April 13, 1954 in Newton, Massachusetts and spent his childhood in Wellesley, Massachusetts. After graduating from Yale University in 1976 (Summa cum Laude, Phi Beta Kappa and recipient of the William H. Massee award in economics) he attended the London School of Economics for one year during which time he earned a M.S. degree in economics. While working towards his Ph.D. at M.I.T. his fields of primary concentration were Industrial Organization and Econometrics, his secondary fields Fiscal Economics and Economic History.

Randy enjoys skiing, tennis, sailing, water skiing, and golf, and while at M.I.T. he participated at the intramural level in basketball, volleyball, squash, table tennis, softball, sailing and at the intercollegiate level in Ultimate Frisbee.

Randy has accepted a position as assistant professor with the Economics Department at Boston University starting in September, 1981.