Issues in Public Utility Regulation

edited by
Harry M. Trebing
Michigan State University

Proceedings of the Institute of Public Utilities
Tenth Annual Conference

1979
MSU Public Utilities Papers

Division of Research
Graduate School of Business Administration
Michigan State University
East Lansing
Contents

Part Two
Evaluating New Pricing Practices in Electricity

Experimental Time-of-Use Electricity Prices
A.K. Miedema, S. B. White, C.A. Clayton, and D.P. Lifgren 51

A Benefit-Cost Framework for Evaluating New Pricing Concepts and Rate Structure Reform
Robert L. Borlick 73

Load Research Concerns in Rate Structure Revision
Samuel Behrends, Jr. 90

Comments
William E. Miller 96
Comments
Craig R. Johnson 99
Comments
Rodney E. Stevenson 104

Part Three
Network Access Pricing in Telecommunications

The Theory of Network Access Pricing
Robert D. Willig 109

Network Access Pricing
Laurence Garfinkel 153

Public Policy Issues of Network Access Pricing
Andrew Margeson 167

Comments
Ronald R. Braeutigam 176
Comments
James R. Nelson 180
Comments
William H. Melody 187

Part Four
Costing Methodologies for Time Differentiated Rates

Time-of-Use Rates: Marginal Costing and the System Planner
Leo T. Mahoney, Jr. 197

Windmills of Costing Methods — 3: Poor Don (Rate Managers) — 0: Sancho Hasn’t Helped Much
William J. Leininger 206

Another Approach to Time-of-Use Rates
Alan Schoor 221

Time-of-Day Rates: Past, Present, and Future
J.R. Crespo 244

Comments
Michael L. Tennican 252
Comments
William H. Melody 257
Comments
Robert S. Gay 263

Part Five
The Impact of New Technology on Public Utility Industries

Diffusion of Central Station Solar Power
David A. Huetten 269

Robert D. Stoner 291

Obsolescence, Depreciation, and New Technology
F.W. Sindel 304

Comments
John E. Cummings 317
Comments
William G. Shepherd 321
Part Six
Rate of Return and Regulatory Oversight Models

Risk Measurement and Rate of Return under Regulation
Ronald W. Meticher 325

The Cost of Equity Capital:
A Model for Regulatory Review
Basil L. Copeland 342

Economic and Financial Models as Regulatory Tools
Robert E. Wayland 367

Comments
John L. O'Donnell 373

Comments
Joseph M. Quigley 378

Comments
Clement T. Loshing 381

Part Seven
New Issues in Telecommunications Regulation

The Impact of Interconnect Rate Restructure on
Local Telephone Rates
Alan C. Hasselwander 391

The Continuing Role of Federal Regulation in the
Transition to Competition in Communications
Walter G. Bolter 401

Competition and Regulation in
Telecommunication Services
Katherine E. Sasseville 418

Comments
Leland L. Johnson 430

Comments
Jeffrey Rohlf 436

Part Eight
New Issues in Gas Regulation

Incremental Pricing and the National Energy Act
John E. Holtzinger, Jr. 445

Financing Large-Scale Gas Supply Projects
John F. Curley, Jr. 460

Public Policy Considerations in the
Pricing of New Gas Supplies
David S. Schwartz 468

Comments
Richard A. Solomon 487

Comments
Richard A. Rosen 493

Part Nine
Introducing Incentives and Efficiency Criteria into Regulation

Integrating Efficiency Standards into the
Regulatory Process
Thomas K. Standish 501

Incentive Regulation
Daniel J. Demlow 531

Comments
Sherwood H. Smith, Jr. 548

Comments
Robert J. Rohr 556
List of Contributors

Stephen C. Ailes
Price Analyst, Office of Price Monitoring
Economic Regulatory Administration
U.S. Department of Energy

Samuel Behrends, Jr.
Vice-President
Corporate Regulatory Policy
Carolina Power & Light Co.

Walter G. Bolter
Chief, Economics Division
Common Carrier Bureau
Federal Communications Commission

Robert Borlick
Chief, Economic and Financial Analysis Branch
Division of Regulatory Proceedings
Economic Regulatory Administration
U.S. Department of Energy

Ronald Braeutigam
Professor of Economics
California Institute of Technology

C. Andrew Clayton
Senior Statistician
Research Triangle Institute

Basil L. Copeland, Jr.
Chief Economist
Division of Energy Conservation and Rate Advocacy
Arkansas Office of the Attorney General

J. R. Crespo
President and Chief Executive Officer
Ebasco Business Consulting Co.

John Cummings
Manager
Solar Energy Program
Electric Power Research Institute

John F. Curley
Vice-President
Morgan Stanley & Co.

Daniel J. Demlow
Chairman
Michigan Public Service Commission

Lawrence Garfinkel
Director
Rate & Tariff Planning
American Telephone and Telegraph Co.

Robert S. Gay
Executive Manager
Rates & Regulation
Virginia Electric Power Co.

Alan C. Hasselwander
Vice-President
Network Services Planning
Rochester Telephone Corporation
Contributors

John E. Holtzinger, Jr.
Morgan, Lewis & Bockius
Washington, D.C.

David A. Huettner
Professor of Economics and
Director, Energy Program Center for
Economic and Management Research
University of Oklahoma

David A. Irwin
Chief, Program Evaluation Staff
Acting Chief, Policy and Rules Division
Common Carrier Bureau
Federal Communications Commission

Craig R. Johnson
Chief, Regulatory Economics and Standards
Office of Utility Systems
Economic Regulatory Administration
U.S. Department of Energy

Leland L. Johnson
Associate Administrator for
Policy Analysis and Development
National Telecommunications and Information Amin.

William J. Leininger
Principal
Ernst & Ernst

Dale P. Lifson
Consultant
Research Triangle Institute

Clement T. Loshing
Treasurer
Cleveland Electric Illuminating Co.

Leo T. Mahoney, Jr.
Vice-President

Andrew Margeson
Corporate Planning Director
Rochester Telephone Corporation

Degna Marmo
Attorney, Office of Price Monitoring
Economic Regulatory Administration
U.S. Department of Energy

Ronald W. Melicher
Professor of Finance
University of Colorado

William H. Melody
Professor and Chairman
Department of Communications
Simon Fraser University

Allen K. Mederna
Manager
Economics Department
Research Triangle Institute

William E. Miller
Manager
Economic Research Department
Baltimore Gas and Electric Co.

James R. Nelson
Professor of Economics
Amherst College

John L. O'Donnell
Professor of Finance
Michigan State University
<table>
<thead>
<tr>
<th>Name</th>
<th>Title and Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Howard F. Perry</td>
<td>Director of Regulatory Assistance, Office of Utility Systems, Economic Regulatory Administration, U.S. Department of Energy</td>
</tr>
<tr>
<td>Joseph M. Quigley</td>
<td>Financial Vice-President and Secretary, Northern Illinois Gas Co.</td>
</tr>
<tr>
<td>Jeffrey H. Rohlfis</td>
<td>Member, Technical Staff, Bell Laboratories</td>
</tr>
<tr>
<td>Robert J. Rohr</td>
<td>Professor of Economics, Brown University</td>
</tr>
<tr>
<td>Richard A. Rosan</td>
<td>Former Sr. Vice-President and Counsel, Columbia Gas System Service Corporation</td>
</tr>
<tr>
<td>Katherine E. Sasseville</td>
<td>Commissioner, Minnesota Public Service Commission</td>
</tr>
<tr>
<td>Alan Schoor</td>
<td>Director of Regulated Industries, Gordian Associates, Inc.</td>
</tr>
<tr>
<td>David S. Schwartz</td>
<td>Public Interest Consultant and Regulatory Economist, Washington, D.C.</td>
</tr>
<tr>
<td>William G. Shepherd</td>
<td>Professor of Economics, University of Michigan</td>
</tr>
<tr>
<td>Frank W. Sinden</td>
<td>Senior Research Scientist, Center for Environmental Studies, Princeton University, and Member, Technical Staff, Bell Laboratories</td>
</tr>
<tr>
<td>Sherwood H. Smith, Jr.</td>
<td>President, Carolina Power &amp; Light Co.</td>
</tr>
<tr>
<td>Richard A. Solomon</td>
<td>Attorney at Law, Wilmer &amp; Schreiner, Washington, D.C.</td>
</tr>
<tr>
<td>Thomas K. Standish</td>
<td>Commissioner, Connecticut Public Utilities Control Authority</td>
</tr>
<tr>
<td>Rodney E. Stevenson</td>
<td>Professor of Public Utilities, University of Wisconsin - Madison</td>
</tr>
<tr>
<td>Robert D. Stoner</td>
<td>Economist, Bureau of Economics, Federal Trade Commission</td>
</tr>
<tr>
<td>Michael L. Tennican</td>
<td>Vice-President and Director, Energy &amp; Environment Studies, Temple, Barker &amp; Sloane, Inc.</td>
</tr>
<tr>
<td>Maurice Van Nostrand</td>
<td>Chairman, Iowa State Commerce Commission</td>
</tr>
<tr>
<td>Robert E. Wayland</td>
<td>Economist, Temple, Barker &amp; Sloane, Inc.</td>
</tr>
</tbody>
</table>
Preface

The rapid proliferation of issues confronting the public utility industries makes it more and more difficult to devote a conference or a collection of papers to a specific topic. This volume, divided into nine parts, treats a range of major problems in electricity, gas, and telecommunications.

Passage of the National Energy Act in 1978 introduced significant changes in the nature and degree of federal participation in public utility regulation. In addition, federal involvement has been affected by White House efforts to control inflation. Part One of the collection considers the changing pattern of federal-state relations in energy regulation. Degan Marmo and Stephen Ailes examine federal antinflation guidelines as applied to public utilities. Howard Perry reviews federal initiatives with respect to electric utility rate design under the new legislation, and Maurice Van Nostrand expresses a state regulator's concerns over growing federal involvement. Rodney Stevenson interprets the changing pattern of federal energy regulation in terms of an emergent philosophy of neofederalism. Charles Zielinski and Mark Widoff comment, respectively, from the perspective of a state regulator and a consumer advocate.

Part Two evaluates three aspects of the introduction of new pricing practices in electricity supply. Allen Miedema and his colleagues examine the effect of experimental time-of-use rates on selected residential markets. Robert Borlick explores the maze of problems associated with the development of a benefit-cost framework for evaluating rate structure reform, and Samuel Behrends calls attention to the need for a wider analytical framework and more empirical
evidence in considering changes in rate design. The discussants provide three different perspectives on these papers.

Part Three focuses on one of the crucial problems stemming from the growth of competitive entry in telecommunications, notably the price for access to the local exchange or common carrier network. The theoretical, applied, and public policy aspects of network access pricing are considered by Robert Willig, Lawrence Garfinkel, and Andrew Margeson. Ronald Braeutigam, James R. Nelson, and William Melody offer comments which, in turn, reflect their perception of the theoretical and institutional issues involved in network access pricing under conditions of structural change.

Part Four contains four contributions that examine costing methodologies for time-differentiated rates. The papers by Leo Mahoney, William Leininger, and Alan Schoor discuss different approaches for measuring the marginal cost of electricity supply in peak and off-peak periods. J. R. Crespo, in contrast, appears more sympathetic to fully allocated accounting costs as the appropriate criterion. Three discussants provide a critique of the methodologies proposed as well as a commentary on the controversy over the relative merits of marginal and fully allocated cost as a guideline. Avid followers of the debate over costing methodologies will note that Charles Cicchetti did not participate in this collection; however, the editor hastens to note that he does contribute to the discussion in a paper presented at the Institute’s 1979 conference, to be published in the forthcoming *Electricity Law and Economics*.

In Part Five, David Huebner and Robert Stoner consider new technology in terms of the diffusion or spread of innovations, and Frank Sinden analyzes the interdependence among technological change, depreciation rates, and future uncertainty.

Part Six turns to the traditional topics of rate of return, and Ronald Melcher and Basil Copeland demonstrate how new analytical concepts could be applied. Particular emphasis is placed on the capital asset pricing model and the discounted cash flow model. The applicability of forecasting, financing, and operations research models to the regulatory process is assessed by Robert Wayland. As before, the discussants fulfill their role as commentators and critics.

In Part Seven, the papers focus on the interrelationship between growing competition in telecommunications, on the one hand, and the price of local exchange service, the future role of the FCC, and the rationale for state regulation, on the other. Leland Johnson and Jeffrey Rohlfs provide alternative policy perspectives on the papers by Alan Hasselwander, Walter Bolter, and Katherine Sasseville.

In Part Eight, John Holtzinger, John Curley, and David Schwartz consider the difficulties arising from growing national reliance on high cost sources of new gas. Holtzinger examines the problems associated with the incremental pricing of such supplies under the Natural Gas Policy Act, while Curley and Schwartz take markedly different positions regarding the appropriate course for public policy. The discussants’ comments further emphasize the schism that exists at the public policy level.

Part Nine deals with the perennial question of introducing incentives and efficiency criteria into regulation. Thomas Stimson evaluates the shortcomings of achieving Pareto optimality through marginal cost pricing, and he concludes that more direct regulatory intervention in the operation of the firm is required. Daniel Demlow examines the problem of implementing a system of incentive regulation, giving special attention to the Michigan Plan, which was introduced under his chairmanship of the Michigan Public Service Commission. The discussants provide the appropriate critical evaluation.

The contributions contained in this collection are essentially the final versions of the papers and comments originally presented at the Institute’s 1978 Williamsburg conference. The Marmo-Ailes paper was added subsequently. Participants were selected from government, industry, the academic community, and consulting and law firms. The academic purist may be troubled at times by an apparent lack of methodological rigor in treating particular topics. However, it is important to remind such readers that one of the objectives of the collection is to provide insight into the major problems confronting the public utility industries and regulation. In many cases, these are highly contentious areas in which the only source of information is the practitioner dealing with applied problems. Public policy decisions can seldom await the completion of a generally accepted body of positive analysis. Furthermore, the work of the practitioner provides the reader with an appreciation of the current state of the art.

Special recognition should be given to the contributions of Mrs. Virginia Michels and Ms. Elizabeth Johnston in producing this volume. Mrs. Michels played a pivotal role in carrying out the administrative duties associated with the Institute’s 1978 conference. Ms. Johnston took responsibility for preparing the papers for printing and for overseeing the publication process. In carrying out this task, she was occasionally called upon to perform feats of literary legerdemain in enhancing the readability of the final product.

Recognition should also be given to the members of the Program Planning Committee, who provided considerable assistance in or-

Harry M. Trebing

Part One

Changing Patterns in Energy Regulation
Applying Anti-Inflation Guidelines to the Electric and Gas Utilities

Degna L. Marmo and Stephen C. Ailes

The objective of President Carter's anti-inflation program is to break the inflationary spiral of wage and price increases. The price standards were designed to cover a wide range of diverse economic activities and yet retain the simplicity necessary for practicability. However, the institutional characteristics and operational realities of certain industries have mandated modifications of the standards to avoid the imposition of gross inequities. Electric and gas utilities, as energy dependent regulated industries, pose special problems.

The Alternative Standard

The original Price Deceleration Standard provides each company with a numerical limitation on price increases during the program year. This limitation is derived by deducting one-half of a percentage point from the average annual rate of price increase over the 1976–1977 period. For example, if a company's prices increased at a rate of 6 percent per year during the base period, then its limitation for price increases during the program year would be 5.5 percent. To take into account the possibility of abnormally low or high base period rates of price change, the standard includes upper and lower bounds on program year rates of price change. Regardless of the base period rate of
price change, a program year rate above 9.5 percent does not comply with the price standard. Similarly, as long as price increases over the program year average 1.5 percent or less, a company will be in compliance with the price standard regardless of the base period rate of increase.

Due to the energy dependent nature of electric and gas utilities, the general Price Deceleration Standard, applied to most other industries, is of limited applicability to these utilities. Sharp rises in fuel and purchased gas cost, which constitute more than 40 percent of total operation expenses for the utility industry, as well as significant variations in the rate of fuel cost increases during 1978, have made it difficult for many utilities to achieve price deceleration. Thus, the Council on Wage and Price Stability ceased to offer an alternative. Any electric or gas utility may now use either the Price Deceleration or the Gross Margin Standard, previously available only to food manufacturers and petroleum refiners, to determine compliance.

Under the Gross Margin Standard, a utility may pass through increases in costs of fuel and purchased gas and power on a dollar-for-dollar basis; however, the nonfuel costs for the utility should not be increased in excess of 6.5 percent during the first program year (13.5 percent during the second program year), plus any percentage growth in physical volume. An electric or gas utility satisfies the Gross Margin Standard if the rate of increase in its gross margin (operating revenues less cost of fuels, purchased gas, or purchased power) between the base quarter (the last complete calendar or fiscal quarter ending prior to October 2, 1978) and the corresponding quarter of 1979 does not exceed 6.5 percent plus any positive percentage growth in physical volume over the same period.

Electric and gas utilities that cannot comply with either the Price Deceleration or the Gross Margin Standard should adhere to the Profit Margin Limitation Exception (Section 765A-6 of the standards).

Recognizing that the prices of most electric and gas utilities are already subject to regulation by state public utility commissions or the Federal Energy Regulatory Commission, the council asks, under the Alternative Price Standard for Electric and Gas Utilities, that these commissions determine compliance with the standards and rule on the applicability of exceptions. Commissions can administer the Profit Margin Limitation Exception (Section 765A-6 of the standards) and exceptions for extreme hardships and gross inequities. Commissions may find it necessary to grant exceptions in order to enable utilities to raise capital to finance construction that is needed either to serve customers or for construction of new facilities to meet federal policies that seek to reduce dependence on oil. Circumstances that the council believes could impose a hardship for utilities committed to such investments include the following: (1) Pro forma interest coverage ratios are inadequate to assure the legal minimum of bond indentures; (2) the market-to-book ratio of outstanding common stock is below minimum levels judged by the commission to be acceptable, a level which in the council's judgment would not, under present circumstances, exceed 0.9 for purposes of granting a hardship exception, but not for other purposes; and (3) cash flows are judged to be inadequate. If increases under these exceptions exceed 9.5 percent, commissions are asked to implement phasing schedules that limit the initial rate increase, to ease the inflationary impact; when such action is permitted by law, it is in the long run interest of consumers and would be reasonable in light of seasonal rate patterns and other relevant factors.

Although the responsibility for demonstrating compliance rests with the utilities, commissions are asked to report to the council on a quarterly basis the extent to which they have succeeded in implementing the anti-inflation standards. A commission that grants an increase in a major rate case that yields an increase in the gross margin exceeding 6.5 percent (or 13.5 percent in the second program year), plus a positive percentage growth in physical volume, is asked to detail the justification in its final order. Likewise, a commission that grants exceptions resulting in rate increases higher than would be permissible under the appropriate Price Deceleration Standard or Gross Margin Standard, or a commission that grants a Profit Margin Limitation Exception, is requested to explain its action in the final decision. Table 1 reports on rate increases, compliance, and exceptions granted in the first nine months of the program.

Since utilities that provide service to more than one jurisdiction maintain separate accounting records for each jurisdiction, the council asks that each commission evaluate compliance within its own jurisdiction. The standards are intended to apply separately to each utility's operations in each jurisdiction. Electric and gas utilities are requested to submit price or margin data only on their nonutility operations, since commissions report to the council on utility compliance.

First-Year Problems

Although the Alternative Standard provides a means of compensating for the sharp rise in fuel and gas costs and regional variations in
these uncontrollable increases, several problems unique to utilities remain. Seasonal variations, inherent in electric and gas utility operation, produce inequities in conjunction with the use of the third quarter of 1978 as the gross margin base quarter. This inconsistency has been remedied in the second-year standards by changing from a base quarter to a base year calculation for the Gross Margin Standard.

The utilities have alerted the council to other problems which are not addressed by the Gross Margin Standard. In each of these cases, the council feels that the profit-margin limitation, undue hardship, and gross inequity exceptions available to commissions under the Alternative Standard are flexible enough to cover circumstances of genuine need and yet allow the commissions to make decisions based on the public interest.

The first of these problems peculiar to the utility industry is escalating capital costs, particularly of new construction to reduce national dependence on imported oil in addition to the normal capital costs the industry would incur. The first program year Alternative Standard explicitly cites the need to raise capital to finance construction of new facilities to meet federal policies that seek to reduce dependence on oil as a circumstance that could warrant an exception to the standards.

A related problem concerns the ever-increasing amounts of construction work in progress (CWIP) due to the rapid rise in the cost of new capacity upon which, generally, no cash return is being earned. Many commissions are now allowing CWIP in the rate base to finance necessary construction, and under the Gross Margin Standard the inclusion of CWIP in the rate base is counted against the 6.5 percent allowable growth in the gross margin. A similar problem arises when a new plant comes on line during the program year and is thus included in the rate base in that year. Since the allowance for funds used during construction (AFUDC) is not presently included in the definition of operating revenues according to standard accounting principles, generally a plant coming on line would not merely be handled by a conversion from noncash to cash revenues from the base period to the program year. This could represent an exaggerated increase in the gross margin in the program year, since the costs of the plant would then be included in the rate base producing revenues. Again, the exception procedures should adequately cover such a case. Since any plant coming on line in the program year could put a utility out of compliance with the Gross Margin Standard, the commissions are at liberty to find a gross inequity exception. The reasoning is that new plants are built to provide service for all customers within the jurisdic-

<table>
<thead>
<tr>
<th>Table 1. Nine-Month Report on Rate Increases, Compliance, and Exceptions Granted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of rate increases granted</td>
</tr>
<tr>
<td>---------------------------------</td>
</tr>
<tr>
<td>Alabama</td>
</tr>
<tr>
<td>Alaska</td>
</tr>
<tr>
<td>Arizona</td>
</tr>
<tr>
<td>Arkansas</td>
</tr>
<tr>
<td>California</td>
</tr>
<tr>
<td>Colorado</td>
</tr>
<tr>
<td>Connecticut</td>
</tr>
<tr>
<td>Delaware</td>
</tr>
<tr>
<td>District of Columbia</td>
</tr>
<tr>
<td>Florida</td>
</tr>
<tr>
<td>Georgia</td>
</tr>
<tr>
<td>Hawaii</td>
</tr>
<tr>
<td>Idaho</td>
</tr>
<tr>
<td>Illinois</td>
</tr>
<tr>
<td>Indiana</td>
</tr>
<tr>
<td>Iowa</td>
</tr>
<tr>
<td>Kansas</td>
</tr>
<tr>
<td>Kentucky</td>
</tr>
<tr>
<td>Louisiana</td>
</tr>
<tr>
<td>Maine</td>
</tr>
<tr>
<td>Maryland</td>
</tr>
<tr>
<td>Massachusetts</td>
</tr>
<tr>
<td>Michigan</td>
</tr>
<tr>
<td>Minnesota</td>
</tr>
<tr>
<td>Mississippi</td>
</tr>
<tr>
<td>Missouri</td>
</tr>
<tr>
<td>Montana</td>
</tr>
<tr>
<td>Nebraska</td>
</tr>
<tr>
<td>Nevada</td>
</tr>
<tr>
<td>New Hampshire</td>
</tr>
<tr>
<td>New Jersey</td>
</tr>
<tr>
<td>New Mexico</td>
</tr>
<tr>
<td>New York</td>
</tr>
<tr>
<td>North Carolina</td>
</tr>
<tr>
<td>North Dakota</td>
</tr>
<tr>
<td>Ohio</td>
</tr>
<tr>
<td>Oklahoma</td>
</tr>
<tr>
<td>Oregon</td>
</tr>
<tr>
<td>Pennsylvania</td>
</tr>
<tr>
<td>Rhode Island</td>
</tr>
<tr>
<td>South Carolina</td>
</tr>
<tr>
<td>South Dakota</td>
</tr>
<tr>
<td>Tennessee</td>
</tr>
<tr>
<td>Texas</td>
</tr>
<tr>
<td>Utah</td>
</tr>
<tr>
<td>Vermont</td>
</tr>
<tr>
<td>Virginia</td>
</tr>
<tr>
<td>Washington</td>
</tr>
<tr>
<td>West Virginia</td>
</tr>
<tr>
<td>Wisconsin</td>
</tr>
<tr>
<td>Wyoming</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>
tion which the utility is obligated to serve, to reduce dependence on foreign oil, or to accommodate other federal policies.

Comments on Exception Criteria

The utility industry has also commented on the three circumstances that the council suggested in the Alternative Standard could impose a hardship on companies committed to investments to finance construction of new facilities to meet federal policies or to provide service to customers. Generally, pro forma interest coverage ratios are seen by both the council and the industry as a valid indication of financial hardship. indentures and charters of most utility companies limit the amount of noncash income allowable for coverage ratio purposes. The interest coverage ratio is basically the pretax operating income before interest charges, divided by the total interest charges on debt. The council coverage standard of 2.0 is generally accepted as being a level indicative of financial stress. The only criticism voiced has been that coverage, while a useful guide, should be considered in conjunction with other financial indicators.

Inadequate cash flow for utilities is also judged by the council and the industry as a possible indication of financial problems warranting an undue hardship exception. Cash flow is defined as net earnings, plus depreciation, plus deferred taxes. Cash flow ratios are important for correctly assessing the financial situation of a company, since this is how investors determine whether a firm will be able to pay off its debt. The first important measurement is cash flow as a percentage of long-term debt, which affects a utility's industrial credit rating. The second is the ratio of net cash flow (funds from operations minus cash dividends) to capital expenditure. This ratio indicates to investors the company's need to finance externally, its ability to finance internally, and the outlook for its financial liquidity. If a company has a low ratio, and the trend continues to deteriorate, this is a red flag to investors that a credit rating may be changed. If a company has to borrow so much that its key financial ratios will be diluted, this is an advance indication to investors of potential deterioration of a utility's credit rating. In essence, the value of a company is measured by the current market value of all outstanding claims on the company's present and future cash flow.

One circumstance has received both criticism and support from the electric and gas utility industry as well as consumer advocates. This is the situation in which the market-to-book ratio of outstanding common stock is below minimum levels judged by the commission to be acceptable, a level in the council's judgment which should not exceed 0.9 for purposes of granting a hardship exception. Consumer advocates have commented that this ratio is of limited value as a spot or even short-run indicator, since the average market-to-book ratio for electric utilities, for example, is below the council standard of 0.9.

In addition, all utility stocks follow the market, and the drop in market-to-book ratios has been unrelated to regulatory decisions. Book value per common share is merely the total equity component of the capitalization divided by the number of common shares outstanding. Industry representatives have argued that utilities issuing large amounts of new common stock in the marketplace at prices below book value will experience a dilution in their equity or asset per share after each sale. The logical outcome of repeated sales of new stock at discounts to book value would be financial collapse. Before reaching this point, a utility would halt the sale of stock, bringing all types of financing to a halt. Thus, additions to plant capacity and servicing of prior debt would end. In the final analysis, the utility representatives state that the marketplace, not the council or the commissions, determines where a utility's stock sells in relation to its book value and that the 0.9 ratio suggested by the council is meaningless. Conversely, some utility representatives have argued in favor of the 0.9 level, stating that, although any figure below 1.0 is potentially confiscatory, the reality of the matter is that under present economic conditions most electric and gas utilities have market-to-book ratios less than 1.0 (averaging about 0.86). These spokesmen feel that without the 0.9 guideline suggested by the council, the commissions would, under present economic conditions, set a much lower market-to-book ratio as an indicator of financial hardship.

Traditional Rate-Making Criteria

The council has also received some criticism from both the industry and the regulatory commissions concerning the differences between the council's methods of determining compliance and the commissions' methods of financial evaluation in rate cases. Traditionally, public utilities are regulated on a cost basis. The regulator is asked to determine a company's total cost of service or, in other words, the revenue requirements of the company that will allow it to earn a just and reasonable rate of return. Revenue requirements are expenses, plus depreciation, plus taxes, plus the return allowed on the rate base. The objective of a fair rate of return is to establish a level of future earnings sufficient to attract capital. This attraction of capital is necessary in order for a public utility to satisfy its legal obligation (since it is a natural monopoly) to supply service to its present and
future customers at all times. The rate of return is defined as revenues, minus expenses, divided by invested capital, which is the rate base.

The regulator is charged with setting rates that balance the interests of both the consumer and the investor. The main focus, then, would be on prices or rates. To set these rates, the regulator looks at a test period, usually a year in the past. The regulator or commission must determine a level of profits that would allow the utility a fair rate of return on its rate base and then set prices (rates) at a level such that, if they had been in effect during the test period, they would have yielded the allowed rate of return. A fair rate of return should be the median between inadequate and excessive earnings. In its determination, consideration is normally given to several factors, including ability to attract capital, economic risk, quality of earnings and comparable earnings, quality of service provided, and cost of capital. Prices and expenses in the test period would be compared to actual earnings. The comparison would yield levels of revenues, expenses, investment, and a rate of return that would have resulted had the test period been normal. This result could then be adjusted to account for recent or extraordinary changes. Unfortunately, the last year, especially in terms of costs to utilities, cannot be considered normal.

The commission, as part of its analysis, must determine what is allowed in the rate base to which the rate of return is applied. The rate base generally consists of the physical assets of the utility. Problems arise when items that are held or are under construction for future use are considered. Needless to say, there is not general agreement among commissions on this issue. One example is construction work in progress. CWIP traditionally was not allowed in the base rate, but capitalization of its financial costs was allowed. This was because commissions did not feel current customers should pay for construction of plants for the use of future customers. However, this allowance for funds used during construction (AFUDC) now accounts for 40 percent of net income (in some cases, up to 70 percent) and is a noncash earning. In other words, AFUDC does not produce cash flow dollars, and a reduced cash flow is one of the most severe of the current problems facing the industry. Notwithstanding the old argument that current customers would be asked to pay for plants used to supply future (who could also be the same current) customers, there appears to be some trend toward allowing CWIP in the rate base and discontinuing the capitalization of AFUDC.

Due to the different methods of financial analysis used by the various commissions, it would be impossible for the council to devise a workable price standard that would conform to each commission's present system of financial evaluation in rate cases. Since the Gross Margin Standard is simple to apply, the council feels that it is not overly burdensome for either the utility companies or the commissions to calculate. The data required under the Gross Margin Standard are merely the utility's operating revenues, costs of fuel, purchased power, or purchased gas, and the physical volume growth in kilowatt hours (kwh) or million cubic feet (mcf) between the base period and the program year. The council also believes that if an exception is necessary, the regulatory commissions are better equipped to make an in-depth evaluation of a utility's financial needs.

Conclusion

The council recognizes that all the aforementioned situations represent valid potential problems for the utility industry, and it has given careful consideration to each of these questions. The council remains committed to the approach used in the first-year standards allowing for the administration of the Gross Margin Standard for electric and gas utilities and exceptions by the appropriate public utility commissions or the Federal Energy Regulatory Commission. The commissions have the statutory and constitutional mandates to regulate electric and gas utility prices based on considerable financial evaluation and adversary proceedings. While the council believes electric and gas utility industries should be covered by the standards notwithstanding their regulated nature, the council also recognizes the expertise as well as the statutory and constitutional responsibilities of the regulatory bodies. The council, therefore, requests that these commissions continue to consider the Council on Wage and Price Stability standards as a part of any decision granting a rate increase. The council also believes the suggested circumstances that could warrant a hardship exception remain valid in the second program year. While the average market-to-book ratio for utilities in approximately 0.85, 0.9 is still an indication of financial problems worth consideration, and it has been suggested that any figure below one borders on being confiscatory. The interest coverage consideration clearly is relevant and is one of the most widely used indicators of financial viability. Finally, inadequate cash flow situations can affect a utility's industrial credit rating, discourage investors, and thus impair the value of a company.

While the final results of the Alternative Standard for the first program year are not in yet, some conclusions can be drawn from the nine-month reports. Of the 115 rate increases granted to electric and
gas utilities during the first nine months of the anti-inflation program, 88 were in compliance with the alternative gross margin standard. This means that the commissions were able to grant increases they deemed necessary and still allow only a maximum 6.5 percent increase in operating revenues, less the cost of fuel since the third quarter of 1978. The remaining 27 rate increases were exceptions granted by the appropriate regulatory body with statutory authority and justified on financial grounds by the granting commissions in their final orders. The results to date seem to support the council's position that the Gross Margin Standard is an effective means of measuring a utility's rate increases without penalizing the industry for its energy dependence. The initial results also show that while the council's standards are a genuine barrier to excessive rate increases, they do not represent an unreasonable or insurmountable barrier to utilities in cases of genuine need.

Electric Utility Rate Design and Energy Management Initiatives

Howard Perry

Title II of the Energy Conservation and Production Act of 1976 (ECPA) directed the Federal Energy Administration (FEA) to undertake several major initiatives in the area of electric utility rate design and to transmit to Congress an annual report with respect to these activities. The first of these covered the limited work accomplished under ECPA in 1976. The 1977 report described the progress of the FEA in carrying out the ECPA Title II initiatives prior to 1 October 1977, as well as the transfer of the program to the new Department of Energy (DOE) after that date. The current report describes the progress of this DOE program during 1978. Activities covered include: financial assistance to state utility commissions and public power systems to field test and implement time-of-use rates, federal intervention in state utility regulatory proceedings, grants for state utility consumer offices, and technical assistance to state commissions and public power systems to help plan and carry out rate design and energy management programs. In addition, the DOE conducted a number of analytical projects to support these primary functions.
New Energy Legislation

Enactment of national energy legislation on 9 November 1978 has significantly increased the responsibilities of the DOE's Office of Utility Systems (OUS), which administers ECPA Title II. As provided in Titles I and III of the Public Utility Regulatory Policies Act of 1978 (PURPA), the OUS is involved in several new initiatives, including voluntary federal guidelines concerning the PURPA regulatory policy standards, grants to state commissions and nonregulated utilities, reports from state commissions concerning their progress in carrying out the act, and a comprehensive natural gas rate study. In addition to these new efforts, the activities initiated under ECPA and outlined above are extended by PURPA and, in some cases, significantly expanded.

Title II of the National Energy Conservation Policy Act (NECPA) also affects the programs; the OUS is responsible for those aspects of the NECPA "utility program" relating to state commissions and nonregulated utilities. Although the main responsibility within DOE for this program is assigned to the Assistant Secretary for Conservation and Solar Applications, aspects of the NECPA utility program that are of particular concern to the OUS include cost accounting and billing, advertising, termination of service, and prohibitions against utility financing of conservation measures. The OUS is also responsible for administering the gas light bill under section 402 of the Powerplant and Industrial Fuel Use Act of 1978 (PIFA) and the fuel conservation measures authority under sections 214, 314, and 501 of PIFA.

PURPA Regulatory Policy Standards

Of the several new pieces of national energy legislation, PURPA is the most closely related to ECPA Title II. PURPA requires that state regulatory authorities and nonregulated electric utilities consider implementing, within a specified time, eleven ratemaking and regulatory standards. The standards will be evaluated against the criteria of end-use energy conservation, production efficiency, and equitable rates for consumers. They are as follows: (1) cost-of-service rates; (2) prohibition of declining block rates except when cost justified; (3) time-of-day rates; (4) seasonal rates; (5) interruptible rates; (6) load management techniques; (7) elimination of master metering; (8) automatic adjustment clauses; (9) information to consumers; (10) procedures for termination of service; and (11) treatment of advertising expenses. All of these standards relate to electricity, and the standards of termination of service and advertising relate to gas utilities as well. States are further required to consider lifetime electricity rates as an exception to the cost-of-service standard.

Utility Rate Demonstration Projects

Section 204(1), Title II, of ECPA authorized DOE (FEA) to fund a variety of rate design, load management, and related projects. In early 1975, FEA began to conduct, in cooperation with state and local utility authorities, a number of electric and gas demonstrations under the more general provisions of the Federal Energy Administration Act of 1974. Twenty projects in sixteen states were initiated, seven in fiscal 1975 and thirteen more in fiscal 1976. Ten received continuation funding in fiscal 1977, and five were continued in fiscal 1978.

The Electric Utility Rate Demonstration Program was established to fulfill three major objectives: (1) demonstrate to utilities and regulators the viability and customer acceptance of innovative electric rates; (2) gather empirical data as to the effects of such rates on customer and class electricity consumption patterns; and (3) transfer the results of comprehensive analysis of these data nationwide.

Of these objectives, perhaps the most important to utilities and state and local regulators is the analysis of customer response data under such rates. In particular, utilities and regulatory commissions considering adopting time-of-day rates need to estimate their customer and system load effects.

Funding for these rate demonstrations is summarized in Table 1. Most of the electric rate demonstrations listed in the table have provided either interim or final customer response data to DOE. The analyses performed to date tend to show that the introduction of time-of-day rates has had a significant impact on the load patterns of residential and larger commercial and industrial customers. More specifically, the analyses indicate that time-of-day rate customers reduce kilowatt-hour consumption during the utility's peak period, compared to control customers on traditional declining block rates. In addition, the data reveal that time-of-day rates generally reduce both residential and industrial class diversified kilowatt demands, even at hours of utility system monthly and annual peak. In the future, as additional data become available, further analyses will be conducted to produce more precise and thus more useful results.

The demonstration program has also made significant progress in other areas. Since 1975, regulatory commissions and utilities involved in the program have acquired considerable skill in overcoming the technical and administrative impediments to the adoption of time-of-day rates. Also, as the demonstration projects are completed, project personnel are conducting follow-up surveys to determine consumer understanding and attitudes toward the new rates. In the coming years, therefore, as the regulatory commissions and major nonregu-
Energy Management Initiatives

Table 1. Funding for Rate Demonstration Projects, Fiscal 1975 – 1978, in Thousands of Dollars

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Arizona</td>
<td>210</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Electric*</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Gas*</td>
<td>79</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Arkansas*</td>
<td>210</td>
<td>160</td>
<td>42</td>
<td>—</td>
</tr>
<tr>
<td>California</td>
<td>—</td>
<td>—</td>
<td>355</td>
<td>215</td>
</tr>
<tr>
<td>Electric</td>
<td>—</td>
<td>—</td>
<td>143</td>
<td>—</td>
</tr>
<tr>
<td>Gas*</td>
<td>—</td>
<td>—</td>
<td>130</td>
<td>—</td>
</tr>
<tr>
<td>Connecticut</td>
<td>210</td>
<td>30</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Edmund, Oklahoma</td>
<td>—</td>
<td>202</td>
<td>182</td>
<td>50</td>
</tr>
<tr>
<td>Los Angeles, California</td>
<td>170</td>
<td>350</td>
<td>270</td>
<td>280</td>
</tr>
<tr>
<td>Michigan*</td>
<td>—</td>
<td>260</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Wisconsin</td>
<td>200</td>
<td>181</td>
<td>—</td>
<td>470</td>
</tr>
<tr>
<td>New Jersey</td>
<td>160</td>
<td>555</td>
<td>450</td>
<td>—</td>
</tr>
<tr>
<td>New York</td>
<td>—</td>
<td>—</td>
<td>555</td>
<td>265</td>
</tr>
<tr>
<td>Electric*</td>
<td>—</td>
<td>—</td>
<td>555</td>
<td>248</td>
</tr>
<tr>
<td>Ohio</td>
<td>220</td>
<td>168</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Electric*</td>
<td>—</td>
<td>—</td>
<td>130</td>
<td>—</td>
</tr>
<tr>
<td>Gas*</td>
<td>—</td>
<td>—</td>
<td>357</td>
<td>292</td>
</tr>
<tr>
<td>Puerto Rico</td>
<td>—</td>
<td>275</td>
<td>125</td>
<td>—</td>
</tr>
<tr>
<td>Rhode Island</td>
<td>—</td>
<td>275</td>
<td>125</td>
<td>—</td>
</tr>
<tr>
<td>Vermont*</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Washington</td>
<td>—</td>
<td>—</td>
<td>261</td>
<td>152</td>
</tr>
<tr>
<td>Total</td>
<td>$1,556</td>
<td>$4,510</td>
<td>$2,168</td>
<td>$1,418</td>
</tr>
</tbody>
</table>

*Project completed.

related utilities comply with the provisions of PURPA, the Electric Rate Demonstration Program will be able to offer a valuable reservoir of practical experience and empirical data to regulatory authorities.

Pilot (Innovative) Projects

Section 204(1), Title II, of ECPA authorized DOE to fund regulatory rate reform initiatives. Ten implementation projects were begun in fiscal 1977; of these, eight received continuation funding in fiscal 1978. The pilot projects are distinctly different from the DOE rate demonstration projects in at least two ways. First, rather than serve experimental purposes, these projects are intended to encourage and assist permanent adoption of cost-based rates and energy management programs. Specifically, the new projects involve either the elimination of declining block energy charges or the adoption of time-of-use rates, or both. Also featured in these projects are customer conservation services, ratemaking guidelines, prohibition of master metering, and energy efficiency criteria for new service hookups.

Second, whereas the demonstration projects have focused primarily on customer response, the pilot projects focus on institutional response. The experiences and lessons learned through this program about successful institutional approaches to utility rate design reform and energy management activities will be useful to the federal government and to state and local regulatory authorities in implementing PURPA.

PURPA authorizes funds to be appropriated in fiscal 1979 and 1980 to continue this innovative effort, which is intended to serve two interrelated purposes. The funds are to defray such costs as ratemaking design and analysis and customer education campaigns, as well as to obtain transferable information about the process of implementing nontraditional utility rates and conservation activities. Funding for these projects is summarized in Table 2.

Table 2. Funding for Pilot Projects, Fiscal 1977 – 1978, in Thousands of Dollars

<table>
<thead>
<tr>
<th>Regulatory Jurisdiction</th>
<th>1977</th>
<th>1978</th>
</tr>
</thead>
<tbody>
<tr>
<td>California</td>
<td>426</td>
<td>673</td>
</tr>
<tr>
<td>Connecticut</td>
<td>581</td>
<td>402</td>
</tr>
<tr>
<td>Grand River Dam Authority, Oklahoma</td>
<td>160</td>
<td>116</td>
</tr>
<tr>
<td>Iowa</td>
<td>284</td>
<td>140</td>
</tr>
<tr>
<td>Minnesota</td>
<td>489</td>
<td>518</td>
</tr>
<tr>
<td>North Carolina</td>
<td>975</td>
<td>269</td>
</tr>
<tr>
<td>Ohio*</td>
<td>335</td>
<td>—</td>
</tr>
<tr>
<td>Seattle, Washington</td>
<td>376</td>
<td>524</td>
</tr>
<tr>
<td>South Dakota**</td>
<td>202</td>
<td>—</td>
</tr>
<tr>
<td>Springfield, Missouri</td>
<td>333</td>
<td>110</td>
</tr>
<tr>
<td>Total</td>
<td>$5,298</td>
<td>$2,551</td>
</tr>
</tbody>
</table>

*Completed.
**Discontinued.

Technical Assistance

Pursuant to the authority of ECPA section 204(1) to fund regulatory rate reform initiatives, DOE has undertaken a comprehensive technical assistance program to help utility regulatory authorities and nonregulated utilities strengthen their capability to carry out rate reform, energy management, and related programs. DOE's assistance activities are directed toward both the gas and electric utility sectors.

Given the great diversity of utilities, it is important that technical assistance be tailored to the particular local characteristics of the utility
Energy Management Initiatives

in question. DOE's technical assistance activities are designed to help each commission (and major nonregulated utility) assess its needs, formulate a new program which meets these needs, and carry out the program in an effective and practical manner.

In order to pursue these objectives, DOE has drawn upon the resources of the National Regulatory Research Institute (NRRI). Operating with DOE 1978 funding of $500,000, NRRI, established under the auspices of the National Association of Regulatory Utility Commissioners, is providing professional assistance to the state regulatory community and major nonregulated utilities. NRRI, originally funded at $1,070,000 in 1977, has become a technical assistance center offering the staff and resource materials needed to assist state and local utility authorities. Specifically, NRRI has provided technical teams for nineteen on-site technical assistance projects, a series of five instructional case studies, and a summary of regulatory computer programs. With fiscal 1978 funding, NRRI is conducting a series of five workshops for state regulatory and energy agencies to address the practical issues of implementing PURPA and NECPA. It will also undertake six technical assistance studies in the area of utility rates and revenues, five studies in the area of state regulatory commission and utility management and operations, and eight on-site technical assistance projects.

Pursuant to section 605 of PURPA, authorization exists for making grants, not to exceed $2 million annually in fiscal 1979 and 1980, to an institute established by NARUC on a cost-sharing basis. The grants may be used to conduct research on electric and gas utility regulatory policy issues, to develop data processing and retrieval methods for electric and gas utility ratemaking, and to perform other functions directly related to assisting state regulatory authorities in carrying out their functions under state law and under PURPA.

State Regulatory Interventions

Section 204 of ECRA authorizes DOE, upon the request of a state, a utility regulatory commission, or a participant, to intervene in electric utility rate and rate design proceedings before state utility regulatory commissions. Section 121 of PURPA authorizes DOE to intervene as a matter of right in state regulatory proceedings which involve consideration of one or more of the electric rate and other regulatory standards established by the act or other concepts which contribute to the achievement of the purposes of PURPA. Section 305 of PURPA gives DOE similar authority with respect to state regulatory gas utility rate and rate design proceedings. Section 305 of the Natural Gas Policy Act of 1978 grants DOE the authority to intervene in state regulatory proceedings relating to proration or other limitations on natural gas production.

The primary purpose of DOE intervention activities is to advocate the implementation of regulatory policies consistent with legislative objectives, national energy policy, and other DOE concerns. During 1978, the pending National Energy Act (NEA) necessitated the reevaluation of positions DOE had taken in past state interventions and the development of new positions consistent with that legislation. As a result, DOE intervened in only five state proceedings during 1978. With the final enactment of the NEA, DOE interventions in state regulatory proceedings are expected to increase significantly and will involve not only the electric utilities but also natural gas utilities. Table 3 lists cases in which the DOE participated during 1978.

<table>
<thead>
<tr>
<th>State</th>
<th>Docket number</th>
<th>Proceeding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colorado</td>
<td>4638</td>
<td>Generic electric rates</td>
</tr>
<tr>
<td>Hawaii</td>
<td>2773</td>
<td>Generic electric rates</td>
</tr>
<tr>
<td>Iowa</td>
<td>76-54</td>
<td>Puget Sound Gas Company (declining block rates)</td>
</tr>
<tr>
<td>Maryland</td>
<td>7226</td>
<td>Washington Gas Light Company (standing service rates)</td>
</tr>
<tr>
<td>Texas</td>
<td>1770</td>
<td>Generic electric rates</td>
</tr>
</tbody>
</table>

Utility Consumer Offices

In September 1977, FEA awarded grants on a competitive basis to twelve states to establish offices to represent consumer interests in electric proceedings before utility regulatory commissions. These awards were authorized by section 205 of ECRA for state offices which operate independently of utility regulatory commissions. In both fiscal 1977 and 1978, $2 million was authorized and appropriated for the program. In September 1977, forty-one states submitted grant applications, and twelve were selected in accordance with published criteria. Because of the grantees' performance during the initial year of the program, the projects were renewed for a second year in September 1978.

During the first year of this program, grantees established and began operating offices pursuant to DOE guidelines. The grantees actively represented consumer interests in their states by performing a number of diversified activities: (1) assessing the impact of proposed rate changes and other regulatory actions on all affected consumers,
(2) providing technical and financial assistance to aid consumers in their presentations before utility regulatory commissions, and (3) advocating before utility regulatory commissions positions determined by the office to be most advantageous to consumers. Grant awards for utility consumer offices are shown in Table 4.

Table 4. Grant Awards for Utility Consumer Offices, Fiscal 1977 - 1978, in Thousands of Dollars

<table>
<thead>
<tr>
<th>State</th>
<th>1977</th>
<th>1978</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arkansas</td>
<td>185</td>
<td>200</td>
</tr>
<tr>
<td>District of Columbia</td>
<td>126</td>
<td>147</td>
</tr>
<tr>
<td>Georgia</td>
<td>188</td>
<td>200</td>
</tr>
<tr>
<td>Guam</td>
<td>60</td>
<td>75</td>
</tr>
<tr>
<td>Idaho</td>
<td>190</td>
<td>200</td>
</tr>
<tr>
<td>Illinois</td>
<td>182</td>
<td>200</td>
</tr>
<tr>
<td>Indiana</td>
<td>200</td>
<td>55</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>196</td>
<td>200</td>
</tr>
<tr>
<td>Michigan</td>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>New Hampshire</td>
<td>40</td>
<td>65</td>
</tr>
<tr>
<td>New Mexico</td>
<td>300</td>
<td>200</td>
</tr>
</tbody>
</table>

| Total           | $1,973| $1,942|

ECPA Support Activities

In addition to the programs described above, DOE performs a number of activities to provide analytical support to the other ECPA Title II program functions. These include comprehensive analysis of data from the demonstration program, analysis of system efficiency improvements attainable with commercially available technology, continuing analysis of electric and gas rate design issues, and analysis and development of associated energy management initiatives. Funding for these support activities is shown in Table 5.

Consolidation of Annual Reports

As indicated herein, the newly authorized PURPA activities are closely related to ECPA Title II activities. Section 116 of PURPA requires state commissions and nonregulated utilities to report annually to DOE, which then must report annually to the Congress.

Since the ECPA Title II activity has been extended and substantially amended by section 142 of PURPA, it appears that the new PURPA reporting requirement supersedes section 206 of ECPA. Accordingly,

DOE anticipates that future developments concerning those activities originally authorized by ECRA will be included in the PURPA annual reports.

Table 5. Funding for Various Department of Energy Activities, Fiscal 1977 - 1978, in Thousands of Dollars

<table>
<thead>
<tr>
<th>Activity</th>
<th>1977</th>
<th>1978</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analytical support related to rate design activities</td>
<td>270</td>
<td>—</td>
</tr>
<tr>
<td>Development and implementation of a comprehensive analytical plan for analyzing data from the rate demonstration projects</td>
<td>250</td>
<td>—</td>
</tr>
<tr>
<td>Technical and economic assessment of district heating and commercially available technologies for reducing transmission and distribution losses</td>
<td>292</td>
<td>—</td>
</tr>
<tr>
<td>Economic and technical analysis of utility rate design issues</td>
<td>192</td>
<td>—</td>
</tr>
<tr>
<td>Technical support for residential conservation service rule making</td>
<td>—</td>
<td>80</td>
</tr>
<tr>
<td>Analytical support related to rate design and impact on commercial and industrial customers</td>
<td>—</td>
<td>250</td>
</tr>
<tr>
<td>Analytical support related to rate design impact on dispersed solar energy systems</td>
<td>—</td>
<td>75</td>
</tr>
<tr>
<td>Analytical support related to regulatory policy standards</td>
<td>—</td>
<td>$914</td>
</tr>
<tr>
<td></td>
<td>$955</td>
<td></td>
</tr>
</tbody>
</table>
Federal-State Relationships in Energy Regulation: A State Commissioner’s Perspective

Maurice Van Nostrand

President Jimmy Carter sent his energy bill to the Congress on 20 April 1977. At that time, it should have been evident to every state utility commissioner that the world in which we operated was going to change, that Uncle Sam was going to get very involved in utility ratemaking.

Some of us got the message several months before. The comments the Congress had made in 1975 and 1976, and the direction the Federal Energy Administration was taking during those years, were ominous that the federal government wanted into our act. And those of us who participated in early discussions with the Carter energy people learned in February and March of the emphasis retail electric and gas rates were going to get in the planned solution to the energy crisis.

I give this early history as a preface to a statement that is blunt but certainly not meant to be critical: State commissioners have not yet developed a real awareness of how massive an impact the 1978 Energy Act is going to have on the already difficult life of a retail utility ratemaker.

The reason the statement is not critical is that the lack of awareness is completely understandable. There are two reasons. First, the unawareness reflects a kind of “what will be, will be” philosophy on the part of state regulators. They were not asked to play much of a part in the development of a national energy program; for the most part (although Paul Rodgers did secure a great victory in getting the Congress to junk the mandates in the Dingell proposals), the views of state commissioners were ignored in the legislative deliberations of energy strategies.

Those of us who testified before congressional committees during the eighteen months of energy debates had our confidence shaken in the system by which this great nation makes major decisions. Not only was it like testifying at a revolving door — with congressmen coming and going, but usually absent — but also one was left with the feeling that, on those issues related to utility regulation, minds were mostly made up.

The second reason for the lack of awareness of what that energy legislation does to state regulation of utilities stems, I believe, from utter disbelief. State commissioners simply cannot believe that a Congress which puts the federal government deeply into retail ratemaking would not — either by accident or intent — do something which would help solve the serious problems faced by state regulators before the act was passed.

I am sincere in that analysis. Let us assume a survey of state commissioners had been made two years ago, asking each to identify the six or eight greatest obstacles to effective regulation of energy utilities in his or her state. I am confident that a review of the new legislation and that list of identified regulatory problems would result in the incredible conclusion that Congress had missed the target entirely.

Would it not have been reasonable for commissioners to expect that Congress, when it started fooling around with utility issues, would solve the nuclear waste problem, or tell us how to account now for nuclear waste costs to be incurred later, or make it clear whether they really want scrubbers attached to plants in which only low-sulfur coal is to be burned? And perhaps some regulators secretly hoped the Congress would tell us how to cope with political demagogues — running for governor or utility commissioner — who say: Vote for me and your utility rates will stop going up.

So it is correct, I think, to say the Congress caught our people by surprise. It ignored the real regulatory problems and concentrated on items apparently perceived as promising maximum political benefit.

But despite the state commissioners’ lack of awareness of the considerable impact the 1978 Energy Act will have on the operation of
their agencies and the companies they regulate, I am optimistic about their actions. They believe in law and the right of the Congress to develop public policy. State regulators will not fight the new law; they will take the pragmatic approach and do their best to make things work.

They know their states well. They are dedicated to performing the responsibilities assigned them, either by the statutes of their states or by the acts of the U.S. Congress. They know that the standards imposed on the rates charged and the service provided by the retail utilities, whatever they are, can be better administered by the states than by the federal government.

As this new relationship begins, the intentions are all good. I believe. Both relations and communications between state commissions and federal agencies have never been better. There is presently a feeling of mutual respect that has not always existed. And I have heard no one express anything but positive motivations to carry out the new and technically heavy role allocated by the Congress.

Having conveyed (for myself, my commission, and, I think, the overwhelming majority of NARUC's state commissioners) a total commitment to making our part of the new energy legislation work, I do want to make some points in a slightly different vein. There are three aspects of the 1978 Energy Act which bother me very, very much. I think most state commissioners would agree, especially those whose responsibilities in their states' energy affairs have gone beyond the area of electric and gas utility regulation.

Oil Is the Problem

Surely, no one can argue that liquid energy — oil — is the most serious component of America's energy problem. The outflow of $45 billion in each of the last two years for oil threatens, as has nothing before in history, the stability of the U.S. dollar. The vulnerability of our economic and political freedom to oil embargoes by foreign powers is a threat almost impossible to overstate.

We must more nearly balance domestic production and consumption of oil. That must be the main goal of energy legislation. Our friends around the world are crying out for action to demonstrate that America's insatiable demand for foreign oil will be stemmed.

But that kind of demonstration is conspicuously absent in the Energy Act. The Congress clearly lacked the courage to tackle the oil problem. Recognizing that, one would have to be superhuman not to resent being asked to alter materially the way we produce, distribute, and price other forms of energy. I do resent it. In my part of the country, almost no electricity is generated from oil. It is hard for us to think the Congress means business when it asks us to tamper with our electric system — which presents no problem — and ignores the energy catastrophe of oil.

We are asked to identify and punish people who consume electricity and natural gas at inappropriate times of day or for inappropriate uses. But Americans are encouraged and even urged to buy 15 million cars and trucks each year and drive them as much as they want. The conflict in those two positions is more than I can tolerate in silence.

When will the Congress have the courage to tell the American people that our serious energy problem cannot be cured by zeroing in on appealing targets like utility companies, or large users of electricity or natural gas, or even the auto makers? Only by getting all 218 million Americans into the act, by putting into effect those measures which will decrease oil consumption and also increase domestic oil production, will the greatest peril to our nation be erased.

The argument can be made, although only weakly, I believe, that the restrictions on utility operations will reduce oil imports. But I think an argument also can be made, especially if the provision for incremental pricing of natural gas is not cautiously applied, that some large consumers of energy will be forced to use more foreign oil.

It would be easier for those of us in utility regulation to administer the new constraints with enthusiasm if we could point to a visible, across-the-board commitment to solve America's energy ills. But we will have to swallow hard as we tell our people they are bad, somehow, for wanting to keep their house cool on a blistering August day, when charging around a lake all day in a speed boat is still in compliance with congressionally developed public policy. I do not know what a court would think about this dichotomy, but I can tell you that a lot of people in Iowa think it is pure nonsense.

Where Do We Get the Technicians?

Even in the so-called good old days, utility ratemaking called for the use of large numbers of people specially prepared for complex responsibilities. There is no one academic degree or college course that teaches how to develop cost-of-service evidence. Only by exposing bright, inquisitive people to a mix of disciplines can a background be secured that will assure a capability to handle such diverse yet related issues as depreciation, cost allocation, heat rates, cost of capital, tax accruals, working capital studies, and so forth.

There is a terrible shortage of such people now. At our commission, where we pay above-average salaries and where, at least in our estimation, we maintain a regulatory climate that is attractive to such people,
we have been unable to hire all the personnel we need. At the present
time, we are short 17 staff members to fill the authorized table of
organization in the Utilities Division.
Even before passage of the 1978 Energy Act, the shortage of people
to do this kind of work was becoming acute. All the old issues were still
there in rate and service cases; in addition, there were new demands for
rate design analysis and development of procedures by which utility
management could be evaluated.
The need for people able to perform broad, cost-of-service
functions was not confined to state commissions. On their own initi-
tive and encouraged by state regulators, utilities were hiring and train-
ing people to do this kind of work so the use of consultants could be
eliminated or, at least, reduced.
The addition of these "cost-of-service" types to utility staffs has
been extremely useful. To withstand today's pressures, companies
must better understand, and be better able to explain, why they do
things in certain ways. The flat statement by a high-priced consultant
that "that's the way this is done" no longer suffices. Accounting, cost-
ing, retirement, and cash handling procedures must be proven appro-
prate if they are to meet the tests of today.
Where are such people to come from? State commissions and utility
companies now need these multidisciplinary individuals not only for
routine rate and service responsibilities, but also for effective applica-
tion of the provisions of the new energy legislation. In the utility
conservation programs alone, it seems to me, if we want results that
justify the considerable expense contemplated, the home audits must
be done by people with the breadth of general knowledge now asso-
ciated only with cost-of-service types.
And the Department of Energy itself is to render valuable assistance
in getting the job done as mandated by the Congress, will need many
people acquainted with the techniques of developing cost-of-service
information. If cost-benefit ratios are to be measurements of the effec-
tiveness of utility conservation programs — and surely we have the
right to expect that — an understanding of the trade-offs inherent in
every utility operation is imperative.
This is not a money problem. I assume the Congress will appro-
priate the money it has authorized for aid to the states in carrying out
those federal mandates. But providing money, although not easy, is far
less difficult than providing trained people.
In my own state, we have had excellent cooperation from the
governor and the legislature; they have given our commission all the
money for staff that we have been willing to justify. But that serious
vacancy factor in our staff persists, and this new bill, by creating a
greater demand for an extremely scarce commodity, will make it worse.
In earlier conversations with Dave Bardin and Howard Perry, I
suggested that the Department of Energy and NARUC work together
to train large numbers of cost-of-service people quickly. Surely we can
find several academic institutions willing to start with a basic discipline
— perhaps accountancy is the most appropriate — and add principles
of regulatory law, engineering, and economics in a crash program to
fill this need.
At this point, I might say I see this as one of the more truly
productive results of the new activities in which we are all going to be
engaged. Utility operation and utility regulation have fallen into the
trap of excessive specialization. No uncompromising articulation of
regulatory law, economic principle, doctrine of conservation, or en-
26
Maurice Van Cottland
reasoning requirement will be very helpful as a guideline in these areas.
If we can come out of this period with a cadre of well-trained
individuals who understand and can explain how those goals and
principles — applied together, and tested against each other — provide
the only sound basis by which utility service can be provided and
priced, we will have made a significant contribution to the well-being
of our nation.
Reasonably Adequate Service at
Reasonable Cost
The third aspect which bothers me about the direction in which we
seem to be heading, and which bothers me most, can be expressed as
who is going to mind the store? Who will assure the public it will be
provided reasonably adequate utility service at reasonable cost? Who is
going to serve as the substitute for the competitive forces that, if they
were present, would compel efficiency and force prices to or near cost?
If utilities as we know them are to survive, that is a responsibility that
must be carried out by state commissions.
As I understand it, the reason for regulation of utilities — as
natural, necessary monopolies — is to guarantee the public interest will
be given consideration when there is a divergence of view between the
utility and its customers, since these customers are denied a choice of
suppliers. I still think that is our primary job, but the Congress, in
drafting the Public Utility Regulatory Policies Act, apparently does not
think so.
Let us examine the many issues on which utility companies and
their customers could have differing views. You will note I said there
could be differing views; my listing these issues does not mean that in
The state commission, therefore, must not allow itself to get carried away with the new, exotic direction in which Congress is driving regulation of utilities. If the state commission does not continue to devote substantial time and resources to rate of return, accounting, and rate base issues, who will assume this absolutely vital task? The answer, I assert, is no one.

The Congress has turned loose, with the passage of the Energy Act, a new bunch of well-financed players in the regulatory arena. We will have public advocates galore and a flood of utility-financed (actually customer-financed) intervenors cajoling through our hearing rooms. But will these new players help us with the tough issues of rate regulation, the issues that, unfortunately, many commission staffs are not competent to develop thoroughly? Not if what we see thus far is any indication.

They are going to be busy on the issues Congress says are important. They are going to be alleging that one category of consumer is profiting at the expense of another, that there is no incentive for certain consumers to conserve, that one type of utility company is trying to eradicate another, that companies are building plants they would not need if they pushed conservation.

In my state, we analyzed and rejected those theses more than six years ago. We concluded we were not going to pit one group of Iowans against another for the demagogic purpose of finding a scapegoat on whom to blame our energy plight. We concluded that if our production, pricing, and use patterns were inappropriate for these new changed circumstances, we would revise them as rapidly as possible, within the economic and political framework in which our people believe.

But we decided we would not assume — nor would we listen to accusations that — these patterns were developed by evil, greedy people out to wreak havoc upon an ignorant, unsuspecting Iowa public. That decision came easily, based upon pure logic, for there was no indication then, and there is none now, that one group of Iowans has profited excessively from the energy mess in which we find ourselves.

Conclusion

In closing, let me reiterate my confidence that state commissioners will do their part to make this new federal-state relationship return the maximum benefit to the consumers. They will do that not because they believe unconditionally in every goal of the new law or the way these goals are to be achieved, but because they are realists, and they respect the law.
Generally, they like and respect the Department of Energy and Federal Energy Regulatory Commission officials with whom they will work. They have no reason to doubt the integrity and intentions of those officials charged with monitoring the progress of the state commissions in achieving the goals identified by the Congress.

I believe, however, that the state commissions to which history will give the highest marks for this period will be those that carry out the new responsibilities without diverting emphasis and resources from their primary task of rate and service regulation of utilities, serving as surrogate for competitive forces. For it is rate and service regulation as we have always known it, carried out more effectively than before, that will give the public the greatest assurance its utility needs will be met at reasonable cost.

Energy Regulation and the Federalist Solution

Rodney E. Stevenson

Despite Jimmy Carter, the National Energy Act (NEA) was not the result of immaculate conception. Some did say that the administration's conversion of Congress did demonstrate the teachings of old: "Thou shall give unto others if thou expect others to give unto thee." The administration gave: promises for more highways, an international trade center for Harlem, a new federal office building for Milwaukee, a review of the administration's antithrader policy, and a host of other offers — as well as intimations of judgships, commission chairmanships, ambassadorships, and free trips to China with James Schlesinger. And Congress delivered. But what resulted was an act of great irony. The keystone (so labeled by the administration) of the NEA is a set of natural gas deregulation provisions remarkably similar to those which only a year earlier were denounced by President Carter as shameful examples of oil industry profiteering. And although the administration ardently stressed that the passage of the NEA was absolutely necessary for showing our "national resolve" to other nations, the dollar slumped in the world money markets on the day following passage and sank to record lows in the weeks that followed. Although the NEA is significantly different from the administration's
original proposals, the act and the Department of Energy which will administer it are, in the classic sense, a "Federalist" solution.

The Federalist Solution

The "Federalist solution" connotes in part the continuing tendencies toward centralizing power in the federal government and concentrating power in the executive branch. The "Federalist solution" also connotes a special relationship between the predominant economic interests and the federal government. Although these centripetal tendencies are not new, the pace of centralization has clearly accelerated since the time of the Great Depression.

The federal government's evolving role in the economy can be charted by several major events: the rapid expansion of federal regulation (both during the 1950s and over the past fifteen years), the legislation of full employment policies (in 1946 and more recently with the Humphrey-Hawkins bill), the "fine tuning" activities of the Federal Reserve Board, the use of tax rates to stimulate aggregate demand (the tax cuts of the 1960s and deficits of the 1970s) and promote plant investment (the allowance of accelerated depreciation and the investment tax credits), and the control of wages and prices through voluntary guidelines and direct intervention, to name but a few. The creation of the Department of Energy and the passage of NEA are in keeping with the federalist trend and may prove to be a major milestone on that path.

The federal tendencies, however, are not thrust upon an unwilling and unsupporting economy.0 Aberrations will occur, but governments and governmental policies tend to mirror the structure and wishes of the predominant economic interests in the society. Put succinctly, "one dollar — one vote" is more aptly descriptive of our process of governance than are the usually espoused nostrums. As our economy becomes increasingly complex and as the major economic agents become increasingly intertwined, the uncertainties of a "truly unregulated" market or "unplanned" economy become less palatable, particularly to those with a major stake in the economy (recall the fears expressed when Lockhead was in danger of bankruptcy). The movement toward federalism can amuse these concerns by stabilizing markets, absorbing risks, and legitimizing private planning activities.

The Federalist Heritage

From the founding of our nation, there has existed a great debate over the desirability of the federalist solution. In the early days of the republic, men such as Alexander Hamilton, Henry Clay, Daniel Webster, and John Adams strove for a strong government. But the essence of federalism was not merely the relative power of the federal and state governments; it was to be found in the issue of who had access to the reins of government. Hamilton believed that no society could succeed "which did not unite the interest and credit of rich individuals with those of the state." According to him, "all communities divide themselves into the few and the many. The first are rich and well-born, the other the mass of people. . . . The people are turbulent and changing; they seldom judge or determine right." The solution for governance was simple: "Give, therefore, to the first class a distinct, permanent share in the government. They will check the unsteadiness of the second, and, as they cannot receive any advantage by a change, they therefore will ever maintain good government." Webster, who believed "power naturally and necessarily follows property," concluded that "it would be the part of political wisdom to found government on property." Chancellor Kent stated the federalist case in the following terms: "The notion that every man that works a day on the road, or serves an idle hour in the militia is entitled as of right to an equal participation in the whole power of government, is most unreasonable, and has no foundation in justice. . . . Society is an association for the protection of property as well as life, and the individual who contributes only one cent to the common stock, ought not to have the same power and influence in directing the property concerns of the partnership, as he who contributes his thousands." The lawyer Jeremiah Mason spoke even more bluntly: "As the wealth of the commercial and manufacturing classes increases, in the same degree ought their political power to increase. . . . I know this aristocracy of wealth is apt to be evil spoken of. But in a country where wealth greatly abounds, I doubt whether any other foundation for a stable, free government can be found."3

In the early days of the republic, the objective of the federalist solution was to bind political power to men of property and to foster and protect the emergent manufacturing sector. Andrew Jackson, however, would describe a government based on the federalist perspective in other terms — as "an engine for the support of the few at the expense of the many."4

Since the time of Alexander Hamilton, we have evolved from a nation of individuals into a nation of institutions. Still, our federalist heritage remains, although now the intent is no longer to bind political power to men of substantial property, but to bind political power to those institutions which gather and control great wealth.
Federalism and Energy Policy

There should be little doubt that the general concerns which prompted the creation of the Department of Energy and the passage of NEA were real. The Arab oil embargo and resultant price increases, the tight supply of gasoline and fuel oil, the growing dependence on foreign oil, and the decline in domestic oil and natural gas production increased the nation's awareness of the energy problem. Federal authorities were both blamed for the crisis and beseeched to find a solution. The federal government was called upon to "coordinate" its energy policies, assume the risks associated with large-scale projects, support emerging technologies, and provide larger economic incentives to stimulate higher levels of production. The solution which evolved is "federalist," that is, it centralized policy in the executive branch, diminished the power of state agencies, and altered the balance between the "first class" (in Hamilton's terms) and the "second."  

If a federalist solution for energy policy is to be effective, decision-making authority must be concentrated, the relevant decision makers must be of like opinion and fairly free of accountability strictures, information and effective input must be limited, and the capability must exist to enable effective direction of those groups or decision-making units which are not directly controlled.

Energy policy decision making at the federal level is now quite concentrated. Not only were the various agencies (the Nuclear Regulatory Commission being the most notable exception) brought together into a single agency, but also the head of the new agency was given broad discretionary power. The title "Energy Czar" is not applied to James Schlesinger without bias. No longer do agencies of differing but overlapping jurisdiction provide a form of "checks and balances" within the executive branch. The Secretary of Energy has (with the exception of certain matters under the jurisdiction of the Federal Energy Regulatory Commission) the power to enforce and coordinate the policy of the agency — and has great leeway in the use of that power. It has been used in a multitude of ways: offering favorable interpretations on a host of oil and other DOE energy regulations, to obtain the oil industry's support for the NEA; withdrawing, on the last day for submission, Alaskan oil pipeline rate of return testimony which had been solicited and enthusiastically endorsed by DOE's Economic Regulatory Administration but was offensive to the oil industry; blocking the importation of natural gas from Mexico because it was either overpriced (1977) and thus damaging to consumers or underpriced (1978) and thus threatening to the successful completion of the Alaskan natural gas pipeline; and limiting the potential for importing liquefied natural gas. NEA, which was promoted quite fervently by DOE, increases the secretary's powers, particularly in the areas of boiler fuel use, natural gas curtailments, conservation monitoring, and so on, consumer group funding. In short, jurisdiction which at one time dispersed through a number of noncoordinated agencies in the federal bureaucracy is now concentrated primarily in one agency in which overall policy can be controlled with a handful of appointments.

Even with a very concentrated decision-making process in place, the effectiveness of a federalist agency will be augmented if there is reasonable conformity of opinion regarding the federalist intent among those who populate the various layers of the administration of organization. At the very least, civil servants in the hierarchy should be ideologically bland and willing to carry out assigned tasks quietly and reasonably promptly. While it would probably be a gross overstatement to say that DOE is a handmaiden of the energy industry, there is a special sensitivity to the concerns of the industry. Such concerns are mirrored in DOE's hiring practices. For example, FERC head Charles Curtis refused to rehire David S. Schwartz (a former employee in the Federal Power Commission) in part because of Curtis's concerns that Schwartz had been "a rather visible advocate" of the public interest point of view on energy issues. Although Curtis described Schwartz as "brilliant, very articulate, and almost uniquely qualified in this field," he expressed fear that Schwartz's background would make it difficult for the petroleum industry to regard him as "fair." At the same time, Laura Coleman, a rather visible advocate of the oil and gas industry's point of view, was nominated for the post of General Counsel in DOE. When questioned as to the propriety of the appointment, Secretary Schlesinger responded that it would be "guilt by association" to disqualify Coleman on the ground that he had spent twelve years defending oil and gas company clients.  

In Washington, one of the great truisms is that "knowledge is power" — or at least a fellow traveler. Knowledge in the area of energy policy has several dimensions: technical knowledge, knowledge of status of emergent policies, and knowledge of the decision makers. For the federalist solution to be effective, access to knowledge must be controlled and inequitably distributed. Technical knowledge will be more concentrated within industry. Such knowledge is required for industry's production, development, and research activities. Few will master the technical complexities of the industry and then choose not to use it. Regarding the other two areas of knowledge, the structure of organization at the federal level assures unequal access. Prior to the concentration of federal energy policy making in DOE, keeping track
of policy was an arduous task. With the creation of DOE, a vastly more complex and insular organization than those which preceded it, the task has become Herculean. Tracking federal energy policy and cultivating federal policy makers requires substantial resources. While the industry can maintain a large and highly qualified staff of representatives there, few of Hamilton’s “second class” can afford such expenditures. Even if the relevant agencies of power were to hold themselves open to all who come to court, the economic imbalance among the affected interest groups would demand an imbalance in the acquisition of knowledge. Yet, knowledge does not flow only one way. An unequal presence in Washington results also in an unequal presentation of parables and pleadings to the government, thus reinforcing federalist tendencies. While it might be argued that various portions of NEA (particularly the funding provisions within the Public Utility Regulatory Policy Act [PURPA]) were enacted to redress these imbalances, it will be argued below that neither the intent nor the effect of these sections of the law are so clear-cut.

DOE’s power is not absolute. Although most federal energy policy jurisdiction has been concentrated in DOE, the individual states retain considerable authority. To replace the state mechanisms with a unified federal authority would be administratively impractical and politically infeasible. Far better, from a federalist perspective, for Washington to concede and the states to deliver. Historically, the states have had great latitude in exercising their regulatory powers. NEA, however, encroaches on that latitude and enables the federal government to “set the agenda” for the states for a number of energy issues. In the area of conservation services to be provided by large utilities to residential customers, DOE will promulgate rules (standards for materials, installation, consumer protection, and so forth), and the states will be responsible for carrying out the regulatory oversight. In the area of natural gas pricing, the state commissions are forbidden to establish end-use pricing policies which would mitigate the impact of the incremental prices which are to be developed by FERC. FERC will be able to shelter industrial cogenerating firms from the implementation of public utility regulations by the state. In electric utility pricing, the state commissions are required formally to consider a number of rate and nonrate issues (time-of-day pricing, elimination of declining blocks, line rate, direct load management, and so forth). DOE will have the responsibility of ensuring that the states carry out the consideration of these policies in an appropriate manner.

Indirect guidance through agenda setting can be made much more effective with an incentive system in place. Such a system can be found in PURPA, which gives DOE authority to administer over $50 million annually (in fiscal 1979 and 1980) to state commissions, consumer protection boards, and various public interest groups. The funds may only be used to underwrite efforts directed at projects in keeping with the PURPA agenda. In the era of Proposition 15, the PURPA funds are attractive to state and local groups — and provide a substantial inducement for those groups to support the passage of NEA. One recalls Hamilton’s ultimately successful efforts at establishing the Bank of the United States, aided by his proposal to assume state debts.

PURPA funding, far from addressing the imbalance in knowledge acquisition and dissemination (mentioned above), can further the federalist cause in a number of ways. Since state regulatory and public interest efforts are not infinitely expandable, work on approved (and hence fundable) projects will of necessity limit the abilities of the groups to pursue projects relevant to alternative paradigms. The ability to discriminate in the granting of funds provides the grantor with a means of penalizing the impudent and rewarding the faithful. As indicated by DOE’s Howard Perry, “not all states will get funds since not all states will be able to use the funds wisely.” Of course, the relevant issue is who defines the dimensions of “wisely.” One other effect of PURPA funding could be to erode local sources of revenue (and perhaps support) for those groups receiving public funds. Such an erosion would limit the fundee’s long-run independence and promote an attitude of deference toward the supplier of federal funds. Even free money has a price.

The political solutions implemented to deal with the energy problem are not merely federal, they are federalist in nature. However, one paradox should be addressed. The “Keystone” of NEA is the Natural Gas Policy Act, which removes federal regulatory control from the wellhead pricing of natural gas. How can we argue that NEA is an instrument of federalism when its centerpiece is a dismantling of federal authority over private business activities? The answer is to be found in the nature of the energy market and in the meaning of the federalist intent. Through deregulation, natural gas producers are not being thrust into a competitive environment in which survival requires sharp wits and low margins. Deregulation will permit natural gas producers to reap substantial windfall gains, for the market they are being “thrust” into is one in which the price of energy is being artificially maintained at levels far exceeding the resource requirements of developing new energy resources. It also should be recalled that the federalist intent is not so much the promotion of large-scale central governments per se, but rather the development of policies which bind
in a symbiotic fashion the state and men (or institutions) of great wealth. The imposition of wellhead price regulation was "accidental" — the result of a 1954 Supreme Court interpretation, many argue, which was not in concert with the intent of Congress. In this instance, men of property were separated from power and lost the right to exploit their resources for maximum financial benefit. But as Benjamin Leigh observed in an earlier age, "power and property may be separated for a time by force or fraud — but divorced, never. For, so soon as the pang of separation is felt — property will buy power, or power will take over property."* If the federal government had used its power to take over the property of the producers, both men and institutions of great wealth would have been alienated from the state. Using federal power to grant substantial economic rents to Hamilton's "first class," at a time when energy markets are being constrained by foreign cartels and domestic producers are growing increasingly monopolistic, is clearly an act intended to benefit the institutions of wealth and bind them to the state. In short, natural gas deregulation is a supreme act of federalism.

Concluding Remarks

The argument pursued here is that our national energy policy is a modern version of Hamiltonian federalism. As such, that policy is "antidemocratic" within the Jeffersonian-Jacksonian meaning of the term. The policies do not augur well for broad-scale democratic participation in the decision-making process. The end results are not equally beneficial to all segments of society.

What we have not addressed is whether another approach to the energy problem would be as effective as those arising from the federalist solution. Indeed, it might be that property and power cannot be separated and that concentrated and insular power is best for dealing with problems which are both volatile and threatening to the nation's well-being. If this is so, we may have no choice but to accept a hierarchical society and accede to the federalist solution. But if feasible alternatives for the social control of business exist, they should be pursued vigorously. As Franklin Roosevelt warned of the growing federal institutions: "We have built up new instruments of public power. In the hands of a people's government this power is wholesome and proper. But in the hands of political puppets of an economic autocracy such power would provide shackles for the liberties of the people."**
energy policy. Indeed, to the extent that many important energy policy problems involve a conflict between greater production of energy and protection of the environment, the most important independent check clearly remains in the form of the Environmental Protection Agency, which has been at odds with the Department of Energy on a number of issues. Moreover, the independence of the Federal Power Commission, renamed the Federal Energy Regulatory Commission (FERC) and placed within DOE, was a major concern of the Congress when it was debating the reorganization legislation. The new FERC was agreed to only when that concern was satisfied through language clearly protecting its status as an independent regulatory commission.

Finally, Stevenson seems to ignore an argument that is often made and undoubtedly correct to some degree: The very concentration of federal energy policy within one department enhances rather than restricts the accountability of that agency by minimizing the buck-passing that can plague decentralized decision making.

More important, I find it difficult to accept the thesis that the recent changes in the apparatus for making energy policy represent a triumph of Alexander Hamilton's politics and economics over those of Thomas Jefferson and Andrew Jackson. It is true that the states have had thrust upon them an "agenda" in the area of electric rate design, and that change, along with several others, implies a shift in power from the states to the federal government; and it is true that this development would no doubt have pleased Hamilton more than Jefferson. But it is, to a large degree, only an agenda. In the process of electric rate design, the states are asked to "consider" a number of issues within a certain time frame; they are not ordered to adopt them. If adoption is compelled at some stage, there will be more to the point Stevenson is making here.

But, even in that event, I could not accept Stevenson's contention that "the federalist intent," which is supposedly furthered by such legislation, "is not so much the promotion of large-scale central government per se, but rather the development of policies which bind in symbiotic fashion the state and men (or institutions) of great wealth." My tenure as a state regulator has increased my admiration of Louis Brandeis's portrait of the states as laboratories of social and economic experiment to a degree I would have previously not thought possible, but I simply do not believe it is axiomatic that the federal government is more favorably disposed to great wealth than are the states. If increased federal authority over energy policy and utility regulation is a bad idea, it is not because it serves the interest of the wealthy at the expense of the poor; and if we are interested in protecting the inter-

Comments

Charles A. Zielinski

The papers prepared by Howard Perry, Maurice Van Nostrand, Rodney Stevenson, and Steven Ailes and Degna Marmo treat various aspects of a phenomenon that a state regulator may or may not like but clearly cannot deny: an increasing federal presence in public utility regulation. Van Nostrand and Perry deal with aspects of the National Energy Act legislation that affect state regulation. The Ailes-Marmo paper discusses in some detail a somewhat different manifestation of the federal presence — the wage-price guidelines — and I shall discuss it last. Let me begin with Stevenson's paper, which attempts to analyze recent federal developments in energy policy from broad political and historical perspectives.

While some of Stevenson's observations seem clearly correct — it is true, for example, that maintaining a presence in Washington is important for understanding and influencing energy policy and that there are disturbing inequalities in the ability of various groups to maintain that presence — others appear overstated or simply incorrect. I question, for example, the assertion that bringing within the Department of Energy a host of functions previously performed by separate agencies would mean, as Stevenson says, that checks and balances no longer exist within the executive branch in the area of
of the poor, asserting states' rights over energy policy and utility regulation is hardly a reliable way of doing so.

One of Stevenson's concerns is that the funding available under PURPA will be provided only to states whose levels of pollution are certified by DOE. This not unreasonable concern that federal authorities will impose their positions by determining the states' agendas is also reflected in Maurice Van Nostrand's paper, although in rather different terms. Van Nostrand fears that Congress, through its stress in PURPA on rate design issues, has signaled to state regulators that revenue requirement questions are to be deemed less important. Given the shortage of competent staff (another serious problem to which Van Nostrand alludes), a federally mandated emphasis on rate design could result in important revenue requirement questions — more precisely, cost reduction and allocation initiatives — being slighted.

While Van Nostrand's concerns about federal agenda setting are legitimate (I, too, want my commission to be the master of its docket), I do not believe that Congress, in enacting PURPA, intended to divert states from their responsibility to see to it that overall costs are kept down and allocated properly among customers through the various rates that are charged. Indeed, sound rate design will itself help hold down overall costs by using price as a signal of the cost of additional consumption. But it cannot be denied that many states, most likely for want of adequate resources, were simply unable to address themselves properly to the issues encountered in purging general cost and subversion of "new" rate design principles into practice. For those states, it was not unreasonable for PURPA to provide an agenda and include funding for carrying it out; indeed, pace Stevenson, one of the important criticisms of Congress in this regard is that the funding it ultimately provided was nowhere near the level initially contemplated and probably needed.

I can wholeheartedly join in one criticism of the federal government developed in Van Nostrand's paper. He is certainly correct in pointing out that our excessive use of imported oil is at the heart of our energy difficulties and that the electric utility industry should not be the only segment of the economy in which oil savings are sought. The failure of the federal government to develop a coordinated attack in all consuming sectors against the total cost — including diminished international influence and political security — of imported oil to our society is unfortunate, to say the least.

I am particularly pleased to see Van Nostrand raising this issue, for when I do so myself, in the context of considering how consumers will be affected by the costs of the coal conversions that are a promising way of cutting New York's dependence on oil-fired electric generation, I am sometimes accused of having too parochial an interest. New York is among the leading states in the efficient use of fuel for transportation, but it is inordinately dependent on oil for electric generation. As a result, the federal government's tendency to seek allocation and rationing for such use of oil, while avoiding like plagues any schemes to control its use in transportation, will impose disproportionate burdens on my state unless adequate federal assistance is provided for coal conversion. Van Nostrand says he would be better able to accept federal tinkering with state regulatory commissions discretion "if we could point to a visible across-the-board commitment to solve America's energy ills." I would be better able to accept the costs coal conversions may impose on my state if I could point to that same commitment.

Howard Perry's paper provides a useful history of NECPA and PURPA as well as an overview of some of the activities DOE and its predecessors have undertaken pursuant to those statutes. My only relevant comment is that many of these projects have generated a wealth of useful information, and it would be helpful if DOE devised some mechanism for better dissemination of that information to interested state regulatory commissions and other parties.

A rather different way in which the federal government has recently superimposed its authority on state regulatory commissions is through the application to public utilities of the Council of Economic Advisers' "utility's wage-price guidelines. As in the case of PURPA, the effect of the federal presence is less to require some specific action than to demand that the state commissions think about what they are doing. Thus, the standards offer a reasonably broad range of exceptions that may be used by state commissions in authorizing rate increases that otherwise would exceed the guidelines, but the commissions are obligated to explain their use of the exceptions.

Ailes and Marmo view the first year's experience with the wage-price guidelines as being generally satisfactory. In particular, they note that $88 of the 115 rate increases granted to electric and gas utilities during the first nine months of the anti-inflation program complied with the Alternative Gross Margin Standard; the remainder were justified on financial grounds by the granting commissions in their opinions. Ailes and Marmo believe these data support the council's position that the Gross Margin Standard is an effective means of controlling a utility's rate increases without penalizing the industry for its energy dependence.
I do not entirely share the authors’ opinion. Our initial reaction in New York was somewhat skeptical; that is, we did not see the standard as providing much more leeway than the standards of either primary price deceleration or exceptional profit margin limitations. Our expectation was borne out, at least in our own experience. Of the seven cases we decided during the first nine months of the program, four complied with the price deceleration standard, while three required justifiable exceptions to the guidelines. (The first four mentioned may be the four that Ailes and Marmo list as complying with the Gross Margin Standard. Perhaps they do, but that is not the basis on which we found them to be in compliance.)

The weakness in the Gross Margin Standard is not only the difficulty commissions have in making the projections needed to determine compliance, but also, and more important, its failure to take into account the effects of the extensive capital investments that electric utilities are often called upon to make even if they do not result in additional sales. When new plants or transmission lines are built, they not only serve expanding loads, but also replace unsafe or unreliable facilities that provide service to existing customers. Including these new facilities in the rate base will increase the utility’s revenue requirements, but the Gross Margin Standard fails to take this phenomenon into account. Accordingly, the rate increases needed in these situations can only be justified under the guidelines through an exception, even though the situations are not “exceptional.”

At the heart of the problem, of course, is an effort to impose a system of guidelines intended to control inflation in a competitive marketplace on a regulatory system that has the purpose of serving as a surrogate for that competition, but that operates by rather different rules. It is a commonplace that regulatory commissions may set rates no lower than those that would allow the company to recover its reasonable expenses and earn a reasonable return on its investment; if rates meeting this standard do not comply with the wage-price guidelines, an exception must be allowed. By adopting guidelines that require the frequent granting of exceptions, the council may unwittingly make the task of state commissions even harder by requiring them to justify those exceptions to an understandably skeptical public. The Gross Margin Standard, I recognize, was adopted in an attempt to obviate the frequent granting of these exceptions, but I am doubtful that it has succeeded in achieving its purpose to the extent Ailes and Marmo suggest.

In reacting to the papers by Rodney Stevenson and Maurice Van Nostrand, I am torn by conflicting inclinations. On the one hand, I am inclined to condemn the federalist solution that Stevenson describes and the intrusion of the federal government into what had hitherto been areas reserved almost exclusively for state public utility commissions. On the other hand, I feel compelled to recognize the reality of the situation, namely, that an effective energy program is only possible if it represents broad national concerns, and these are likely to be developed only at the federal level.

Based on my experience as the Consumer Advocate in public utility matters for the Commonwealth of Pennsylvania from 1976 to 1979 (and other experiences as a state official), I must express my general agreement with Stevenson’s distrust of the federalist solution. In particular, I agree with his statement that “even if the relevant agencies of power were to hold themselves open to all who come to court, the economic imbalance among the affected interest groups would ordain an imbalance in the acquisition of knowledge. Yet, knowledge does not flow only one way. An unequal presence in Washington results also in an unequal presentation of parables and pleadings to the government, thus reinforcing federalist tendencies.” The difficulty, of course, is that the same basic observation can be made regarding
regulatory activity at the state level. By simply substituting "Harrisburg" for "Washington," one comes very close to the sentiments that many in Pennsylvania have expressed, even after the creation of a tax-supported institutionalized advocate for consumer interests in utility matters. Thus, the unfortunate tendency at the state as well as the federal level has been the development of that same "special sensitivity to the concerns of the industry" which, quite rightly, troubles Stevenson.

All the same, as Stevenson says, it is probably true that a further concentration of authority in Washington in the energy area will make it even more difficult for the interests of ordinary citizens to be effectively presented and heard as important decisions are reached and implemented. Certainly, compared to many state utility commissions, I have found federal regulatory agencies to be much less responsive to grass roots concerns and much more responsive to the desires and needs of large corporate institutions. I am sure that this is so for the reasons expressed or implied by Stevenson. But we are still faced with the question of how to deal with the energy problem effectively. In this respect, Van Nostrand's paper is also troubling. While he strongly criticizes the congressional approach, he also neglects to discuss the defects and problems of state regulation that have made a national approach to the problem so terribly necessary.

I should state that I am quite aware of how increasingly frustrating it is to be a public utility commissioner in today's political and economic climate. I detect in Van Nostrand's paper a clear expression of the frustration that he and many of his colleagues throughout the country feel, not only with regard to federal-state relationships, but also with regard to the increasing difficulty of doing the job that state statutes have set out. Consider the difficulty faced by all utility commissions in attempting to simulate the free market for monopoly enterprises. One need not be a distinguished economist or experienced regulator to recognize how far we are always likely to be from achieving this goal. Add to this traditional problem the increasingly trying experience of attempting to keep consumer and utility companies happy and simultaneously maintain a semblance of credibility at a time of devastating dislocation with government and the public service. One begins to understand what conscientious public utility commissioners face. To this mixture, add the economic, political, and political dislocation occasioned by the energy crisis; segments of our population are literally finding it necessary to choose between heating and eating. Add the apparent disappearance or diminution of the economies of scale which were thought a permanent characteristic of the industry, and do not forget rampant inflation. One must be utterly irrational to accept a position as a utility commissioner.

As if all these factors were not serious enough, stir into this crazy mix the new institutionalized consumer advocates or rate counselors, of whom I, until recently, was one. These folks appear to make a living driving commissioners into a rage by telling them how easy it is to do the job right. Van Nostrand shows that he is very kind and diplomatic when he avoids numbering this burgeoning group among the demagogues running for governor or state utility commissioner. One might conclude, therefore, and perhaps not even with tongue in cheek, that the arrival on the state regulatory scene of the "leds" would appear to be inconsequential.

In a more serious vein, I believe it was unfortunate that the National Association of Regulatory Utility Commissioners spent much time and effort attempting to block presidential and congressional initiatives in utility rate reform rather than helping the Congress identify the areas of genuine need and concern to which Van Nostrand refers. It is no coincidence that the utility industry in general joined with NARUC in convincing Congress to weaken significantly the original proposals contained in the Carter administration's submission regarding public utility rate reform. The same utility companies and business people who argued in Pennsylvania against time-of-day rates, marginal cost pricing, and conservation pricing of electricity and/or gas, on the grounds that Pennsylvania would be uncompetitive with other states not having such pricing, found no difficulty in arguing before Congress that it would be an improper infringement on state rights to preempt state regulatory commissions in these areas. Those of us who favored serious consideration of such initiatives were extremely frustrated to see official spokesmen for state utility commissions, utilities, and business organizations play both ends against the middle rather than discuss the issues on their merits. It is no wonder that the quality of the debate and the resulting legislation left quite a bit to be desired.

As Van Nostrand correctly points out, the legislation is pitifully weak in critical areas, and it does not call for equal commitment and equal sacrifices from all Americans. Nor does it contribute the visible, across-the-board commitment that is absolutely essential. Instead, the de facto "solution" slowly but inexorably asserts itself, namely, the massive transfer of this country's wealth, which has taken centuries of labor and sacrifice to accumulate, to the oil-producing states and the multinational energy conglomerates. Our slavery to our wasteful energy consumption habits will soon be matched by a more dramatic
and painful kind — the slavery of the poor and the weak to the wealthy and the strong.

However inadequate the legislation may be, therefore, I must express my basic agreement with its central theme. *All* energy consumption must be cut back. There should be an across-the-board reduction in consumption of electricity and gas, as well as other forms of energy — primary or secondary — while we work as a nation on a sane and sensible program ensuring adequate supply for the future. Load management, cogeneration, conservation-oriented rate structures, energy audits, weatherization grants — all these and more must be pursued with vigor and purpose and without delay so as to set the proper example and create the proper atmosphere for elimination of the senseless waste in the transportation, recreation, and other sectors. Conservation (both voluntary and mandatory) will be in the best interests of consumers during what it is hoped will be this transition period from the age of petroleum. Furthermore, the resources we will save (by not building multibillion-dollar electrical generating plants, for example) can and should be invested in the promising renewable energy sources of the future.

We cannot afford to confine our interests and concerns to the traditional areas of state regulatory commissioners. Like it or not, we all must become part of the coordinated and concerted national program, or we will be swept away by the consequences of our failure. We must use our skill and experience to strengthen the national program and help federal regulators succeed, else posterity shall surely condemn us all.
Experimental Time-of-Use Electricity Prices

A. K. Miedema, S. B. White, C. A. Clayton, and D. P. Lifson

This paper has two primary objectives: to summarize the analyses being performed by the Research Triangle Institute (RTI) on time-of-use electricity prices, and to describe broadly our findings from residential rate demonstration projects in Connecticut, Ohio, and Arizona.

In general, the Institute's findings to date show that experimental time-of-use rates significantly affect the load patterns of residential customers. The results suggest two major overall conclusions. First, residential customers on time-of-use rates had almost uniformly lower kilowatt-hour consumption during peak rating periods than did control customers on existing declining block rates. Second, compared to usage patterns of customers on the existing rates, the usage of customers on time-of-use rates implied lower diversified demand of the residential class at the time of the system's monthly and annual peaks. That is, the total power demand on the utility system by the residential class would have been lower at the time of the observed monthly and annual system peaks if the class load was like that of the time-of-use customers.

Note: This work was sponsored by the Department of Energy, Office of Utility Systems, under contract number CR-04-80220-08.
Other tentative findings include these: (1) There was a tendency for kWh consumption to decline sometime during intermediate rating periods; (2) there was no evidence of needle peaking before and after peak periods; (3) base period consumption sometimes tended to increase but never appeared to decline under time-of-use rates; (4) total electricity usage appeared either to remain the same or decline slightly under time-of-use rates as compared to standard rates; and (5) the limitations of these rate demonstration projects are very important to the proper interpretation of the results.

The Rate Demonstration Program

High energy and capacity costs in the 1970s have forced electric utilities, public utilities commissions, and federal energy authorities to consider a number of load management policies. These policies, designed to redistribute the time pattern of electricity demand, are potentially beneficial because they can reduce both energy and capacity costs for electric utilities. The challenge of load management is to reduce these costs without imposing undue restrictions on the conditions of service, for example, on the maximum periods of service interruptions.

To meet this challenge, the Federal Energy Administration (now part of the Department of Energy) in 1975 began a cooperative rate demonstration program with public utility commissions and private utility companies. The options tested included load management policies both direct and indirect, such as interruptible service, and indirect, such as time-of-use pricing. Although the sixteen rate demonstration projects emphasize time-of-use rates for residential customers, other alternatives are also being studied, for example, flat rates, inverted block rates, and direct load management options. In addition, several load management policies for industrial and commercial customers have already been tested or are scheduled for testing.

The rate demonstration program aims primarily to provide (1) a basis for evaluating the administrative feasibility and customer acceptance of innovative electric rates and (2) a substantial data base for estimating the effects of such rates on class load patterns. These potential effects are important to both utilities and utility regulators because implementing load management policies can affect peak capacity requirements, the optimal mix of generating units, and utility revenues.

Scope

As part of the rate demonstration program, individual project teams are preparing final reports, some of which include the estimated effects of the experimental load management policies. These evaluations will provide an initial investigation of all policies tested in each study. The final reports from the projects, however, will vary considerably because of differences in interests, background, approaches, and resources of the project teams.

Realizing this diversity, the Department of Energy (DOE) in 1977 began a comprehensive program to analyze the test data from the residential component of the rate demonstration projects. The analysis is being carried out in large part under a contract with the Research Triangle Institute. This project, expected to continue into 1980, started with the development of an Analytical Master Plan which now guides the comprehensive analysis of the demonstration data [Research Triangle Institute 1978].

This paper is limited to a summary of the analysis of experimental time-of-use electricity prices as applied to residential customers in the Connecticut, Ohio, and Arizona demonstration projects. It presents only the core of the results and does not attempt any detailed explanation of the analysis procedures.

The results, of course, depend on the restrictive design characteristics of the experiments analyzed and, perhaps most significantly, apply only to those customers who were eligible to participate and who, when offered time-of-use rates, would volunteer. The reliability of the results also is hampered by limitations in the experimental designs, sample designs, and data files from each study.

The following section summarizes the RTI analysis approach. The next section broadly describes the results of the Connecticut, Ohio, and Arizona studies.

Analysis Procedures

Because of the manner in which the experimental rates were applied in the DOE experiments, RTI is using the two major types of analyses: comparative and demand. Comparative analysis (analysis of variance and covariance) is used if inferences are limited to the particular variable or set of rates (treatments) actually applied in the experiment. Such limitations may be imposed by either the experimental design or the particular analytic objective. Demand analysis is used if the experimental design enables inferences to be made about rates that were not specifically applied to customers in the experiment. In either case, when enough observations are available for a particular rate (or rates) to provide minimal statistical reliability, simple (noninferential) displays of various load-characterizing descriptive statistics are provided.
Comparative Analysis

For the comparative analysis, we first examine graphic displays of average daily residential load curves. Table 1 shows the four average daily load curves that are being developed for each experimental rate administered to 40 or more customers and for the control group. These load curves are estimated for at least six calendar months per year and cover those seasons pertinent to the particular utility.

Table 1. Average Daily Load Curves

<table>
<thead>
<tr>
<th>Type</th>
<th>Averaging period</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Day of monthly system peak</td>
</tr>
<tr>
<td>2</td>
<td>Weekend days during the month</td>
</tr>
<tr>
<td>3</td>
<td>Weekdays during the month</td>
</tr>
<tr>
<td>4</td>
<td>All days during the month</td>
</tr>
</tbody>
</table>

The next major procedure involves statistically comparing responses observed under one pricing scheme with those observed under another. Our objective is to determine, for some population of customers, whether the level of response for a particular variable depends on the treatment or rate being applied. Our approach in determining whether a response under one rate differs significantly from that under another is based on the concept that there exists a definable population about which inferences from samples can be made. This concept implies that weights must be used to expand the sample to a finite population. If meaningful population cannot be defined or if weights are not available, then unweighted sample data are analyzed. The problem then is to make valid inferences about some population larger than the sample itself.

After we identify the population of inference, we estimate the population variables that are listed in Table 2. For each selected treatment, our estimates include the population means and their associated variances, which are computed directly using sample weights. These estimates, coupled with distributional assumptions, are then used to test whether significant differences exist between the true population parameters for the control group and the parameters for each experimental rate. These variables are computed and compared for at least six calendar months per year (the same six months for which load curves are constructed).

In addition to the variables in Table 2, we also compute estimates of seven other descriptive parameters, listed in Table 3. These parameters cannot actually be compared statistically because each can vary in two ways: from customer to customer and from hour to hour. Procedures for obtaining the variances of such estimates are not available.

Demand Analysis

The demand analysis involves so-called regression techniques to estimate what are known as price elasticities. Basically, these elasticities are estimates of how the time pattern of electricity usage will change
when the time pattern of electricity prices changes. They are estimated using statistical relationships that account for all the factors, demographic and experimental design, that must be held constant to isolate the effects of time-of-use price changes. For example, in the time-of-use rates, many rate design factors can be altered, including, for example, starting time for the peak period and the length of the peak period. Our analysis attempts to measure changes in load curve shapes that are caused only by changes in energy charges given constant facilities charges, starting time of peak period, and so on. For example, we try to determine what might happen to the shape of the average daily load curve if the base price were increased and the peak price decreased from levels included in the rate demonstration project. Our results will be presented as a set of price elasticity estimates that are as detailed as possible given the design of the rate demonstration project.

The general form of our demand analysis model is a simple linear (or loglinear) relationship between electricity consumption variables and several independent variables, including electricity prices and customer characteristics. The model includes so-called interaction terms between prices and other independent variables. These terms help distinguish differences among customers’ responses to price as functions of the types of appliances they own. For example, people with central air conditioners may respond differently from those who do not have them.

Our approach is to estimate models of this form for several depen-

dent variables averaged over different days. Usually, the dependent variables will be kwh consumption during each of the daily rating periods, such as peak, intermediate, and base periods. When a daily rating period is split (for example, when there are two peak periods in a day), consumption during both split portions is used as a dependent variable. These variables are developed for each customer for both the day of system peak within each season and for the average weekday within the two or three months of the summer and/or winter peak season. For example, if a year-long experiment has single daily peak and base periods, eight regressions are estimated, one for each daily rating period for both the peak day and the average weekday of each of the two seasons.

Also, for some of the rate demonstration projects that include several time-of-use rates, the consumption data from all test customers are combined. This increases the number of daily time intervals that can be used to define the kwh consumption variables and increases the number of price variables included on the right-hand side of the regression equation (see Research Triangle Institute 1978, pp. 25–27, for details). Essentially, this approach is designed to facilitate more precise estimates of residential customer responses to variations in time-of-use prices.

Both the comparative and demand analyses are subject to the constraints of the experimental designs, the sample designs, and the data files from each of the projects analyzed. The specific constraints that applied to the three studies reported here are noted as limitations in the next section. For a fuller discussion of these limitations, see A. K. Miedema and C. A. Clayton (1978).

Findings

In this section we summarize the characteristics and analysis results for the Connecticut, Ohio, and Arizona studies.

The Connecticut Experiment

The Connecticut time-of-use pricing experiment began in October 1975. It was a one-year controlled experiment in which 200 Connecticut Light and Power Company residential customers were placed on a seasonal time-of-use rate. An equal number of customers on the existing rate served as a control group. The rate utilized three (peak, intermediate, and base) rating periods during weekdays and two (intermediate and base) rating periods during weekends and holidays. The peak, intermediate, and base prices were 16 cents, 3 cents, and one
cent per kwh. (See White et al. [1978] for a fuller report on the Connecticut project.)

The experiment was designed principally to determine customer acceptance of a time-differentiated pricing scheme and customer ability to alter usage patterns (in terms of conserving electricity or shifting usage from one time period to another) in response to a time-of-use rate structure.

Under the sponsorship of the Department of Energy, the study was conducted through the joint efforts of the Connecticut Public Utilities Control Authority, the Connecticut Department of Planning and Energy Policy, the Connecticut Office of Consumer Counsel, the Connecticut Light and Power Company, and Northeast Utilities.

The population of customers from which the sample was drawn consisted of all 1975 CLP nonfarm residential customers who had a twelve-month billing history and did not have estimated bills in 1975. This sampled population (numbering 296,776) was stratified into five groups according to 1975 annual consumption. Simple random sampling within each stratum produced the desired number of participants. All experimental and control customers volunteered to participate. Customers on the experimental rate were given a lump-sum payment prior to the start of the test. The amount ranged from $50 to $150, depending on the level of consumption.

Limitations on the Scope of the Analysis. During the sampling process, some customers selected to participate were not included for various reasons. (Meter installation problems, refusal, seasonal customers, and bad credit are typical examples.) Based on the combined number of rejections for the experimental and control groups, the number of residential customers represented in the actual study sample is estimated at 134,788, or approximately 45 percent of the sampled population. This group is defined as the effective population because it represents the population to which (in our opinion) valid inferences can be drawn from the sample results. Inferences to the sampled population (or to any population larger than the effective population) require the assumption that excluded customers will respond to the experimental rate in the same manner as those represented in the sample. Since this assumption was not made, all test results shown herein are applicable only to the effective population.

Summary of Analysis Results. Two methods of analysis were chosen for the data. The first characterized the electricity consumption patterns of the sample by several parameters and then statistically compared differences in estimates of these parameters between the control and experimental groups. The second method computed various types of load curves that provided visual comparisons of the consumption patterns of the two groups. Figures 1 and 2 illustrate four representative load curves.

![Figure 1. Connecticut Day of System Peak Load Curves](image-url)
The major results of the study, summarized in the following observations, are considered applicable only to the population of residential customers represented in the sample. This population is estimated to contain approximately 135,000 customers, whereas the number of customers in the sample population was 296,776.

Table 4 summarizes tests of treatment differences for all months analyzed and identifies parameters for which the experimental time-of-use rate response differed significantly from the control group response. Entries in this table correspond to the 20 parameters of Table 2 and are shown as percentage changes for those parameters which are considered significant by tests of hypotheses. Blank cells indicate that the associated differences were not statistically significant.

Several patterns are evident in Table 4. For example, some parameters show significant differences only in the summer months, whereas significant differences show up only in the winter months for others. These patterns may be summarized as follows. First, during peak periods, the experimental group consumed significantly less electricity than did the control group. For example, Table 4 shows significant differences of 14 to 39 percent. This result was evident for each month. Second, during the intermediate period in the summer months, the experimental customers used significantly (from 14 to 22 percent) less electricity than did the control customers. During the winter months, however, the two groups used about the same amount of electricity during the intermediate period. These results suggest that incorporating intermediate periods into time-of-use rates averts needle peaking problems. Third, the experimental group used significantly more electricity in the base period during the winter months than did the control group. During the summer months, however, the two groups consumed about the same quantity of electricity. Fourth, customers in the experimental group used significantly less electricity than did their counterparts in the control group during the summer months (due to the peak and intermediate patterns mentioned above). Joint consideration of peak, intermediate, and base period behavior suggests that during the summer months the experimental customers conserve energy, whereas during the winter months these customers shift usage from the peak period to the base period. These same patterns were evident for weekdays and weekend days within a month. Fifth, diversified demand at the time of system peak (parameter 13) was significantly lower for the experimental group as compared to the control group. The associated load factor (parameter 20) was significantly higher in each month for the experimental group. Sixth, for the peak winter month (January), the estimated effect of implementing the experimental rate on the effective population was a reduction in the system peak of 8 to 13 megawatts. For the peak summer month (June), a reduction of 70 to 85 megawatts was estimated.
Table 4. Significant Differences between Control and Experimental Groups in the Connecticut Study Expressed as Percentage Change

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Dec. 72</th>
<th>Jan. 76</th>
<th>Feb. 76</th>
<th>Mar. 76</th>
<th>Apr. 76</th>
<th>May 76</th>
<th>Jun. 76</th>
<th>Jul. 76</th>
<th>Aug. 76</th>
<th>Sep. 76</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak day kWh</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermediate before peak</td>
<td>-22.5</td>
<td>-14.4</td>
<td>-22.0</td>
<td>-20.9</td>
<td>-20.0</td>
<td>-22.8</td>
<td>-28.8</td>
<td>-50.6</td>
<td>-22.1</td>
<td></td>
</tr>
<tr>
<td>One hour before peak</td>
<td>-18.0</td>
<td>25.1</td>
<td>23.8</td>
<td>19.6</td>
<td>18.5</td>
<td>26.0</td>
<td>28.5</td>
<td>26.5</td>
<td>21.1</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>-14.3</td>
<td>-15.3</td>
<td>22.8</td>
<td>-11.6</td>
<td>-15.3</td>
<td>22.8</td>
<td>28.5</td>
<td>-18.1</td>
<td>-14.5</td>
<td></td>
</tr>
<tr>
<td>Average day kWh</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermediate before peak</td>
<td>-29.2</td>
<td>-16.7</td>
<td>-21.5</td>
<td>-25.4</td>
<td>-25.4</td>
<td>-25.4</td>
<td>-24.8</td>
<td>-22.4</td>
<td>-19.2</td>
<td>-24.8</td>
</tr>
<tr>
<td>One hour before peak</td>
<td>-12.9</td>
<td>14.6</td>
<td>20.6</td>
<td>14.1</td>
<td>15.6</td>
<td>14.8</td>
<td>15.8</td>
<td>13.3</td>
<td>14.5</td>
<td>14.5</td>
</tr>
<tr>
<td>Total</td>
<td>-22.0</td>
<td>-12.9</td>
<td>-12.9</td>
<td>-12.8</td>
<td>-12.8</td>
<td>-12.8</td>
<td>-12.8</td>
<td>-12.8</td>
<td>-12.8</td>
<td>-12.8</td>
</tr>
<tr>
<td>Demand during peak</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermediate before peak</td>
<td>-32.0</td>
<td>-12.2</td>
<td>-12.2</td>
<td>-12.2</td>
<td>-12.2</td>
<td>-12.2</td>
<td>-12.2</td>
<td>-12.2</td>
<td>-12.2</td>
<td>-12.2</td>
</tr>
<tr>
<td>One hour before peak</td>
<td>-32.0</td>
<td>-12.2</td>
<td>-12.2</td>
<td>-12.2</td>
<td>-12.2</td>
<td>-12.2</td>
<td>-12.2</td>
<td>-12.2</td>
<td>-12.2</td>
<td>-12.2</td>
</tr>
<tr>
<td>Total</td>
<td>-32.0</td>
<td>-12.2</td>
<td>-12.2</td>
<td>-12.2</td>
<td>-12.2</td>
<td>-12.2</td>
<td>-12.2</td>
<td>-12.2</td>
<td>-12.2</td>
<td>-12.2</td>
</tr>
<tr>
<td>Nonresidential kWh</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average before peak</td>
<td>12.4</td>
<td>31.4</td>
<td>49.4</td>
<td>37.4</td>
<td>37.4</td>
<td>37.4</td>
<td>37.4</td>
<td>37.4</td>
<td>37.4</td>
<td>37.4</td>
</tr>
<tr>
<td>Average one hour before peak</td>
<td>-10.3</td>
<td>15.8</td>
<td>21.9</td>
<td>15.8</td>
<td>15.8</td>
<td>15.8</td>
<td>15.8</td>
<td>15.8</td>
<td>15.8</td>
<td>15.8</td>
</tr>
<tr>
<td>Average total</td>
<td>-3.5</td>
<td>8.3</td>
<td>14.5</td>
<td>8.3</td>
<td>8.3</td>
<td>8.3</td>
<td>8.3</td>
<td>8.3</td>
<td>8.3</td>
<td>8.3</td>
</tr>
<tr>
<td>Case load factor (#/121 #/131 x 12)</td>
<td>29.8</td>
<td>27.8</td>
<td>27.8</td>
<td>27.8</td>
<td>27.8</td>
<td>27.8</td>
<td>27.8</td>
<td>27.8</td>
<td>27.8</td>
<td>27.8</td>
</tr>
</tbody>
</table>

Note: Percentage change is calculated as $100 \times \frac{E - P}{P}$, where $E$ and $P$ are means for experimental and control groups, respectively. Values are shown only for those cases deemed significant at the .05 level. See Table 2 for parameter definitions. NA indicates not applicable.

The Ohio Experiment

The Ohio time-of-use pricing project began in June 1976. It was an eighteen-month controlled experiment in which 100 of Dayton Power and Electric Company's residential customers were placed on a seasonal time-of-use rate. Two rating periods (peak and base) were established during weekdays, and a single rating period (base) during weekends and holidays. During the spring and fall, the peak rate was 5.1 cents per kWh, and the base rate was 0.4 cent per kWh; during the summer and winter, they were 9.1 cents and 0.4 cent respectively. Sixty customers on the existing rate served as a control group. (See White et al. [1978] for a full report on the Ohio project.)

Under the sponsorship of the Department of Energy, the study was conducted through the joint efforts of the Public Utilities Commission of Ohio and the Dayton Power and Electric Company.

The population from which the sample was drawn consisted of all residential customers from April 1974 to April 1975 after excluding (1) customers without a twelve-month billing history, (2) bulk-metered residences (apartment houses, trailers), (3) company employees, (4) residents on "frozen" rates, and (5) customers with less than 6,000 kWh annual usage. This sampled population (approximately 196,000) was then stratified according to annual consumption from April 1974 to April 1975. One hundred strata were formed with an equal number of customers in each. Random sampling was employed within each stratum to obtain one customer willing to be placed on the experimental rate. A similar procedure was followed for the control group, except that only 60 strata were formed. All participants volunteered, and they received no direct compensation.

LIMITATIONS ON THE SCOPE OF THE ANALYSIS. Of the customers selected in the sample, about 80 percent were rejected. Some of the major reasons for the extremely high rejection rate included refusal to participate, moving, etc. If meter inaccessible. Based on the combined number of rejections from the experimental and control groups, the number of residential customers represented in the actual sample (the effective population) is estimated at 39,053, or approximately 20 percent of the sampled population. We did not assume that excluded customers would respond to the experimental rate in the same manner as those represented in the sample; hence, all test results shown herein are applicable only to the effective population.

SUMMARY OF ANALYSIS RESULTS. Our report covers analyses made on the following select months: August, September, and December 1976, and January, May, and July through September 1977. For each month, we estimated the population parameters of Table 2 for the
effective population for the experimental and control rates (that is, assuming these rates were applied to the effective population). These estimates were compared to determine whether the true population parameters for the two groups differ statistically. The average daily load curves of Table 1 were also developed for each month. Figures 3 and 4 illustrate representative load curves.

The major results of the study, summarized in Table 5, are considered applicable only to the 39,000 residential customers represented by the sample. The patterns of the results shown in Table 5 are summarized below. First, during the six-hour peak period of high prices, customers on the time-of-use rate used significantly less electricity as compared to customers on the existing (declining block) rate. The reduction in usage was relatively consistent over the period examined (which encompassed all seasons). Monthly reduction in usage during the peak period ranged from 21 to 38 percent, or 1.73 to 2.62 kWh per customer per day. Second, some shifting of usage from the peak to the off-peak period occurred during the base period, as evidenced by an increase in usage during the base period in seven of the eight months examined. Statistically, these increases were not significant; however, such small shifts in usage are not detectable with the present sample sizes and the large variation in usage from customer to customer observed during the base period. Third, concerning total consumption, during three of the eight months the customers on the time-of-use rate used slightly more electricity than those on the existing rate; during the other five months they used slightly less. In no case were the differences in usage statistically significant. Fourth, diversified demand at the time of monthly system peak could be computed for only those two months for which system load data were provided. In those two months, diversified demand was 30 percent or more lower for the customers on time-of-use rates.

The Arizona Experiment

The Arizona rate demonstration project began in 1975 and involved experimental evaluations of time-of-use rates both alone and in combination with load control devices or with special educational efforts by the staff of the participating utility, the Arizona Public Service Company (APS). The analysis completed by the Research Triangle Institute deals only with data from customers on the time-of-use rates. (See Miedema et al. [1978] for a fuller report on the Arizona project.)

Three groups of experimental time-of-use rates were offered to residential customers in the Yuma and Phoenix metropolitan areas. These two areas account for most of the summer peaking on the APS
Figure 4. Representative Load Curve, Peak Summer Day

Table 5. Significant Differences between Control and Experimental Groups in the Ohio Study Expressed as Percentage Change

```
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Parameter number</th>
<th>Aug. 76</th>
<th>Sep. 76</th>
<th>Dec. 76</th>
<th>Jan. 77</th>
<th>May 77</th>
<th>July 77</th>
<th>Aug. 77</th>
<th>Sep. 77</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak day kWh</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peak</td>
<td>1</td>
<td>-39.7</td>
<td>-26.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermediate</td>
<td></td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base</td>
<td></td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One hour before peak</td>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One hour after peak</td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average day kWh</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peak</td>
<td>2</td>
<td>-35.0</td>
<td>-35.2</td>
<td>-24.7</td>
<td>-27.9</td>
<td>-23.6</td>
<td>-27.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermediate</td>
<td></td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base</td>
<td></td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One hour before peak</td>
<td></td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One hour after peak</td>
<td></td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diversified demand</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Noncoincident maximum demand</td>
<td></td>
<td>13</td>
<td>-40.2</td>
<td>-29.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average hourly kWh — 80%</td>
<td></td>
<td>15</td>
<td>-19.6</td>
<td>-22.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average hourly kWh — 90%</td>
<td></td>
<td>16</td>
<td>-34.4</td>
<td>-39.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average hourly kWh — 95%</td>
<td></td>
<td>17</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#1/#6</td>
<td></td>
<td>18</td>
<td>-37.0</td>
<td>-25.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class load factor (#12/#15 × 24)</td>
<td></td>
<td>20</td>
<td>60.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

Note: Percentage change is calculated as 100(\(\bar{Y}_e - \bar{Y}_c\)), where \(\bar{Y}_e\) and \(\bar{Y}_c\) are means for experimental and control groups, respectively. Values are shown only for those cases deemed significant at the .05 level. See Table 2 for parameter definitions. — indicates zero data.
1975 while they remained on the conventional APS residential rate. Then, during May through October 1976, all participants were billed under time-of-use rates. The rate schedules were adjusted, however, to ensure that no household paid more than its previous year's bill. Rather, if a household's total usage and time pattern of usage were identical to those of corresponding months of the previous year, the customer was assured of a bill 30 percent lower than under the prevailing conventional rate. Any divergence from each household's previous year's usage within each of the three daily rating periods was billed or credited at the variable kWh charges specified by one of the 28 unique time-of-use rates.

Since five customers were placed on each of the 28 rates, a total of 140 volunteers were involved. Of these 28 rates, the 16 Group I rates incorporated a three-hour peak period; the 6 Group II rates, a five-hour peak period; and the 6 Group III rates, an eight-hour peak period. Since the same clock hours (9:00 a.m. to 10:00 p.m.) were covered by both the peak and intermediate periods for all three groups, the lengths of the intermediate periods were 10, 8, and 5 hours for Groups I, II, and III, respectively. All other hours were designated as base hours.

LIMITATIONS ON THE SCOPE OF THE ANALYSIS. The 140 participants in the experiment were among those eventually retained from simple random presamples of the Phoenix and Yuma service areas. At the time the presamples were selected, these service areas contained 165,000 and 16,000 households, respectively, or about 63 percent of the total 286,000 residential accounts in the APS system. A substantial number (59 percent) of those in the presample were rejected — 18 percent because their billing histories were shorter than one year, and 41 percent because of either inaccessible service entrances or insufficient space to accommodate the time-of-use meter. Because of these restrictions, the results, strictly speaking, apply only to those Phoenix and Yuma residential customers who would have volunteered for the study and would not have been excluded for the above reasons. Thus, the nature of the sample finally selected for the Arizona project limits the population to which these results can be projected accurately.

The sample data were further pared by editing out test customers who did not report certain demographic data and others who (as a result of the peculiar definition of the time-of-use rates) were actually billed under the existing declining block rates. The sample of customers on the time-of-use rates that was actually analyzed was about 70 percent of its original size, depending on the particular month(s) and group of customers involved.

SUMMARY OF ANALYSIS RESULTS. Two types of analysis were completed. A comparative analysis sought to determine whether time-of-use rates, as compared with conventional rates, created a difference in household consumption of electricity. A regression analysis (of variations in demand) attempted to measure the elasticity of demand response to marginal prices (that is, the percentage change in consumption of electricity corresponding to percentage changes in the marginal time-of-use prices incorporated in the experimental rates).

The comparative analysis was made of time-of-use consumption for customers in Group I, the group that in both 1975 and 1976 had identical lengths of time in the peak, intermediate, and base periods (late afternoon, morning and evening, and night). These comparisons reflect the "average" time-of-use rate effect in that the analysis averages over all 16 sets of time-of-use rates administered to Group I customers.

This analysis fairly convincingly shows that when customers in Group I were on time-of-use rates, they reduced their electricity consumption during the peak rating period by 6-17 percent and in the intermediate period by about one to 9 percent, depending on the month of analysis. To a less significant degree, the evidence indicates that some of this electricity consumption was shifted to the base period and that overall consumption declined slightly.

These results are presented graphically in the four panels of Figure 5, which provides plots of average daily kWh usage for each of the three daily rating periods for 1975 and 1976, together with average total daily kWh usage. These charts reflect both the 6-17 percent reductions in peak period kWh usage from June to October and the reductions in intermediate period usage in each of the months. Finally, the lack of significance in either the average increase in base period consumption or in the average decrease in overall electricity consumption is apparent from the small betweeen-year differences shown in the corresponding panels of Figure 5.

By contrast, the regression analysis failed to detect any significant price elasticities for peak, intermediate, or base period consumption; that is, there was no significant change in kWh consumption as time-of-use prices were varied. For example, Table 6 shows estimated price and income elasticities of demand for each of the three rating periods. The particular estimates given there are average estimates for July through September 1976 for the Group I customers, but they are characteristic of the results for individual months and for Group II and III customers as well. In an appendix to the Arizona report, we used data identical to those in Table 6 to estimate a more restrictive model.
For that model we did find "significant" elasticities, as have other authors [for example, Atkinson 1979], using restrictive models similar to theirs in analyzing the Arizona data. Our conclusion was that "sig-
ificant" in these restrictive models is very likely attributable to the strong assumptions imposed on the data in model estimation.

There may have been many reasons for observing between-year differences in peak and intermediate consumption in the comparative analysis and yet nonsignificant elasticities. However, one possible explanation is the substantial emphasis that the billing scheme, as conveyed in the customer education literature, placed on the discount formula and its association with peak, intermediate, and base times of the day, as opposed to the exact cost of peak, intermediate, or base kwh. Another is the relatively small number of observations which hampered our ability to detect significant elasticities even if they existed.

Notes

1. Cost-benefit analysis is a commonly accepted criterion for evaluating the appropriateness of conditions of service.

2. In 1975, ten projects were initiated in Arizona, Arkansas, Connecticut, California (Los Angeles), New Jersey, Ohio, Wisconsin, Vermont, Michigan, and New York. Subsequently, six projects were added in California, North Carolina, Oklahoma (Edmond), Puerto Rico, Rhode Island, and Washington.

3. For projects that involve a single experimental rate and control group, all types of load curves will be presented regardless of sample size.

4. An alternative for estimating these variables is to conduct an analysis of covariance. This procedure involves the estimation of a linear model which,
Time-of-Use Electricity Prices

for example, explains variation in the parameters of Table 2 as a function of several independent variables called covariates, such as income and appliance wattages. This procedure will be considered for projects in which the stratification used in the sample design suggests that covariates would help reduce the experimental error.

5. Some of the parameters could not be estimated because system load data were not available for all months.

References


A Benefit-Cost Framework for Evaluating New Pricing Concepts and Rate Structure Reform

Robert L. Borlick

The current interest in rate structure reform in this country began about four years ago, roughly corresponding with the entry of the postembargo period, which brought with it dramatic increases in electricity bills and concomitant declines in the financial positions of most investor-owned utilities. Thus, it was no coincidence that different interest groups, each with divergent objectives, viewed rate structure reform as an opportunity to alleviate their particular problems or to advance their positions. Consumers wanted lower bills; utilities wanted improved earnings and smaller capital outlays; environmentalists wanted less new plant construction; the federal government wanted lower oil imports.

The campaign for marginal cost-based rates began in a few progressive state regulatory commissions (principally New York, California, Wisconsin, and Massachusetts) and subsequently spread throughout the nation. At the federal level, interest in retail rate structure reform has culminated in the signing into law of the Public Utility Regulatory Policies Act of 1978 (PURPA).

Despite the extensive amount of attention dedicated to the subject, the debate continues over whether the benefits offered by new rate structures exceed the costs of implementation. This paper develops a framework for evaluating the net benefits of rate structure reform.
based on three goals identified in PURPA. While the effect is not intended to resolve the current debate, it is hoped that it will at least illuminate the nature of the problem.

In the next section, three goals of rate reform are defined. Quantitative criteria are then derived for use in evaluating new rate designs in terms of their contributions to these goals. These evaluation criteria are conceptually applied to the segments of the economy primarily affected by the new rate designs. Based on the insights gained in that exercise, the limitations to estimating the benefits of rate reform are subsequently examined. Finally, the main points of the paper are summarized and some conclusions drawn.

Before proceeding, the author wants to make clear that this paper represents his personal views. While these views do not conflict with current Department of Energy (DOE) policy, neither do they represent that policy.

**Goals of Rate Structure Reform**

In order to evaluate the benefits offered by a specific rate structure, it is necessary to determine the goals of rate structure reform and to develop appropriate criteria for measuring the contributions of the proposed rates to those goals. The recently passed PURPA provides a good starting point. Section 101 of the act states the following three purposes: (1) Conservation of energy supplied by electric utilities; (2) The optimization of the efficiency of use of facilities and resources by electric utilities; (3) Equitable rates to electric consumers.

Based on the language of the related Conference Manager's report and on extensive discussions within DOE, the author has made the following observations regarding these purposes, or goals.

**Conservation**

The conservation goal is satisfied when each kwh of electricity is consumed in an end use which is valued by each customer at least as highly as the resource cost to the economy of making that kwh available to him. This goal is promoted when the customer faces price signals which make him recognize the underlying resource costs of his decision to consume electricity, thereby encouraging him to use it in efficient, nonwasteful end uses. Of course, this implies that some end uses which are currently being served may not be sufficiently valuable to warrant continuation. But the definition also recognizes that conservation means more than just using less; that energy consumption creates value and that such usage should occur, indeed should be encouraged, so long as the value created at least equals the resource cost of production.

The above notwithstanding, it is anticipated that rate structure reform will encourage more efficient allocation of the nation's energy (and other economic) resources and most likely will reduce overall consumption of energy.

**Efficiency**

A narrow reading of the language in PURPA would interpret efficiency to mean technical efficiency — the utility's engineering and managerial effort to meet a given demand for electricity at the lowest resource cost to the nation. This encompasses such factors as constructing the optimal mix of new generating plants, avoiding the construction of excess or insufficient capacity, and substituting cheap fuels, such as coal, for expensive fuels, such as oil.

This narrow interpretation ignores the question of whether the level and time pattern of electricity demand is consistent with both its value to the customer and the resource cost of producing it. Thus, it can be argued that the utility's use of facilities and resources is not optimal and efficient unless such consistency exists.

The retail rate structure is the factor which links the customer's economic values to the utility's cost structure; furthermore, rate policy is the subject of Title I of PURPA. Thus, this broader interpretation of PURPA equates efficiency, as defined in Section 101 of the act, with economic efficiency.

**Equity**

Unfortunately, PURPA provides little guidance as to exactly what is meant by equity. However, it is the author's view that, to the extent practicable, rates should reflect costs imposed on the utility by the customer. This implies that subsidies should not occur among different customers, utility stockholders, and the public in general.

Many would argue that the current distribution of income is inherently unfair; hence, some such subsidies may be justified. But even accepting this, the logical point of departure for determining what subsidies ought to be provided (if any) is a cost-based rate structure. Without such a benchmark there is no basis for designing those subsidies such that they minimize the adverse impact on economic efficiency.

A second view of equity, reflected in the Conference Manager's report, is that rate structure reform should not impose undue economic hardships on rate payers through sudden and extreme shifts in their rates. This dimension of the equity goal is consistent with the
future accountability aspect of cost-based rates, while at the same time recognizing that such rates might unduly penalize some customers for past investments, or other decisions, made when the rules of the game were different.

The author has nothing further to offer regarding the equity goal except to suggest that decisions involving equity judgments are best carried out at the levels of government closest to the people affected by them, thereby allowing such decisions to be more responsive to local attitudes and political desires.

Conflicts among the Goals

The three goals of rate structure reform do not inherently conflict, since they share the common characteristic of embracing cost-based rates. To the extent that the equity goal is redefined, it may be in conflict with the other two for some proposed rate designs. More specifically, some rate structures which increase economic efficiency will not equitably distribute the net gain in benefits.

Some Consequences of Rate Structure Reform

At this point it is worth digressing briefly to consider some of the consequences of implementing rate structures that better reflect the resource costs of production. More specifically, implementation of such rate structures is likely to accomplish the following: improve utility load factors; slow the growth in peak electricity demand; increase the utilization of utility plant; reduce the need for utility capital outlays; reduce utility oil and natural gas use; and reduce the average price of electricity.

While each of these results has merit, as evidenced by the fact that they have their own respective constituencies, their collective value is derived from the contributions they make to one or more of the goals of conservation, efficiency, and equity.

Criteria for Evaluating Rate Structures

The three goals described above should be simultaneously achieved when electricity prices reflect their respective resource costs of production. Indeed, the conservation and efficiency goals are merely different dimensions of economic efficiency; thus, only one criterion is needed to measure a specific rate design's contribution to these goals. The equity goal necessarily involves considerable judgment, and it is probably not feasible to develop a satisfactory unidimensional evaluation criterion. Still, quantification of the distributional effects of a proposed rate design should be useful in clarifying the nature of the judgments to be made. Each of these evaluation criteria is discussed below.

---

**Economic Efficiency Criterion**

For any fixed level of productive resources, the generally accepted measure of economic efficiency is the sum of consumer's plus producer's surplus. This concept is illustrated in Figure 1, which shows the demand and marginal cost functions for some arbitrary economic good.

The demand function represents the value that consumers place on the next unit of the good at any level of consumption. Thus, the first unit is worth $P_a$ to some consumers, whereas all units in excess of $Q_{max}$ have no positive economic value. To maximize his satisfaction, each customer should continue to purchase additional units of the good until his highest valued use for the next unit is less than the price which he must pay for it. In Figure 1 this occurs at a price equal to $P^*$, which means that the quantity $Q^*$ would be consumed. At that level of consumption, the total monetary equivalent value derived by consumers

**Figure 1. Illustration of Consumer's and Producer's Surplus**

- **Price**
- **Consumer's surplus**
- **Producer's surplus**
- **Resource cost**

**Marginal cost function**

**Demand function**

**Quantity demanded**
from the economic good is approximately represented by the area under the demand function between the vertical axis and the dotted line at $Q^*$ (the combined areas of the striped, stippled, and cross-hatched portions). Since the consumer paid to the seller the sum $PQ^*$, the consumer’s collective net benefit from the transaction is approximately represented by the striped triangular area, to denote the consumer’s surplus. Thus, the consumer’s surplus approximates the total monetary value by which an economic good consumed exceeds the total revenue paid by consumers in purchasing it. Another way to illustrate this is to say that consumers would have to receive individual payments totaling a dollar amount approximately equal to the aggregate consumer surplus in order to be indifferent to a price increase beyond $P_c$ (in which case they would choose not to buy any of the good). The accuracy of this approximation is discussed further in the Appendix.

Turning now to the producer side, the marginal cost function shown in Figure 1 represents the cost to the producer of making available the next unit of the economic good at any level of output. Thus, the first unit costs $C_w$ and the cost increases with output. When a producer faces a fixed price for all his units of output, as does a regulated utility, it is in his interest to provide consumers with additional units of the good until the marginal cost of the last unit is equal to the market price. This point occurs at the intersection of the demand and marginal cost functions. At that level of output, $Q^*$, his cost of production is represented by the crosshatched area under the supply function between the vertical axis and the dotted line at $Q^*$. The total revenue he receives is represented by the rectangle consisting of the stippled and crosshatched areas. His profit (equal to revenue less costs) is thus represented by the stippled area alone and is denoted the producer’s surplus. Thus, the producer’s surplus is the amount by which the revenue received from the sale of an economic good exceeds the seller’s cost of production, or acquisition.

The change in the sum of the consumer’s and producer’s surplus is an approximate measure of the net benefits accruing in a single period, for example, one year. This static measure can be made dynamic by estimating the time series of such annual net benefits over an appropriate planning horizon, then discounting the terms in the series to obtain the net present value of the entire stream of benefits. The result of this calculation is the net present value of the economic gain accruing to the group of consumers and producer(s) of the economic good.

The fact that implementation of a new rate structure causes one consumer (or producer) to experience a gain in his time-discounted consumer’s (or producer’s) surplus which exceeds another’s loss does not necessarily mean that the collective economic well-being of the group has increased over the status quo. But it does mean that the potential exists for each to be better off, since the “market basket” to be divided among them is now larger. Exactly how this gain should be divided requires an equity judgment.

**Equity Criterion**

An equitable distribution of the increase in time-discounted consumer’s and producer’s surplus is very subjective. Nonetheless, there is considerable value in initially assuming that equity is largely served when each customer pays the full cost which he imposes on the utility through his usage of electricity. The distribution of benefits which results from such a cost-based rate structure is quantifiable in that the customer population can be subdivided into classes and the change in time-discounted consumer’s plus producer’s surplus estimated for each. The resulting distribution of these benefits explicitly reveals the winners and losers under a purely marginal cost-based rate structure. While such an analysis is more revealing, and therefore more valuable, as the level of customer disaggregation increases, it has substantial value even if performed only for the existing residential, commercial, and industrial customer classes.

The distribution of benefits can next be modified to reflect other concepts of equity, such as protection against sudden rate increases. What finally emerges is a distributional standard which can be used as the desideratum for new rate structures. Thus, the more closely the distribution of benefits resembles this ideal pattern, the more equitable the proposed rate structure is judged to be.

**Assessing Benefits and Costs**

In assessing the benefits and costs offered by a particular rate design, it is sufficient to estimate the change in consumer’s and producer’s surplus which occurs when moving from existing rates to the proposed rates. This analysis also needs to account for the additional implementation costs as well as for any external benefits (or costs) associated with the new rate design.

Before attempting a practical benefit-cost analysis, it is necessary to decide how much of the economy needs to be examined. In theory, the entire economy should be included since the impact of rate reform will ripple throughout the nation and will even affect the international balance of trade (and, it is hoped, the world oil market). But, to consider minor effects beyond the utility service area, or on products that are neither direct complements, substitutes, nor major inputs to
the production of electricity, probably would not merit the substantial additional effort required. Moreover, to the extent that the affected markets, both within and outside the utility service area, are reasonably efficient and set prices for goods and services close to their marginal resource costs, the change in producer's surplus in those markets will be negligible.

It is true that prices do not approach marginal resource costs for many economic goods, due to the presence of monopolistic/oligopolistic elements, government price regulation, and/or public services being offered at subsidized prices. This divergence of prices and costs gives rise to the "second best" distortions frequently cited in regulatory proceedings by opponents of rate reform. To the extent possible, these distortions should be accounted for when they appear to be significant, which is most likely to be true only for close substitutes or complements within the utility service area, and for the inputs to the production of electric power.

Thus, in broad terms the evaluation of the net benefits of any proposed rate reform should consider each of the following components in the benefit-cost analysis: effect on utility customers; effect on utility stockholders; significant localized externalities; and significant national externalities.

**Effect on Utility Customers**

For the reasons cited above, most of the impact of rate reform is likely to fall upon either the utility customers or the utility stockholders. However, to the extent that the regulatory process is effective, the producer's surplus of the utility will be little affected; thus, essentially all of the net economic benefit (or loss) will be captured by the utility customers. Of particular importance in this regard is that regulatory authorities ensure that the long-run allowed rate of return approximates the market cost of capital, thereby avoiding Averch-Johnson allocative distortions.

To estimate the change in consumer's surplus requires knowledge of the nature of the electricity demand functions in each pricing period over the range of potential change. This is shown in Figure 2, which illustrates the impact of a hypothetical rate design on one customer class. Price \( P_1 \) is the existing rate, assumed to apply in each period. Prices \( P_2 \) and \( P_3 \) are the proposed rates, respectively, for the on-peak and off-peak periods. The relevant ranges of change shown there are between points 1 and 2 on the on-peak demand function and points 1' and 2' on the off-peak demand function; thus, only these portions of the demand functions need to be estimated. The crosshatched area lying under the on-peak demand function is an approximate measure of the direct loss in economic benefits by customers due to higher on-peak rates, while the stippled area approximately measures the gain by those customers due to lower off-peak rates.

Not all of the change in consumer's surplus is captured in the calculation described above because the demand for other economic goods will also change significantly in response to electricity rates. Accordingly, the consumer's surplus associated with these other goods must also be examined. In practice, this investigation can be limited to the close substitutes for electricity, such as natural gas or distillate oil used in space heating, and to the major complements of electricity usage, such as appliances.

Assuming the allowed rate of return is maintained close to the utility's cost of capital through frequent rate adjustment, revenue excesses or deficits generated from rate structure changes will only be temporary. For ease of illustration, it is assumed in Figure 2 that the off-peak rate is set sufficiently low to just offset the increased revenue collected on-peak. In practice, there are other more complex revenue balancing adjustments which promote greater economic efficiency and/or distribute the gains in a more equitable manner.

![Figure 2. Changes in Consumer's Surplus in Response to Moving On-Peak and Off-Peak Rates Closer to Marginal Costs](image-url)
The method chosen for avoiding the collection of excess (or deficient) revenues is itself important to the benefit-cost analysis because it affects the mechanics of calculating the change in consumer's surplus. This is because the rate adjustments required to meet the utility's total revenue constraint have both allocative efficiency and distributional effects. For example, adjustment of fixed/mo customer charges, on the one hand, and prorate adjustments to the peak and off-peak rates, on the other hand, will have significantly different resource allocative and distributional effects. In short, the choice of method for holding the total revenue constant is an important issue; however, it is likely that any reasonable adjustment technique will result in a final rate structure that is far superior to those currently in existence, in terms of meeting the three PURPA goals.

Utility Stockholders

If the utility's long-run average rate of return is effectively constrained to approximate the firm's cost of capital, the utility should be relatively indifferent to rate structure changes. However, it can be argued that the uncertainty over how customers' demand levels and patterns will respond to such rate structure changes will translate into less stable (or at least less certain) utility earnings. But it is also possible that a new rate structure could more closely track changes in short-run operating costs, thereby improving earnings (but not revenue) stability.

If the investment community perceives greater earnings instability (whether real or not), then the likely result of rate structure reform will be to increase the cost of capital, at least until sufficient experience is gained by investors to convince them that the new rate design will not adversely affect earnings. This factor must be viewed as a cost of rate reform (at least in the short run) and must be included in the benefit-cost calculus. If the regulatory authority allows the utility to recover its full costs of service, including an increase in the rate of return required by any increase in the utility's cost of capital, then these higher costs ultimately will be borne by the rate payers and will be implicitly included in the change in consumer's surplus. Moreover, to the extent that the regulatory authority provides such rate level relief in a regular and expeditious manner, that action itself will significantly reduce investor uncertainty over earnings instability. Finally, there are precedents for retroactive adjustments to protect the utility (and the rate payers) from over- or underrecovery. The utility could keep track of any revenue shortfalls that might result from implementing a new rate design, so that the regulatory commission would have a basis upon which to make the company whole. That guarantee, in and of itself, would do much to reduce the earnings uncertainty associated with rate reform.

Localized Externalities

Some of the effects of rate reform occurring within the utility service area (be they benefits or costs) will not be directly imposed upon or accrue to either the utility or its customers — at least not in the form of higher (or lower) costs of production and electricity rates. That is, some of the benefits and costs resulting from rate reform will be external to the utility. Examples of local externalities are potential changes in the degree of environmental degradation and in industrial activity and employment levels resulting from electric rate structure reform.

Despite the fact that many states and federal environmental regulations already impose a heavy compliance burden on the industry, the state of “zero discharge” is not attained today and never will be. Thus, the generation and delivery of electric power results in environmental insults which vary among various technologies. Accordingly, alterations in general levels and mix, in response to rate structure changes, will alter the magnitude of these external costs. One way to internalize residual environmental costs is to impose an excise tax on electricity, as several states have already done. In the absence of such a tax, the environmental impact of rate reform should be recognized, be it positive or negative. The problem is to determine the value of the impact and therefore the appropriate size of the adjustment.

Rate structure reform may also affect both the level and composition of local employment. The direction of these effects will depend largely on the type of rate structure under consideration and the electricity usage characteristics of local industry, including the relative importance of electricity in the production process. In some instances the proposed rate reform could result in plant closings and increased unemployment, or at least in a reduction in the industrial growth and employment of the region. In those instances the regulatory authority may wish to offer price concessions to the affected industrial customers, perhaps in the form of gradual rather than immediate implementation of the new rate structure. However, such concessions can be carried too far, since indefinitely providing local industry with electric power priced below cost is tantamount to subsidizing uneconomic production. In any event, these local employment effects, be they positive or negative, should be included, to the extent practicable, in the determination of the net benefits of the proposed rate reform.
National Externalities

Some important externalities transcend the utility service area in terms of the general public they affect, only one of which will be mentioned here — the nation's importation of oil and natural gas. Because reliance on foreign producers carries with it the risk of sudden supply interruption, the nation incurs a cost which exceeds the price paid by the purchaser. Furthermore, it does not much matter whether the oil or natural gas consumed is of domestic or foreign origin, because if that fuel were conserved it would almost certainly replace imported energy being consumed elsewhere — at least until energy imports drop to zero, a very unlikely event in the foreseeable future. In the benefit-cost calculus the gain to the nation due to reduced energy imports must be added to the net economic benefits offered by a proposed rate design.\(^9\)

Some Practical Limitations

To carry out an assessment of the benefits and costs of rate reform as described in the previous section, it is necessary to estimate portions of the demand functions for electricity and for its close substitutes and complements. Such estimates need only apply across the ranges of potential change in the usage of these goods in response to the new elasticity rates. Thus, at a minimum, estimates of the respective price elasticities and cross-elasticities of demand are required which pertain at the levels of current rates. Unfortunately, even these data are not completely available, thereby precluding sufficiently precise estimates of changes in consumer's surplus.\(^9\)

Given that estimates of changes in consumer's surplus are imprecise, what conclusions can be drawn regarding the net benefits offered by rate structure reform? There are still valid arguments to be made in its favor.

It is an easy theoretical exercise to show that a rate structure which moves rates closer to marginal resource costs will increase the sum of producer's plus consumer's surplus, exclusive of the costs of implementing the new rates. In fact, this measure of net economic benefits should approach its maximum value when rates are set equal to their respective marginal resource costs.\(^11\) Thus, to justify the adoption of a new marginal cost-based rate structure on economic efficiency grounds, it is only necessary to demonstrate that the resulting improvement in consumer's surplus, in the absence of implementation costs, would likely exceed such costs. This is a much easier task than that of explicitly quantifying the net economic benefits offered by the proposed rate structure, because only a lower bound on those benefits needs to be established.\(^12\)

The evaluation framework outlined here for estimating the time-discounted change in consumer's surplus suggests a practical means for establishing a lower bound on the benefits of a specific rate proposal. As more detailed demand data become available, this methodology also has promise as a means of comparing the relative attractiveness of different rate proposals, both in terms of economic efficiency and distributional equity.

Conclusion

Over the past several years much of the rate reform debate in the United States has been conducted without the benefit of a general analytical framework for quantitatively evaluating the benefits of alternative rate structure proposals. Furthermore, the various advocates of rate structure reform have too often held overly narrow views of the purposes to be served. Much of this myopia originates from a perceived need to assess the benefits of rate structure reform in terms of the operational parameters of the utility industry, such as improving load factors or reducing the growth in peak load. As a result, the analytical efforts to date have primarily concentrated on the utility system itself. Certainly, it is important to understand how utility systems will be affected by new rate structures, since these firms must accurately forecast future loads in order to plan their capacity expansion and fuel requirements effectively. But such considerations are merely tangential to the issue of justifying the implementation of new rate structures; only explicit consideration of the effects on the customer side of the meter can do that.

It is ironic that the concepts underlying the benefit-cost framework presented in this paper have long been available as part of the traditional literature of welfare economics.\(^13\) While far from perfect, the methodology outlined here would permit a significant improvement in the assessment of new pricing concepts and rate structure reform.

Many will view this paper as being one more esoteric exposition of economic theory; such is not the case. The dollar-and-cents benefits to be realized through rational assessment of alternative rate structures are substantial. Revenues in the electric utility industry exceeded $70 billion in 1978. Thus, even a relatively small percentage increase in consumer's surplus would represent a large monetary gain in absolute terms. Clearly, caution is required in devising and implementing new rate structures, particularly when time-of-day metering of small customers is involved. But the absence of perfect knowledge regarding the
benefits to be derived from rate structure reform must not be used as an excuse for inaction when the potential gains to the nation are so large.

Appendix

A Few More Words about Consumer's Surplus

In the text of this paper the net change in consumer's surplus is used as a proximate measure of the benefits enjoyed by utility customers due to a change in rate structure. Actually, this measure is not precisely correct; the preferred measure is the sum of the "equivalent variations" of these individual customers.

The equivalent variation is defined as the nominal dollar amount a consumer would have to receive (pay) to make him as well off as he would be if subjected to an increase (decrease) in one or more of the prices of economic goods available to him. This measure, rather than the net change in consumer's surplus, is the precise indicator of an individual's welfare change brought about by a rate structure change. Unfortunately, the equivalent variation cannot be directly derived from the demand functions estimated on the basis of observable data.

In the past, many welfare economists have assumed that a change in consumer's surplus reasonably approximates the corresponding equivalent variation. In his doctoral research, Robert Willig derived the quantitative relationship between the equivalent variation and the change in consumer surplus. His findings [13] are summarized as follows:

If

\[
\frac{A - \eta}{2M_0} \leq .05, \quad \frac{A + \eta}{2M_0} \leq .05
\]

and

\[
\frac{A}{M_0} \leq 9,
\]

then, for a price change,

\[
y |A| \leq \frac{A - E}{2M_0} \leq \frac{\eta |A|}{2M_0}
\]

where:

- \(E\) = equivalent variation corresponding to the price change;
- \(M_0\) = consumer's initial income; and
- \(\eta, \eta\) = respectively, the smallest and largest values of the income elasticity of demand over the region of change.

If the three preconditions described above are not fulfilled, the bounds on \(A - E |A|\) are given by more complex formulas, also derived by Willig [13].

Notes

1. In this example, it is assumed that marginal cost increases with output, which is typically (but not always) true. However, the alternate assumptions of constant or declining costs would unnecessarily complicate the analysis while adding little to the discussion.

2. This is not easy task because the relevant discount rates vary among individuals and businesses, ranging from the interest rates offered by commercial banks on time deposits (currently about 6 percent) to rates exceeding those charged for consumer installment credit (currently about 18 percent). It is important to realize that the utility's cost of capital is not the relevant discount rate, since that only reflects the preferences and investment opportunities of the firm's stockholders, who are not likely to be representative of the population of its customers.

3. This caveat warns against making interpersonal comparisons of individual welfare. To the author's knowledge, there is no sound economic basis for making such comparisons. In theory, consumers who gain could financially compensate those who lose so that everyone is better off. In practice, such compensation is not feasible, thereby forcing one to assess explicitly the distributional effects of rate structure changes.

4. In a recent article, Alfred Kahn [9] took a pragmatic approach similar to that proposed in this paper; his article provided the writer with the initial impetus to develop the benefit-cost framework described here. The validity of this approach was also recognized by William J. Baumol [3, p. 144]: "In fact, there are probably a great many interrelationships within the economy that are weak enough to be ignored. Thus, for all practical purposes, the demand for most goods is likely to be dependent only on the demands for a few other items at least judging by the relatively small number of statistically derived interrelationships that turn out to be significant. It may, then, be possible to partition the economy more effectively than some might have suspected."

5. Perhaps the most renowned article on the theory of second best is that by economists R. G. Lipsey and Kelvin Lancaster [10], in which they conclude: "To apply to only a small part of an economy welfare rules which would lead to a Pareto optimum if they were applied elsewhere, may move the economy away from, not toward, a second best optimum position." Note, however, that the second best problem is circumvented in the benefit-cost methodology described in this paper because the change in welfare is
actually estimated — not merely inferred from the application of normative welfare rules such as setting prices equal to marginal costs.

6. Originally formulated by Harvey Averch and Leland Johnson [1], the basic A-J hypothesis states that a utility has an incentive to inflate its rate base excessively when the rate of return allowed by the regulatory commission significantly exceeds the utility's cost of capital. Subsequent research has revealed that it is sufficient, but not necessary, that the allowed rate of return be set equal to the utility's cost of capital to avoid the A-J effect. Such resource allocation distortions can also be avoided by not strictly enforcing the rate-of-return constraint, thereby allowing the utility to keep some of the savings resulting from efficiency improvements. In practice, the existence of regulatory lag performs this function, as was rigorously shown by Elizabeth Bailey and Roger Coleman [2].

At a more fundamental level, Paul Joskow [8] has questioned the validity of how well the A-J model represents the regulatory process. He concludes that the empirical evidence does not support the A-J hypothesis.

7. The Cost of Service Index employed by the New Mexico Public Service Commission is a good example of such a cost adjustment mechanism.

8. During 1978 this issue arose in Wisconsin when an environmental intervenor group (Wisconsin's Environmental Decade) challenged the implementation of proposed time-of-day rates on the grounds that they would excessively stimulate the construction and operation of baseload nuclear and coal-fired power plants with consequent adverse effects on the environment.

9. The cost to the nation of importing oil exceeds the dollar outlays because of the induced reliance on an uncertain supply. As a result, an implicit risk premium is incurred which is added to the landed price of each barrel of imported oil. This risk premium represents the expected cost of economic disruption due to future supply interruptions. One way to internalize these costs in cost-benefit calculations is to quantify the risk premium and increase the world price of oil by that amount.

10. The primary reason why there are limited data on demand responses to innovative rate structures is the lack of real world experience with these rates. But as rate structure reform occurs, the data base will become richer, permitting improved estimates of the required price elasticities.

11. The cautious tone of this statement is used because many prices do, in fact, diverge from their respective marginal costs; therefore, second-best distortions will occur to some degree.

12. This lower bound on benefits is itself derived by placing lower bounds on both demand elasticities and on the estimates of the equivalent variations associated with the price changes. This latter calculation was performed by Willig [13] and is described in the Appendix.

13. The works of Jules Deupuis [4], A. C. Harberger [5], J. R. Hicks [6, 7], L. M. D. Little [11], Alfred Marshall [12], and R. D. Willig [18] survey the more important advances in welfare economics spanning almost a century of economic thought.

References


Load Research Concerns in Rate Structure Revision

Samuel Behrends, Jr.

The subject of electric utility rate design revision has occupied numerous academic economists to a substantial degree over a number of years, but during the last five years these matters have been churned and thrashed among regulators, legislators, utilities, consultants, and single-interest groups. The tremendous efforts of these five years have been an invaluable exploratory learning process. From the analyses and advocacy have emerged possibilities not heretofore deemed productive but perhaps now worthwhile. From the challenges and denials brought forth to confront ardent proponents, I think they have learned that the paths they advocate often would not bring the panaceas they have sometimes intimated.

I suggest that the churning and thrashing have now been of sufficient depth and breadth that we can begin to separate the wheat from the chaff, to develop strong determination and pursue it, and that the pressures in the public arena require that we do so. The probing, the challenging, and the advocacy have laid bare many flaws in existing rate design and in the proposed solutions.

We must recognize the present lack of two elements fundamental to an orderly approach to the problem. The first of these, I believe, can be supplied to a degree that, while far from having the completeness that is desirable, would be adequate to enable a reasonable assessment. This missing element is a set of cohesive statements of an anatomy for decision making with respect to each of the rate revision proposals now being considered in the public forum. From the thoughts and recognitions that have emerged piecemeal, we can and must develop a realistic and cohesive approach, to be articulated in the national forum and to be applied to specific proposals in specific circumstances. From the enormous amount of material that has emerged from the Rate Design Study, and from the work that is continuing in that project as well as from the work of others, there now can be drawn together in orderly form the bases on which decision making concerning the proposals should lie. For each such proposal, there is someone who feels very strongly that application of that proposal to rate structures is important; for each such proposal, proponents have offered arguments that are at least plausible, and opponents have made arguments that cannot be dismissed. It is time to face those arguments squarely, and this can be done rationally only if we develop a manageable method of systematically dissecting the substance of the proposals and the means for reaching conclusions. If this has been done to date, I have not seen it. Within the limitations that I shall discuss later, I believe that we have now reached the point where we can bring our efforts to bear successfully on the problem. In reality, the passage of the Public Utility Regulatory Policies Act of 1978 requires that we do so. The alternative is that decision making will occur in a disjointed fashion that will produce sound results only by happenstance.

The second fundamental missing element is, of course, the input that enables application of the anatomy for decision making. That input consists of data but also, at the present time, of methodology not yet perfected. With respect to our ability to supply these inputs, I am much less confident than I am that we can now develop the framework for decision making. My particular interest begins, as the title of this presentation indicates, with the condition of load research into the way customers use electricity.

For about forty years a small group within the electric utility industry — the AEIC Load Research Committee and its predecessor — has conducted research into the way customers use electricity and has encouraged others to make such studies. The resources made available to the committee were usually meager, and the equipment available for their work was primitive compared to their needs. Their activities were more tolerated than welcomed, and the results of their research interested only a few. However, they developed many sound principles for conducting and interpreting load research, and their urgings led to
the development of more sophisticated equipment. Given the limitations under which they functioned, the results which they produced served as a guide to the development of basic principles of load design, and they accumulated a body of experience in the difficulties of load research that is not duplicated elsewhere. The problems which they addressed were those of another day. Some of the answers they obtained continue to be relevant, but, more important, the rigorous standards for data collection, analysis, and interpretation which they imposed on themselves should command the attention of all who are interested in this aspect of rate work.1 To their persistence and high standards, I pay tribute. From my experience of association with this group, I have acquired a great respect for their insights. From the knowledge which I have received from them, and from my short, but I believe sound, direct experience in load research, arise the concerns which I express here about load research as it applies to contemporary proposals for rate structure revision.

What, then, are my load research concerns about the proposed rate revisions? First, I continue to see generalizations drawn from meager data. However, it is encouraging to note that the willingness to draw such generalizations seems to be more restrained than previously, and more attention is being given to the limiting conditions of the data. Examples of this restraint are the recent article in Electrical World by executives of the Department of Energy and the Research Triangle Institute analyses of the Arizona and Connecticut experiments.2 Despite these recognitions of the limitations, the analyses are still often assigning numbers.3

Note that the assigned numbers tend to be so broad and uncertain that they provide little realistic aid in computing effects. They remind me of the recent announcement, attributed to the Secretary of Energy, that the 1985 expected savings in imported oil to be gained by the Public Utility Regulatory Policies Act provisions is from zero to 160,000 barrels per day. Whether the subject be elasticity, oil savings, or what have you, from estimates so broad, what use can be made?4

A study of the conditions surrounding the experiments to date tends strongly to negate the probability of pooling results to obtain some nationwide relevant band of elasticity. Furthermore, it renders transferability of results, even when limited to descriptive statistics, of uncertain worth.

From the statistics and experiments thus far, and from the explorations that took place at the San Diego workshop sponsored by the Electric Power Research Institute, the question arises as to whether we can have, in the reasonably near future, a satisfactory answer to the elasticity problem. It may be that we must be satisfied for now with gaining some basic insights not translatable into numbers worthy of direct application without doing violence to load research principles. I am concerned that efforts still seem to be directed toward achieving specific numbers when perhaps they should be oriented more toward a better understanding of fundamental but nonquantifiable results. It is my hope that the work coming from the second phase of the Electric Utility Rate Design Study will produce valuable insights along this line.

Another fundamental concern of mine is the misuse of results even when they are carefully circumscribed. That misuse lies in jumping to conclusions in which load research results play only one part. This is particularly true with respect to time-of-day rates. The tendency, again somewhat ameliorated in recent months, has been to leap from conclusions that time-of-day rates produce significant changes in load patterns to the conclusion that the public interest is served by the adoption of time-of-day rates.5 Even granting the legitimacy of the first conclusion, which I do not, the great leap to the second conclusion is not warranted. The most obvious reason is that monetary costs have not been reckoned with. At the present time the capability to reckon with those costs is far from fully developed, particularly with respect to cost in the distribution system. Certainly, methods for asaying the “costs” to customers’ life-styles and working conditions, and for appraising the so-called societal effects, fall short of practical use.

These concerns lead us to a larger matter, namely, a total framework for decision making in which the results of economic and engineering research play roles that may not be decisive. It is here that we encounter, in addition to strict cost-benefit analysis, concerns over uncertainty of effects on society and its energy sources. It is here that great uncertainties arise as to what short- and long-range results would follow the adoption of specific load management techniques, both on the supply and demand side, and under both pricing and load control methods. When we think about the vital role that elasticity knowledge plays in judging the feasibility of specific rate revision proposals, and our paucity of solid information on that subject; when we review the current state of the art in determining overall cost-benefit based on just the effects on the utility’s cost and on the customer’s cost; when we consider the limited information we have on customer preferences and responses other than usage pattern responses, then we recognize that we have a multifaceted problem for which, in many of the key areas, we have little to offer except generalizations of dubious validity and few specific numbers which can be applied.

By offering these impressions with respect to the elasticity and total
cost-benefit methodology, am I saying that we should de-emphasize the current move toward increasing load research? The answer is an emphatic no. Elasticity is not the only proper subject for load research; many applications for its results are needed throughout the industry. Even in the area of elasticity we need whatever insights further studies of variance or response surfaces can produce. It is to be hoped that the time might come when quantification will be more feasible and useful. But if not, load research is essential for other aspects, including costing, rate design (other than time-of-day), forecasting, and distribution planning.

My thoughts on this point are so strong that I am disturbed that more resources are not being devoted by the industry to load research, particularly to the analysis of data and to the development and articulation of the role of load research in decision making. For many in the industry, load research is an unknown. For others, it is just getting under way. Even for those with longer standing and more sophisticated systems, the lack of trained personnel devoted to analyzing and interpreting the data, and the lack of formal programs for training such personnel, is severe. Moreover, the industry is not allocating sufficient resources to the development and presentation in the national forum of the role and limitations of load research and the development of sound procedures and usage. A casual review of the standards and special rules of Sections 111, 113, and 115 and the data requirements of Section 133 of PURPA make readily evident the need for greater dedication of industry resources to the study of customer pattern usage.

Can I then assert that we can indeed develop a minimum reference case? While I cannot speak with complete certainty, I think that from the vast materials prepared in the Rate Design Study probably there can be developed some base reference cases for specific types of load management components, both in pricing and in the use of load control devices. However, I wonder if the circumstances of these minimum cases will occur in sufficient number to make their application more than minimal.

I am convinced that we can provide the regulator with much more orderly guidance to decision making than has been done to date. It is my hope that the Rate Design Study will produce a type of road map that will spread before the regulator, in orderly fashion, the steps for decision making and a guide to the use of the available materials. I believe that the atmosphere and the degree of understanding requisite to the task now prevail and that we can suggest paths, filled though they be with uncertainty, that will better aid us in distinguishing the good from the bad.

True definitive progress is possible only when the smoke is cleared, theory is exposed to critical examination, data are developed and their applicability appraised, and a path to ultimate decision delineated with passionate discernment of components of decision. If that path contains forks of great consequence to the correctness of the ultimate decisions, but little proven evidence or even theoretical certainty is available to guide the selection at the forks, the decision maker will undoubtedly experience the dismay and frustration that always accompany a high degree of uncertainty. But he will have traveled an honest path, cognizant of those points for which concrete support exists, and equally aware of those points for which objective support is highly fragile. Such an outlook does not mean that sensible decisions are beyond possibility. It means only that assumptions must be made with extraordinary care, that prudence in subjective judgments must be the order of the day, that the course chosen must minimize exposure to harmful consequences, and that the system must enable expeditious correction.

Without such a framework — such an anatomy for decision — I do not believe the rate manager and the regulator can expect to achieve sensible appraisals; bewilderment will prevail in rate design regulation. With such a framework — filled though it be with lack of data and methodology at points of great consequence, and thus dominated by great uncertainty — I believe that regulators and rate managers nevertheless can function effectively to explore the potential for improvement in the design of electric utility rates. It will not be easy, and the answers will not be obvious, but I think it can be done.

Notes
1. Association of Edison Illuminating Companies, "Standards for Preparing Reports of the Load Research Committee," 4th rev. ed. (November 1971). This document is now in process of revision to reflect changes in the format of the committee's Annual Report and to include a format for reporting on studies specifically providing time-of-use data.
3. Ibid. See also the Research Triangle Institute analyses referred to in the text.
4. Ibid.
Allen Miedema's report probably is the most sophisticated analysis of pricing experiments made public to date. Time-of-day rates "significantly" altered the load profiles of the test customers; however, the reliability of these results should be questioned, Miedema pointed out, because the groups were volunteers. The Arizona, Connecticut, and Ohio time-of-day tests sponsored by the Department of Energy (DOE) showed that residential diversified demands were less at the times of the annual and monthly system peaks; furthermore, these customers reduced on-peak energy uses and increased their usages during the off-peak periods. Thirteen other time-of-day pricing experiments funded by DOE remain to be analyzed.

Samuel Behrends' paper strikes a vital nerve when focusing on the "condition of load research into the way customers use electricity." It is easier said than done to decide upon and implement a load research effort. At the Ninth Annual Conference, I emphasized the importance of time in the rate reform effort. In 1979, faced with the requirements of Section 135 of PURPA that relate to the collection of load data, readers will be interested in some hard facts of a situation with which I am familiar. A utility was directed to study the time-of-day usages of about 200 large residential consumers. "Large" was defined so as to provide a reasonable population size — about 15,000 customers. The process of selecting and purchasing more than 200 magnetic tape load survey recorders went forward. Several company departments cooperated in laying the groundwork for the test. A general procedure was written to assure that all participating groups would perform consistently. A statement assuring customers that the test would have no effect on billing was prepared for use only when customers questioned the recorder-meter installation or when the instrument could not be pole-mounted (about one-third of the locations). The first recorders were received two months after the date of the order. Their installation began two months later, but tape cartridges were not activated for another two months to avoid field servicing and processing work initiated on too small a number of locations. More than eighteen months elapsed before the first data were processed. If all other work could be temporarily sidetracked and expenses ignored, and if all meter manufacturers' production could be channeled your way, and no labor or mechanical problems arose, and no design defects were encountered, this elapsed time probably could be reduced to about a year! My conclusion is that one should include twelve months in the planning cycle of a study in order to cover the time before the start of actual data collection, and this applies to a utility which already has the load research capability. In my judgment, a longer lead time will be needed for the inexperienced organization. A twelve-month test period is necessary to develop a good data bank of the customers' uses of energy, and a similar period of data collection will be needed to monitor the impact of whatever price signal or control may be applied to the test group.

About costs: Two-track load survey recorders with electronic contact devices were involved. These package units cost about $600 each, as received, which amounts to well over $100,000 for the hardware. Installation cost was approximately $100 per location and, eventually, another $30 may cover removal at each location. It was further estimated that annual recorder service maintenance may be $10 for each unit. Continuing expenses prevail during the test period. Meter tapes are changed at the premises once a month, and on a systemwide basis this can average up to one hour per tape and cost as much as $15, which covers labor and transportation, which amounts to $5,000 monthly for a 200-customer test. Costs of translation of data from meter record to hard copy or other acceptable data retrieval system vary widely.

In conclusion, it is easier said than done to decide upon and implement a load research effort.
Robert Borlick develops a benefit-cost framework for evaluating utility pricing concepts and reforming rate design — his "personal views." He discusses the three goals of PURPA: conservation of energy supplied, efficient use of resources, and equitable rates to electric consumers. These are served when a customer values a kilowatt-hour as highly as the resource cost of making it available to him, and when the level of demand and time pattern of that demand are consistent with its value to the user and the resource cost of producing it; equitable rates occur when they reflect the costs imposed on the utility by the customer. The key point about cost-based rates is reflected in Borlick’s suggestion that "decisions on equity judgments are best carried out at the levels of government closest to the people affected by them." Much of the discussion is argumentative, and concerning justifying the implementation of new rate structures, "only explicit consideration of the effects on the customer side of the meter can do that."

Comments

Craig R. Johnson

The purpose of these remarks is to analyze the papers by Samuel Behrends, Jr., Robert L. Borlick, and Allen K. Miedema et al. Each presentation will be discussed in turn, but because all three deal with the common topic of analysis of innovative rate designs, a few general observations will be made in conclusion.

Behrends offers perhaps the most thoughtful discussion of the need, from a decision maker's point of view, for a framework for comprehensively analyzing the effects of various rate designs. His emphasis on the importance of such an "anatomy for decision" is entirely well placed, and it accurately reflects the manager's, and the regulator's, need to make decisions on the basis of information that is often incomplete and concerning outcomes that are often uncertain. Nonetheless, there are methods to manage uncertainty (and to make decisions) that are perhaps more readily available to the decision maker than Behrends suggests, and they are available today.

The importance of approaching the question of rate reform in a systematic and orderly fashion has, as he points out, often been overlooked. Too frequently, proponents of innovative rates, as well as proponents of maintaining the status quo, have based their positions on rhetoric rather than hard systematic analysis. Two elements, Behrends suggests, are critical to a systematic approach: first, a work-
able methodology for evaluating alternative rate proposals and for making decisions based on that evaluation, and second, the data to drive the evaluative process and hence to form the basis of decision making.

It is Behrends’s conclusion that the first of these elements, while yet to be fully developed, is presently adequate to facilitate “a reasonable assessment” of rate proposals. Clearly, this conclusion is, if somewhat cautious, warranted by the progress that has taken place in rate design analysis over the past two or three years. Methods for evaluating the impacts, both direct and indirect, of a variety of rate proposals now exist and have been successfully implemented in several utility decision-making environments.

What is still lacking in methodology, however, is a practical approach to determining the costs and benefits of rate design and other load management options. Although a number of models now commercially available can estimate, if provided with appropriate parameters, the system load and generation planning effects of specific rate options, no straightforward method exists to integrate these factors into an overall cost-benefit assessment. Such an approach, designed to evaluate the net cost-benefits of alternative investments from the perspective of the utility manager and regulator, is required under the Section 111 standards of the Public Utility Regulatory Policy Act (PURPA).

Behrends is much less sanguine about the availability of the other key element of his anatomy for decision making, the necessary data to feed into the evaluative models. The difficulties he points out fall into several categories. First, and perhaps most important, he notes the meagerness of data about customer response to time-of-day and other rates. Second, he doubts that the problem of developing satisfactory elasticity estimates will be solved in the near future. Third, he questions whether use can—or should—be made of any of the rate response measurements developed to date. Finally, he points out, as noted earlier, the lack of refined method for assessing, under conditions of uncertainty, the full costs (and benefits) of rate options.

With respect to Behrends’s first two concerns, it is evident that much remains to be done in developing the response estimates that provide utility and regulatory executives with the data they need for decision making. However, as reflected in his concern that “analyses are still often assigned numbers” and in his urgings that future efforts be directed toward nonquantifiable results, he may be defeating the very objectives he is pursuing. If the ratemakers is ever to have the kind of data necessary to fully evaluate alternative rate designs, he simply must have specific numbers. To provide these parameters, load and research data must be directed toward developing a base of quantified response estimates. Some of this data base now exists, due to the research such as that conducted by Miedema and others, sponsored by the Department of Energy, EPRI, and other organizations.

Behrends’s concern about the uses to which the currently available data are put also appears to be slightly overdrawn. Clearly, a great deal of care must be exercised in interpreting response data, in recognizing its limitations, and in applying it to populations outside the effective sampled population. Nonetheless, a rapidly growing—and generally consistent—body of rate design response data is now available and should be taken into account rather than ignored. Decisions are currently being made, whether required under PURPA or motivated by independent initiative, for which this information would be useful, and although the data may be far from perfect, they will provide at least some practical information.

Behrends’s final point, the need to assess the costs and benefits of various rate strategies under conditions of uncertainty, is very well taken. To date, no such methodology exists. Although several individuals are currently working to develop an approach, utility and regulatory decision makers have yet to be presented with a practical assessment method. In light of the cost-effectiveness test required under PURPA, and in light of the utility sector’s own rate-making initiatives, the need for such a method appears compelling.

If Behrends’s paper can be viewed as calling for the development of an analytical framework for evaluating rate design alternatives, Borlick’s can be seen as an innovative attempt to provide it. As Behrends suggests, a major lack is a method for determining the net costs and benefits of rate and other load management strategies. In Borlick’s paper, just such a method, at least in the conceptual stage, is provided.

Essentially, Borlick calls for an evaluation of the effect of rate changes on utility customers and stockholders, as well as of major local and national externalities. The most important is a cost-benefit analysis of alternative rates, since he suggests that most of the impact of rate reform is likely to be felt by utility customers. Borlick’s method calls for estimating the sum of the consumer’s plus producer’s surplus associated with the demand function for electricity. In simplest terms, this will involve subtracting the loss in customer economic benefits due to higher on-peak rates from the gain accruing to these customers from lower off-peak charges. In addition, changes in the consumer’s surplus associated with the demand for other goods—most important, electricity substitutes—must also be investigated.
Borlick's approach shows a great deal of promise and certainly points the way toward the kind of cost-benefit analysis that must be conducted. However, several factors may limit the practical application of the method by utility rate makers. First, it is unclear whether the demand function identified in Borlick's Exhibit 2 would provide an adequate basis for determining changes in consumer surplus. To do so, the rates charged each customer would have to fully reflect all cost changes incurred by the utility system due to changes in system load patterns resulting from customer response to new rates. Given the nature of the regulatory process, and the uncertainty that long-term cost changes will be reflected in current rates, it is doubtful that the rates selected for the demand model would accurately reflect such utility system costs.

Second, it is difficult to estimate the demand function for electricity, setting aside the question of what rates may be set. To date, no fully practical method of modeling this demand function by time of day has been developed. For a utility to make use of Borlick's cost-benefit framework, it would have to devise such a model and generate the data necessary to drive it. Both are tall orders, a fact Borlick, to his credit, fully recognizes.

Third, how would the results of Borlick's analysis be integrated? Although his identification of the issues which need to be evaluated — including effects on customers and shareholders and other externalities — is entirely appropriate, he does not make clear how he would weight these factors and incorporate them into a single cost-benefit framework. For the framework to be useful to the rate design decision maker, such an integrative mechanism is essential.

Perhaps the most important contribution Borlick makes is the approach he suggests toward uncertainty in utility rate decision making. He is fully cognizant that much of the data required by his analytical framework is lacking, yet he also recognizes that some decisions cannot wait for perfect data or perfect estimates of the net costs and benefits of innovative rate designs. Most important, he suggests that systematic decision making need not require precise estimates of how much the discounted benefits of new rate designs exceed (if they do) the costs of implementing them, but only that the estimates show that the net benefits do exceed the net costs.

Miedema's discussion of the Research Triangle Institute's analysis of time-of-use electricity pricing is an excellent companion piece to the papers by Behrends and Borlick. Miedema summarizes the findings of analyses he conducted for the DOE rate design projects in Arizona, Connecticut, and Ohio. In so doing, he begins to provide precisely the kind of response data that Behrends and Borlick point to as so important. 2

Miedema's findings are most important with respect to two parameters of key interest to utility and regulatory rate executives. First, his results show that time-of-use residential customers almost uniformly reduced their kwh consumption during peak periods, compared to control customers on declining block rates. Second, the diversified kwh demand of time-of-use customers, at the time of system monthly and annual peaks, was lower than that of other residential customers. Other significant findings were that there was no evidence of needle-peaking, either before or after peak ratings periods, and that total electricity usage remained about the same or declined slightly under time-of-use rates.

Miedema's analysis was less successful in defining a time-of-day electricity demand function and thus in providing price elasticity estimates. His regression analysis did not detect any significant elasticities in Arizona for any of the three rating periods under consideration. As Miedema points out, the likely reasons for this finding, despite the consumption differences identified, are the company's complex billing scheme and the small number of empirical observations.

The question remains, however, of just how useful Miedema's findings and similar findings of other researchers can be to the executive faced with actual rate decisions. Certainly, the limited base of empirical knowledge now available cannot answer all of the questions that will occur to the decision maker, especially in the absence of the cost-benefit framework all three authors would agree is so necessary. Nevertheless, it is equally important that decision making can and will proceed — as it almost invariably does — on the basis of less than perfect information. The need to reach timely decisions on rate and, more broadly, load management issues is a fact of life, due both to PURPA's requirements and to the utility industry's interest in improving operational efficiency. The key at this point is to develop a practical cost-benefit framework that will permit appropriate decisions to be made, at the right time, on the basis of the information that then available.

Notes

1. Including this author.
2. It should be pointed out that the author, then a colleague of Borlick's at DOE, provided comments on earlier drafts of his paper.
3. The author, in his previous capacity with DOE, helped to supervise the work conducted by Miedema and has thus advised him on previous versions of his paper.
Collectively, the three papers by Robert Borlick, A. K. Miedema and others, and Samuel Behrends demonstrate some of the reasons the debate over electric utility pricing reform has been so protracted. We lack a clear consensus on appropriate criteria for evaluating alternative rate proposals. We lack estimates of demand and cost estimates exacting enough to sway those critical of the need for rate reform. And we lack enthusiastic industry endorsement for changing traditional rate practices.

Borlick tries to sketch out a “cost-benefit” framework for evaluating alternative rate reforms. Although he has made an honest effort, his attempt falls far short of providing a useful and usable analytical framework. Borlick offers the conventional economic nostrums: pricing based on marginal cost and “benefit” measures based on consumer plus producer surplus. The message that he brings is not new, but the problems are legion.

Borlick’s framework is bound in partial equilibrium analysis. He equates the consumer surplus of high income residential customers with that of low income customers. He assumes the comparability of a residential customer’s consumer surplus (which arises out of a utility function maximization process) with that of an industrial customer (for whom electricity is an input in a production process and the “surplus” arises out of a profit function maximization problem). Borlick presumes that residential consumer surplus could be converted into producer surplus (another name for excess profits) with no net welfare loss. And he presumes a world in which prices reflect economic resource costs rather than imbalanced market power. While the condition Borlick has imposed may be desirable for reasons of theoretical simplicity, the net effect is to cast the analysis in a distorted and unrealistic framework.

But even if the theoretical problem were set aside, the usefulness of Borlick’s model for analysis is doubtful. As Edward Mason has noted: “The theoretical techniques of price analysis have been constructed without regard to their empirical applicability.” Application of Borlick’s cost-benefit model requires explicit estimates of marginal costs and relative demand elasticities over a reasonable domain. Although models of marginal cost determination now abound, there is little agreement among practitioners as to the respective merits of reliability of the alternative approaches — as evidenced by a set of papers appearing elsewhere in this volume. And implicit in Miedema’s paper is a message that we are far from having a clear and unambiguous idea of current demand relationships, let alone the likely future values of demand elasticities.

There is some irony in Borlick’s presentation. His proposal, in keeping with the standard position of traditional economists, is basically passive. Borlick accepts existing preferences and supply conditions and seeks optimization within those constraints. Yet, the Department of Energy, where Borlick is employed, is an interventionist agency seeking to change preferences and supply conditions. DOE’s success may be Borlick’s downfall.

Miedema and his colleagues have embarked on a task which is somewhat more tractable than Borlick’s: the evaluation of the results of three specific time-of-use experiments. Miedema’s paper is carefully written and, in light of the previous EPRI/EEI rate study director’s protestation last year about wispy data, quite timely. The analysis of the experiments appears to demonstrate a time-of-use pricing effect, although the estimates of elasticities (crucial to Borlick’s procedures) appear to be unreasonable. The results investigated by Miedema are only one subset of the time-of-use pricing experiments underwritten by DOE. It appears that it will be many years before the final results can be evaluated.

Behrends believes that the “five years” of thrashing over rate reform issues has brought us to a point at which we can “separate the
good from the rot." All that is now missing, Behrends comments, is an
anatomy for decision making, sound data, and a methodology for
evaluation. We are cautioned to proceed with caution; "prudence in
subjective judgments is the order of the day."

While I would be loath to advocate action without analysis, a dou-
ble standard is clearly evident. In the 1970s many regulatory changes,
of consequence equal to or greater than the proposed pricing re-
forms, have been pushed through by the utility industry: adoption of
automatic adjustment clauses, use of future test years for revenue
requirement determination, placement of construction work in pro-
gress in the rate base, use of interperiod tax normalization, expansion
of investment tax credits, change in depreciation procedures, and so
forth. In all of these cases the industry supported the change. While
several might argue that the changes were appropriate, in none of the
cases was there the protracted effort to study, investigate, experiment,
and otherwise mediate about the consequences of the action. For
pricing reform, however, great value is suddenly found in cautious
analytics. The difference between the proposed pricing reforms and
the regulatory reforms already adopted in this decade is that the
utility industry does not enthusiastically endorse the pricing reforms;
indeed, the response of many utilities has been downright hostile.

At a certain point, the call for "further study" becomes a quest to
delay, sidetrack, or bury those proposals to which a group is opposed.
A primary vehicle for further study of the rate reform proposals has
been and continues to be the EPRI/EEI rate study. To date, that study
has failed to provide summary guidance to the regulatory community.
We can only hope that the conclusions of those investigations will be
timely. Otherwise, a great deal of money and effort will have been
spent to provide welfare for consultants and fodder for historians.

Part Three

Network Access Pricing in
Telecommunications
The Theory of Network Access Pricing

Robert D. Willig

This paper presents a theoretical normative analysis of prices of access to a network. While the immediate motivation for the study is the bevy of presently pending policy issues concerning the telephone network, the analysis will clarify the principles of pricing the services of other networks as well. For example, with few modifications, applications would be appropriate to the postal system and to computer networks.

There are two basic features of networks that create special pricing problems and issues. First, the value to an individual of access to a network depends on the size and composition of the group that also has access. For example, telephone service is more valuable the more of one's friends and associates are subscribers so that they can be conveniently reached by telephone. This effect, termed network externalities, causes private decisions about whether or not to purchase network access to have ramifications for others who are not parties to those decisions. As such, it may be in the public interest for prices of network access to be lower than they would otherwise be, with the difference reflecting the value to others of individuals' purchases of access.

The second basic feature of networks that creates special pricing problems is the technological feasibility of interconnection. Firms
other than the one operating the network could, if they had technical access to the network (that is, if they were interconnected), offer services to consumers with network access. These (called network services) might be perfect or close substitutes for those sold by the network operator. The financial success of such a competitive offering may or may not be in the public interest, and, either way, it may or may not be in the financial interest of the network operator. In any case, the profitability of such competing network services will depend critically on the structure and levels of prices for the technical network access necessary for interconnection.

Thus, network externalities and interconnection are basic features of networks that create special complexities in the analysis of prices for network access — both that of final consumers and that of vendors of network services (henceforth called technical network access). These complexities, their interactions, and the consequent lessons for public interest network access prices are the focus of this paper.

The next section describes the abstract stylized network to which the subsequent analyses will pertain. The third section develops a theory of network externalities that is cast within the framework of the standard theory of consumer welfare. Much care is taken to show that network externalities need not be viewed as metaphysical phenomena and that, instead, they can be measured with observable demand data organized as the incremental consumers' surplus in network services. In the fourth section, characterizations of first best network access prices are derived, and it is shown that it is plausible that they are significantly below marginal production costs due to the network externality effects. Consequently, first best prices can be expected to yield financial losses that may not be feasible. Hence, profit constrained welfare optimal (Ramsey) network prices are characterized. These, too, are affected by network externalities, and they have the additional property of being demand as well as cost differentiated.

The fifth section discusses technical network access prices from the viewpoint of public interest desiderata concerning incentives for competitive entry. It is shown that the Ramsey optimal network prices are consistent with these desiderata only if the technical network access prices are differentiated by final use and if each is set equal to the unit contribution to network profit of the final use. Finally, in the concluding section, relationships between such prices and some (abstract) institutional mechanisms are discussed.

The Stylized Network

For the sake of clarity, the analyses will be conducted in the context of a highly simplified and stylized network that is just complex enough to represent the features at issue. As pictured in Figure 1, the network links those consumers who are equipped with the good or service called network access. Each local switching center is the termination point of lines to all the connected consumers in a local area. Flows along the network from one consumer to another within the same local area traverse the access lines of the two consumers, which are temporarily interconnected for the duration of the flow by means of facilities in the local switching center. Such flows will be termed local network services, or local usage. Other, nonlocal network services, or just network services, are flows between more distant consumers. These are switched at local switching centers from consumers’ access lines to transmission facilities that interconnect the various local switching centers. In Figure 1, the junctions between transmission facilities and local switching centers are symbolized as squares.
In this abstract physical model, it is clear that local usage serves as an input needed for the production of network services. And, as depicted in the figure, it is technologically feasible for a competing vendor of network services to construct its own, junction and transmission facilities, while purchasing technical access and local usage from the network. The network services thereby produced need not be identical to those offered by the network operator, and, even if identical, they need not incur the same production costs. For example, the network operator may enjoy economies of scope from the integration and coordination of network access, local usage, and network services. Such economies of scope would be unavailable to a competing vendor of only network services. On the other hand, the latter may be able to employ a newer and less costly transmission system than that embodied in the sunk plant of the network.

For present purposes, let the cost of efficient operation of the network be represented as

\[ C(N, S^i, S^f) . \]  

(1)

Here, \( N \) is the vector whose \( i \)th component, \( N_i \), is the number of consumers of type \( i \) with access to the network. Consumers in different local areas or with different locations within the same area are of different types as long as their connections to the network incur different costs. Also, consumer types may be distinguished by such demand-related factors as business or residential status. Here, too, \( S^i \) is the vector whose \( j \)th component, \( S^i_j \), is the amount of the \( j \)th local network service produced. Similarly, \( S^f \) is the vector of the quantities produced of nonlocal network services (\( T = \) transmission). It will often be convenient to combine the vectors \( S^i \) and \( S^f \) into the single vector \( S \). Throughout, network services that cause different costs must be segregated into different categories. The categories may be distinguished by such factors as various dimensions of quality, time of day, distance, special handling, and nature of end-point areas. However, it is also possible that categories will be distinguished by purchaser characteristics.

Each of the variable elements of the network's output is captured in a component of the vectors \( N \) and \( S \), and each can, in principle, have a distinct price. The vector \( P \) represents the prices of consumer access to the network, and the vectors \( P^l \) and \( P^u \) represent the prices of all local and nonlocal network services, respectively. It need not be assumed that the prices, \( P \), are borne by the party that initiates the network service. Rather, as a matter of terminology, the prices will be borne by the purchaser, who is assumed to have the power of decision over the purchase. The notation precludes neither the possibility that some prices of distinct categories are identical, nor the possibility that some prices are zero.

The next section develops some new theory of the behavior and welfare of consumers who may purchase network access. This theory is essential for a careful analysis of the public interest pricing of network access and network service.

**Network Externalities and Incremental Consumers' Surplus**

It is certainly intuitively plausible that the value to a consumer of network access be sensitive to the size and composition of the group that also has access. As such, a consumer's decision on whether or not to purchase access would depend on the group with access. More important for present purposes, the welfare of a consumer committed to purchasing access in the relevant situations would be increased if the population of network subscribers were to grow more favorable. The aims of this section are to explain and then formalize these intuitive notions in a manner that enables the size of the effects to be assessed by means of observable market behavior. In particular, it will be argued and then shown that the magnitude of network externalities can be estimated with a recently developed consumers' surplus tool called incremental consumers' surplus.

Roughly speaking, a consumer's surplus for a particular service is the value to the consumer of the total number of units purchased, over and above the money paid for those units. Such surplus would be positive, even at relatively high prices, as long as the consumer would continue to buy some, albeit less, of the service if its price were to rise. Each unit purchased contributes to the surplus an amount equal to the difference between the highest price the consumer would pay for it (the reservation price) and the price the consumer actually must pay. Consequently, consumer's surplus can be measured as the area between the demand curve, which depicts the schedule of reservation prices, and the horizontal drawn at the market price.

Consumer's surplus in a particular service can be validly interpreted as the value to the consumer of the opportunity to purchase all he wants of the service at the going price, on average the comparison scenario in which the service were unavailable, or its price were so high that none of it would be demanded. Rather loosely, then, for heuristic purposes, we could express this relationship as: (consumer's surplus in
the service) = (real income at market prices) – (real income in comparison scenario). The following rearrangement will provide an analytic handle on network externalities: (real income at market prices) = (real income in comparison scenario) + (consumer’s surplus in the service).

From this viewpoint, consider a change in the economic position of the consumer that would have no effect on him in the comparison scenario, but that might have an impact when the service is available at the going price. For such a change, (change in real income at market prices) = (change in consumer’s surplus in the service). This heuristic relationship suggests that the effect, in dollar terms, of a change in the external circumstances of a consumer can be measured as the resulting incremental consumer’s surplus in a purchased service.

The crucial stipulation is that the change would have no impact on the consumer in the comparison scenario in which the service were not purchased at all.

This requirement can be plausibly assumed in a variety of circumstances. For example, a change in the quality of a good or service would certainly have no direct effect on an individual who did not consume the good or service. Also, there are public goods (such as highways) which are “asymptotic complements” to certain private goods (such as automobile and motorcycle services) in that the one would be valueless to a consumer without the other. Consequently, the change in the real income of a consumer caused by a change in product quality and by a change in the level of a public good can be measured by the associated incremental consumer’s surplus in, respectively, the good with changing quality and in the good which is an asymptotic complement to the public good.

In the context of a network, a change in the population of subscribers with access to the network would have no direct effect on a consumer utilizing no network services, because it is only through network services that a consumer has access to other subscribers. This observation would seem to indicate that incremental surplus in network services can be used to measure the effects on others of an individual’s decision as to whether or not to purchase access. Such effects on others are exactly what have been termed network externalities.

Here, however, there are two intrinsic complications that do not necessarily arise in other contexts. First, although only one consumer is the purchaser of a unit of a network service, the service is a flow between two consumers, and it may benefit both. As such, while it is self-evident that changes in the population of subscribers would have no effect on a consumer with no usage of network services, such changes may nonetheless affect a consumer who purchases no network services, but who is a recipient of network flows purchased by others. Thus, incoming flows must be explicitly analyzed. The treatment of this complication is deferred until later in this section; in the interim, it is assumed that incoming flows have no value to the recipient.

The second complication arises from the fact that there are many distinct network services. As such, it can only be asserted that changes in network size would fail to affect an individual consuming none of these diverse services. (This contrasts sharply with changes in the quality of a product which would fail to affect an individual consuming none of that one product.) Consequently, it is the incremental surplus in the totality of network services that must be used to measure network externalities. If the demands for the different network services were independent of one another’s prices, then this would be just the sum of the incremental surpluses in each of the services. Otherwise, however, some care (explained in the next subsection) must be exercised in forming the correct sum.

Figure 2 displays a representative component of the incremental consumer’s surplus used to measure network externalities. Each of the curves is a demand curve of a particular consumer for the $i$th network service, as a function of its price $P_i$. The inner curve gives the consumer’s demands for the service for a given population of network subscribers. If that population were larger, the consumer’s demands at all prices would expand to those along the outer curve. The concomitant increase in the value to the consumer of the availability of the service at a price $P_i$ is given by the shaded area — the incremental consumer’s surplus. It is the sum of areas of this form, over all network services, that measures the benefit to the consumer of the expansion of the network. Of course, the reverse also holds. That is, the loss to the consumer from a shrinkage of the network would be measured by the incremental surplus areas eliminated by the contractions in the demand curves of the network services.

Thus far, the discussion has been limited to the effect on one consumer of changes in the population of network subscribers. Yet, such changes can simultaneously affect all of those who remain connected to the network. Fortunately, a straightforward extension of the analysis can be made to measure the total effect of network externalities.

Demand curves of individuals can be summed, horizontally, to form the market demand curve which gives total market demand as a function of price. The consumers’ surplus area under the market demand curve is the sum of the individual consumer’s surplus areas. The incremental consumers’ surplus associated with the market de-
mand curve is, likewise, the total of those of all consumers. It follows that the sum of the effects on individuals' real incomes of an expansion or a contraction of the network can be measured by inframarginal consumers' aggregate incremental surplus in network services.

This discussion has indicated, in heuristic terms, how network externalities can be measured with consumers' surplus tools. The use of these tools involves intuitive manipulations of data on observable market demand behavior. As such, it must be recognized that network externalities need not be considered metaphysical phenomena; instead, they are quantifiable and amenable to scientific analysis.

Furthermore, the analytic representation of network externalities as incremental consumers' surplus enables qualitative insights into their sources and qualitative assessments of their importance. For example, a change in the population of network subscribers that did not affect demand for network services at market prices, and at all

higher prices, would yield no incremental surplus and would therefore cause no network externalities. In contrast, sensitivity of inframarginal consumers' demand for network services to the size of the network, at market prices, would indicate the existence of network externalities. Yet, the inverse does not hold, because it is plausible that demand for network services would be relatively sensitive to composition of the group with access at high prices and be relatively insensitive at low prices. This would occur if consumers were willing to redirect their relatively low value network flows, while, at higher prices, the remaining high value flows were less flexibly directed to particular recipients. In any case, the more sensitive to network size and composition are inframarginal consumers' demands for network services, the more important are network externalities.

With this heuristic groundwork in place, it is now necessary to give a careful treatment of the theory. Only a formal development can eliminate remaining feelings that network externalities are metaphysical. Only the formal framework will enable more general consideration, in a later section, of incoming flows. And only the use of the accepted theoretical paradigm will enable the determination, in the fourth section, of the implications of network externalities for optimal network access pricing.

**Consumer Theory and Network Externalities**

In this section, the standard theory of the welfare of an individual consumer is extended to allow for network externalities. A principal maintained assumption is that incoming flows have no value to the recipient. The relaxation of this assumption is the subject of the next section.

The consumer is assumed to have preferences that are representable by the utility function

\[
U(x, \bar{x}, N),
\]

(2)

where \(x\) is the vector of purchased network services, \(\bar{x}\) is the vector of all other goods and services consumed, and \(N\) is the vector of the numbers of consumers of different types (as in the previous section) who have access to the network. The utility function is assumed to be regular in that it is continuously differentiable in all arguments, and it is nonde-
creasing and strictly quasi-concave jointly in \( s \) and \( \hat{x} \). The consumer has income \( m \) and faces the vectors of prices \( \hat{\mathbf{p}} \) and \( \hat{\mathbf{r}} \) for the goods and services in the vectors \( s \) and \( \hat{x} \), respectively. The consumer also faces a price of \( E \) for network access, without which the vector \( s \) of purchased network services must be zero.

The critical presumption that will permit network externalities to be linked to observable demand behavior is that utility is insensitive to changes in \( N \) when \( x \) is zero. Thus,

\[
\frac{\partial U(o,\hat{x},N)}{\partial N} = 0, \quad \text{for all } i, \hat{x}, \text{ and } N. \tag{3}
\]

An associated condition that will also be assumed throughout is that network services are inessential in that there are levels of other goods and services great enough to compensate the consumer for the absence of any network services. This condition is needed to impute economic meaning to equation (3) because, otherwise, the consumer would be unsaddled by the absence of network services, and no changes would be of consequence to him in that state.

It is hypothesized that the consumer chooses his purchases in accord with his own preferences and within the limits of his income. Consequently, if he were to purchase network access, he would choose his other purchases, \( s \) and \( \hat{x} \), to maximize utility \( U \). That is,

\[
\text{maximize } U(s,\hat{x},N), \quad \text{subject to } \hat{\mathbf{p}} \cdot s + \hat{\mathbf{r}} \cdot \hat{x} = m - E. \tag{4}
\]

Let \( s(p,\hat{p},m-E,N) \) and \( \hat{x}(p,\hat{p},m-E,N) \) denote the choices of \( s \) and \( \hat{x} \) that maximize utility, subject to the budget constraint, as in equation (4). These are conditional demand functions that give the consumer's optimal purchases of \( s \) and \( \hat{x} \), as functions of prices, disposable income, and the population of network subscribers, all conditional on the consumer purchasing access to the network. Correspondingly,

\[
\ell(p,\hat{p},m-E,N) \tag{5}
\]

is the conditional indirect utility function that gives the optimal level of utility, from equation (4), as a function of the parameters facing the consumer, conditional on the purchase of network access.

The analogue of equation (4) that is conditional on the consumer not purchasing network access is:

\[
\text{maximize } U(o,\hat{x}), \quad \text{subject to } \hat{\mathbf{p}} \cdot \hat{x} = m. \tag{6}
\]

The associated indirect utility function that gives the optimized level of utility in equation (6) is denoted by \( \ell(p,m) \). Note that \( N \) does not appear as an argument of utility here, due to the assumption in equation (5).

This simple theory predicts that the consumer will purchase access if his optimized utility conditional on such purchase exceeds the maximum utility he can obtain without it, that is, if

\[
\ell(p,\hat{p},m,E,N) > \ell(p,m). \tag{7}
\]

Then, the consumer's unconditional indirect utility function which takes endogenous account of the access decision is

\[
\ell^*(p,\hat{p},E,m,N) = \max \left\{ \ell(p,\hat{p},m-E,N), \ell(p,m) \right\}. \tag{8}
\]

Now, it is clear that \( N \) is of no concern to a consumer who is not purchasing network access, so that attention can be limited to the case in which \( \ell^* = \ell(p,\hat{p},m-E,N) \). Here, network externalities can appear through the effect of \( N \) on indirect utility, \( \ell \). However, in this form, the effect would be in terms of units of utility, which are essentially arbitrary.

Thus, it is most useful to transform \( \ell \) into the ordinarily equivalent indirect utility function that is scaled in terms of real income, relative to some base. This is the income compensation function,

\[
\mu(p,\hat{p},n|m,\hat{p},N,m), \tag{9}
\]

defined to be the income necessary for the consumer, when facing the exogenous parameters \( p,\hat{p},N \), to reach the same level of utility attainable with income \( m \) and parameters \( p,\hat{p},N \). Formally, this definition is:

\[
\ell(p,\hat{p},n,m) = \ell(p,\hat{p},m). \tag{10}
\]

Hence, equation (9) is the real income of the consumer in the situation summarized by \( p,\hat{p},N,m \), relative to the base defined by the parameters \( p,\hat{p},N \). Of course, if the parameters are the same in the two situations, then real and nominal income coincide; that is,
\[
\mu(p^*, \bar{p}, N^0 | p^*, \bar{p}, N^0, m) = m. \tag{11}
\]

Otherwise, real and nominal incomes differ by an adjustment for the effects on the consumer of the changes in parameters. In particular,

\[
\mu(p^*, \bar{p}, N^0 | p^*, \bar{p}, N^0, m - E) - \mu(p^*, \bar{p}, N^0 | p^*, \bar{p}, N^0, m) = \mu(p^*, \bar{p}, N^0 | p^*, \bar{p}, N^0, m - E) - (m - E) \tag{12}
\]

is the change in real income caused by the change in the population of network subscribers from \(N^0\) to \(N^0\).

In other words, equation (12) formally represents the network externality effect on the consumer, in dollars of real income. It is this representation that can be linked to consumer surplus. The linkage is effected through the fundamental differential equation of welfare economics,11 which here takes the form

\[
\frac{\partial \mu(p, \bar{p}, N | p^* \bar{p}^*, N^* \bar{N}^*, m^*)}{\partial p_i} = s_i(p, \bar{p}, N, \mu(p, \bar{p}, N | p^* \bar{p}^*, N^* \bar{N}^*, m^*)) \tag{13}
\]

This equation can be interpreted as saying that, to the first order, the increase in \(\mu\) from an increase in the price of the \(i\)th service is just that price increase multiplied by the number of units of the \(i\)th service that is demanded. The relevant level of demand is found along the compensated demand curve, because the income argument of \(s_i\), in equation (13), holds \(\mu\) instead of the consumer's actual income.

To study the effect of \(N\) on real income, equation (13) can be differentiated with respect to \(N_i\), the \(i\)th component of \(N\). Suppressing the arguments of \(\mu\), and substituting \(\frac{\partial \mu}{\partial N_i} \partial N_i \partial p_i\), for \(\frac{\partial \mu}{\partial N_i} \partial N_i\), the result of this operation is:

\[
\frac{\partial \mu}{\partial N_i} \partial N_i \partial p_i = \frac{\partial \mu}{\partial N_i} \partial N_i \partial p_i + \frac{\partial \mu}{\partial N_i} \partial m \partial N_i \tag{14}
\]

These relationships, for all of the network services, comprise a system of differential equations which could, in principle, be solved for \(\frac{\partial \mu}{\partial N_i}\) if a boundary condition were stipulated. The requisite boundary condition is

\[
\lim_{m \to -} \frac{\partial \mu(p, \bar{p}, N | p^* \bar{p}^*, N^* \bar{N}^*, m)}{\partial N_i} = 0, \tag{15}
\]

which is the dual equivalent of equation (3), given that network services are inessential.11 Here, the notation \(\rightarrow\) is shorthand for \(\to\), for all \(i\).

Equation (15) has the intuitive interpretation that \(\mu\) must not be affected by \(N\) when \(p\) is so large that no network services can be afforded.12

The clearest solution to the system given by equations (14) and (15) obtains in the case in which there are no income effects on the demands for network services. Then, \(\delta_{N_i} \partial m = 0\), and equation (14) becomes

\[
\frac{\partial \mu_i}{\partial N_i} \partial N_i \partial p_i \partial \mu_i = \frac{\partial \mu_i}{\partial N_i} \partial N_i \partial p_i \partial \mu_i, \tag{16}
\]

and integration of equation (16) for all services, between \(p\) and \(p^*\), yields

\[
\frac{\partial \mu_i}{\partial N_i} \partial N_i \partial p_i \partial \mu_i = \frac{\partial \mu_i}{\partial N_i} \partial N_i \partial p_i \partial \mu_i = \int_{\Gamma \partial p_i} \sum_{N_j} \frac{\partial \mu_i}{\partial N_i} \partial N_i \partial p_i \partial \mu_i \partial N_i \partial \mu_i, \tag{17}
\]

where \(\Gamma(p, p^*)\) is any smooth path between \(p\) and \(p^*\). Now, taking the limit of equation (17) as \(p^* \to\), invoking the boundary condition in equation (15), and letting \(\Gamma(p, p^*)\) denote the limit of \(\Gamma(p, p^*)\) as \(p^* \to\), results in

\[
\frac{\partial \mu_i}{\partial N_i} \partial N_i \partial p_i \partial \mu_i = \int_{\Gamma \partial p_i} \sum_{N_j} \frac{\partial \mu_i}{\partial N_i} \partial N_i \partial p_i \partial \mu_i \partial N_i \partial \mu_i. \tag{18}
\]

This is the solution to the system comprised of equations (14) and (15), when \(\delta_{N_i} \partial m = 0\). It can be immediately applied to the representation of network externalities that is defined in equation (12) through these relationships:

\[
\mu(p^*, \bar{p}, N | p^* \bar{p}^*, N^* \bar{N}^*, m - E) = \mu(p^*, \bar{p}, N | p^* \bar{p}^*, N^* \bar{N}^*, m - E) - \mu(p^*, \bar{p}, N | p^* \bar{p}^*, N^* \bar{N}^*, m - E) = \int_{\Lambda(N^0 \bar{N}^0 \bar{N}^0)} \sum_{N_j} \frac{\partial \mu(p^*, \bar{p}, N | p^* \bar{p}^*, N^* \bar{N}^*, m - E)}{\partial N_i} \partial N_i \bar{N}^0 \bar{N}^0, \tag{19}
\]

where \(\Lambda(N^0 \bar{N}^0 \bar{N}^0)\) is a smooth path from \(N^0\) to \(N^*\). Substituting the
equalities in equation (18), reversing the order of integration between \( N_i \) and \( q_i \), and utilizing the fundamental theorem of calculus yield

\[
\mu(p^* \bar{p}_i, N^0 | p^* \tilde{p}_i, N', m - E) = (m - E) = \int_{\Gamma_{p^*}} \sum_i \tau_i(q, \bar{p}_i, N', m - E) - \\
\tau_i(q, \tilde{p}_i, N^0, m - E) \, dq_i. \quad (20)
\]

Equation (20) is a precise version of the incremental consumer's surplus representation of network externalities that was introduced in an earlier section. The left-hand side of equation (20) is the change in the consumer's real income, base \( \tau_i(p^*, \bar{p}_i, N^0) \), caused by the change in the population of network subscribers from \( N^0 \) to \( N' \). The right-hand side is the difference between the total consumer's surplus in network services with network population \( N' \) and that with network population \( N^0 \). Each total consumer's surplus is a line integral calculated over any smooth path, \( \Gamma \) (\( p, q \)), that runs from the vector of market prices, \( p \), outward so that each component price, \( p_i \), increases without bound.

Perhaps the most natural such path is the one that unboundedly increases each price, in turn, from its market level. Along this path, when the \( i \)th price is increased, the \((i - 1)\)th, \((i - 2)\)th, and so forth, have already been increased without bound, and the \((i + 1)\)th, \((i + 2)\)th, and so forth, are still at their market levels. Then, the total incremental surplus that appears in equation (20) is the sum of areas like that depicted in Figure 2. However, it must be recognized that the relevant demand curves for the \( i \)th network service are defined in terms of the levels of other prices that obtain on the segment of the path along which the \( i \)th price is increased.

Furthermore, it must be recognized that the relevant demand curves in equation (20) are conditional on the consumer purchasing network access. By hypothesis, these are the demand curves in force at the market prices. Yet, it may well be the case that somewhere along the path \( \Gamma (p, q) \) the consumer would choose to leave the network if he actually faced those prices for network services. Consequently, at such a point, unconditional demands would discontinuously drop to zero and thereby diverge from conditional demands. Nonetheless, the derivation of equation (20) shows that it is the conditional demand curves that yield the incremental consumer's surplus measurement of network externalities.

Finally, it must be recognized that the above derivation of equation (20) rests upon the unrealistic assumption of no income effects on the demands for network services. And, in fact, equation (20) does not precisely hold absent this condition. However, it can be shown that equation (20) holds as a tight approximation in a broad range of realistic circumstances.

The demonstration of this fact proceeds from the explicit solution of equations (11), (13), (14), and (15) in the case that all the network services exhibit the same constant income elasticity of demand, \( \eta \):

\[
\mu(p^* \tilde{p}_i, N^0 | p^* \tilde{p}_i, N', m) = m \left[ 1 + \frac{(1 - \eta \Delta A)}{m} \right]^{1/\eta}, \quad (21)
\]

where

\[
\Delta A = \int_{\Gamma_{p^*}} \sum_i \tau_i(q, \tilde{p}_i, N', m) \, dq_i - \int_{\Gamma_{p^*}} \sum_i \tau_i(q, \tilde{p}_i, N^0, m) \, dq_i. \quad (22)
\]

Here, \( \Delta A \) is the change in total consumer's surplus in network services that is caused by the change in network population from \( N^0 \) to \( N' \), and by the change in prices from \( p^* \) to \( \tilde{p}_i \). Note that equation (20) is the specialization of equation (21) that obtains when \( \eta = 0, p^* = \tilde{p}_i \), and \( m \) is replaced by the disposable income \( m - E \). A Taylor Series expansion of the right-hand side of equation (21), together with numerical analysis of the remainder term, shows that

\[
\mu(p^* \tilde{p}_i, N^0 | p^* \tilde{p}_i, N', m) = m + \Delta A. \quad (23)
\]

provided that \( \eta(\Delta A)^2/m \) is small. This condition will hold as long as the income elasticity of demand is in the usual range and the surplus changes at issue are only a small portion of disposable income.

The specialization of equation (23) when only \( N' \) is varying is this restatement of equation (20) as an approximation:

\[
\mu(p^* \tilde{p}_i, N^0 | p^* \tilde{p}_i, N', m - E) = (m - E) \cdot \Delta A. \quad (24)
\]

Finally, equation (24) will hold with nonconstant income elasticities of demand, for a monotone path \( \Gamma(p^*, q) \), as long as \( \eta(\Delta A)^2/(m - E) \) is small in absolute value for all values of \( \eta \) equal to the income elasticity of any of the network services at any of the relevant points. \( ^{16} \)
Hence, with the exercise of considerable care, it has been shown that the standard theory of consumer welfare can be extended to incorporate network externalities. Within this framework, it has been demonstrated that the real income of a network subscriber can be affected by changes in the population that can be reached via the network. Moreover, the actual change in real income can be closely approximated by the incremental consumer’s surplus in network services that is caused by an expansion or a contraction of the network.

**Network Externalities with Valued Incoming Flows**

The preceding analysis maintained the unrealistic assumption that incoming network flows have no value to the recipient consumer. However, it is intuitively clear that if they do have such value, then the linkage between network externalities and incremental surplus may be considerably complicated. The aims of this section are to expose and elucidate some of these complications, and to determine the requisite modifications of the preceding results.

Let $\bar{s}$ denote the vector whose $i^{th}$ component, $\bar{s}_i$, is the quantity of incoming network flows of type $i$. The consumer’s utility function is now

$$ U(s, \bar{x}, \bar{s}, N) $$

(25)

instead of that in equation (2) which explicitly rules out any effect of $\bar{s}$. Clearly, if $\bar{s}$ were invariant to changes in $s$, $\bar{x}$, $N$, and prices, then the analysis that was based on equation (2) would apply without alteration.

There is another and far more interesting case in which the previous analysis is left unchanged by valued incoming flows. Suppose that the quantity of each type of incoming network flow is a function of only the quantities of network services and other goods purchased by the consumer. That is,

$$ \bar{s}_i = h_i(s, \bar{x}) $$

(26)

This specification does permit $\bar{s}$ to vary with $N$ and with prices, but only through their effects on $s$ and on $\bar{x}$. For example, equation (25) would be satisfied if all incoming flows were responses to purchased outgoing ones, with some fixed response ratio. Or, the response ratio could vary with the quantity of network services, or with other purchases, $\bar{x}$.

In this case, equation (25) can be rewritten as $U(s, \bar{x}, \bar{s}, N) = U(s, \bar{x}, h(s, \bar{x}), N) = V(s, \bar{x}, N)$, where $V$ is a new utility function that depends only on $s$, $\bar{x}$, and $N$, as required by equation (2) and the subsequent analysis. Here, the effects of $s$ and $\bar{x}$ on $V$ arise from two distinct sources: the direct effects from the value of consuming the purchases, and the indirect effects from the value of the induced incoming flows. Because these indirect effects fully exhaust the variability of $\bar{s}$, the analysis can proceed in terms of $s$, $\bar{x}$, and $N$ alone, as in the earlier section.

This conclusion is left unchanged by the following technically minor, but economically meaningful, generalization of equation (26); if

$$ \frac{\partial U(s, \bar{x}, \bar{s}, N)}{\partial \bar{s}_i} = 0, $$

then

$$ \bar{s}_i = h_i(s, \bar{x}) . $$

(27)

In this extension, categories of incoming flows that have no value to the consumer are not required to bear the special relationship to the consumer’s purchases that is stipulated in equation (26).

In summary, network externalities can be measured by the incremental consumer’s surplus in network services if the variability in all valued incoming flows is internalized by the consumer. For, in this case, the demands for network services will fully reflect the value of the thereby induced incoming flows. And the effects of the size of the network on the consumer will be fully reflected in its effects on the consumer’s demands for network services.

This does seem to be a plausible scenario, to the extent that valued incoming flows arise in the context of ongoing relationships between others and the consumer. However, the restrictiveness of equation (27) is most clear in the limiting case of $s=0$. Here, for equation (27) to hold, there must be no valued incoming flows, or the remaining variability in valued incoming flows must be internalized through the purchase of $s$. In particular, valued components of $\bar{s}$ may not vary with the population of network subscribers. But if any valued incoming flows remain when $s=0$, then surely they would be eliminated if their purchasers left the network. Hence, an implication of equation (27) is that there be no valued incoming flows when $s=0$.

To relax this restriction, equation (27) can be generalized as follows: If

$$ \frac{\partial U}{\partial \bar{s}_i} \neq 0, $$

Then...
then
\[ z_i = z_i(s, \tilde{x}, N) \]  
(28)

Then, as before, equation (25) can be rewritten:
\[ U(s, \tilde{x}, \tilde{z}, \bar{N}) = U(s, \tilde{x}, \tilde{z}, \bar{N}, N) = \nu(s, \tilde{x}, N) \]  
(29)

But here, even though \( \partial U / \partial N_i = 0 \) when \( s = 0 \), at such a point \( \partial U / \partial N_i = \) \( \Delta \mu(s, \tilde{x}, N) \), which is not necessarily zero. Thus, equation (3) may fail to hold, and, consequently, the analysis of the preceding section must be modified. In particular, the boundary condition in equation (15) must be replaced with an arbitrary one. As a result, equation (24) becomes:
\[
\mu(p^*, \tilde{p}^*, N^0|p^*, \tilde{p}^*, N^*, m) = m + \Delta \tilde{A} +
\]
\[
\lim_{p^* \to \infty} \mu(p^*, \tilde{p}^*, N^0|p^*, \tilde{p}^*, N^*, m) - \mu(p^*, \tilde{p}^*, N^0|p^*, \tilde{p}^*, N^*, m) = 0.
\]  
(30)

Here, the new term is the change in real income that the consumer would experience due to the change from \( N^0 \) to \( N^* \). If the prices he faced for network services were unboundedly large and if he still had access to the network, it is this term that equation (3), or equation (15), guaranteed to be zero. But now, in view of equation (28), it captures the value of the change in incoming flows that would be caused by the change from \( N^0 \) to \( N^* \) at \( s = 0 \). As such, presuming that incoming flows are positively related to the size of the network, and assuming that the newly induced incoming flows have positive value, the new term is positive if \( N^* > N^0 \), and it is negative if \( N^* < N^0 \).

Hence, in this case, the positive incremental consumer's surplus in network services understates the consumer's gain in real income from an expansion of the network. And, conversely, the negative incremental consumer's surplus in network services understates the consumer's loss in real income from a contraction of the network. These understatements would be smaller, the less sensitive to network size would be the value of incoming flows were no network services purchased by the consumer.

In the most general case, incoming flows would depend directly on the prices of network services, as well as on \( s, \tilde{x}, \) and \( N \). It greatly facilitates the analysis of this case to separate conceptually the prices of network services, \( p^* \), faced by the consumer in question from those, \( \tilde{p}^* \), faced by others for their purchases of what will be the consumer's incoming flows. Then, the consumer's utility function is:
\[ U(s, \tilde{x}, \tilde{z}, \tilde{p}^*, N) = \nu(s, \tilde{x}, \tilde{p}^*, N). \]  
(31)

To derive an expression for the effect of network externalities, \( \tilde{p}^* \) can be held fixed as a parameter, and the development based on equations (28) and (29) can be utilized. Consequently, this extension of equation (30) holds:
\[
\mu(p^*, \tilde{p}^*, N^0|p^*, \tilde{p}^*, N^*, m) = m + \Delta \tilde{A} +
\]
\[
\lim_{p^* \to \infty} \mu(p^*, \tilde{p}^*, N^0|p^*, \tilde{p}^*, N^*, m) - \mu(p^*, \tilde{p}^*, N^0|p^*, \tilde{p}^*, N^*, m) = 0.
\]  
(32)

where
\[
\Delta \tilde{A} = \sum_{i} \int_{N^0} (\sigma_i \cdot \tilde{q}_i \cdot \tilde{p}^*, N^*, m) d \sigma_i.
\]  
(33)

The interpretation of equation (32) is the same as that given for equation (30), with one exception. The incremental surplus, \( \Delta \tilde{A} \), is defined in equation (33) in terms of demand functions for network services that include \( \tilde{p}^* \) among their arguments. In view of equation (31), this is sensible because \( \tilde{p}^* \) affects \( \tilde{z} \), which may well itself affect demands for outgoing flows. However, in the incremental surplus calculation given by equation (33), \( \tilde{p}^* \) remains fixed at the level \( p^* \), while the prices \( p^* \) move along the path \( \Gamma(p^*, \tilde{p}^*) \). Consequently, the demand curves which define the relevant incremental consumer's surplus pertain to situations in which the prices faced by the purchasers of the consumer's incoming flows are numerically different from the prices paid by the consumer for his outgoing flows. Such situations arise in most cases of practical interest, strictly hypothetical and unobserved.

Thus, there is a considerable loss in analytic power concomitant with the extension of equation (29) to the completely general formulation of equation (31). Nonetheless, even equation (31) permits the economically meaningful decomposition of network externality effects
given by equation (32). The first portion is the change in value of the hypothetically purchased network services caused by the change in the population of network subscribers. The second portion is the change in values of incoming flows, if the consumer was purchasing no network services.

The first portion of equation (32), \( \Delta d \), can differ from the observable incremental surplus, \( \Delta d \), because the general formulation allows for uninternalized values of incoming flows at all levels of \( s \). Here, with \( p^s = p \), each consumer is injured by increases in network service prices not only because they make his own purchases more expensive, but also because they cause the loss in the uninternalized values of discouraged incoming flows. Then, such uninternalized values would argue for subsidization of the prices of network services. The counterargument would question the rationale for social or network-wide subsidization of flows from one consumer to another, when that same subsidization, if appropriate, could be effected by the very recipients of the valued incoming flows.

This counterargument seems to be compelling. And, in the vocabulary developed in this section, it suggests the plausibility of assuming that the variability in valued incoming flows is internalized by network subscribers. Consequently, the normative study of network prices below is based on equation (29), the model of consumer welfare that incorporates this assumption, rather than on the more general model in equation (31).

**The Gradient of the Indirect Utility Function**

Necessary ingredients of the normative analysis of network access prices are measures of the effects on consumers' utilities of changes in the prices of network access and network services. This section sketches the derivation from the preceding theory of the gradient of consumers' indirect utility functions.

Consider a consumer, \( h \), who is an inframarginal network subscriber and whose behavior is described by the model inherent in equation (29). His indirect utility function, \( \mathcal{U} (p, \hat{p}, E, N, m^h) \), is the maximum level of his utility that he can achieve by choices of \( s^h \) and \( \hat{s}^h \) that remain within his budget. The Lagrangian of his utility maximization program is:

\[
\mathcal{L} = U^h (s^h, \hat{s}^h, p^h (s^h, \hat{s}^h, N), N) + \lambda (m^h - \hat{p}^h - s^h - \hat{s}^h),
\]

where \( E^h \) is the component of \( E \) that pertains to consumer \( h \). Then, by the envelope theorem, the derivatives of \( \mathcal{U} \) are equal to the derivatives of \( \mathcal{L} \), evaluated at the levels of the decision variables that are optimal for \( \mathcal{L} \).

Hence, the marginal utility of income is given by

\[
\frac{\partial \mathcal{U}}{\partial m^h} = \lambda,
\]

and

\[
\frac{\partial \mathcal{U}}{\partial \hat{p}^h} = -\lambda = -\frac{\partial \mathcal{U}}{\partial m^h},
\]

\[
\frac{\partial \mathcal{U}}{\partial s^h} = -\lambda \hat{s}^h = -\left( \frac{\partial \mathcal{U}}{\partial m^h} \right) \hat{s}^h,
\]

where \( q^h \) is the consumer's demand for the \( h \)th network service. Moreover, it can be shown that in a broad range of cases,\(^{10} \) the marginal network externality effect is given by:

\[
\frac{\partial \mathcal{U}}{\partial N_j} = \sum_{T \in \Upsilon, j} \frac{\partial \mathcal{L}}{\partial N_j} (\partial \mathcal{U}/\partial m^h) d_q^i + \lim_{p^h \to \infty} \left[ \frac{\partial \mathcal{L}}{\partial N_j} (p^h \hat{p}, E, N, m^h) \right] \frac{\partial \mathcal{U}}{\partial m^h}.
\]

The divisions by the marginal utility of income transform the terms into units of real income. Thus, equation (38) is exactly the marginal version of equation (30), which represents the network externality effect of a noninfinitesimal change in \( N \). Hence, the discussions of equation (30) pertain to equation (38) as well.

The above results apply to a consumer who is a network subscriber. A consumer, \( k \), who chooses not to join the network is described by the indirect utility function \( \mathcal{U} (p, m^h) \). This depends on only \( p \) and \( m^h \) because the parameters of the network are of no concern to a consumer without network access.

**Public Interest Network Access Pricing**

This section is devoted to the study of prices for network access that are in the public interest. Unavoidably, because of the interrelationships of the network system, the study must deal with the prices of network services as well. The normative stance underlying the analysis is a social decision maker's social preference ordering that respects
individuals’ preferences and that reflects the Pareto principle. As such, the public interest objective function takes the form of a Bergson-Samuelson social welfare function whose arguments are the levels of each individual’s utility: \( W[c^v_1, c^v_2, \ldots] \).

To keep the focus sharply on the effects of the intrinsic properties of a network, various other complications that might be important in particular applications are assumed away. Thus, below, assumptions are made on the social welfare function that preclude concerns about the distribution of income. Furthermore, possible interactions are ignored between the network prices and prices and quantities of other sectors of the economy. Analysis of potential and actual entry of competing vendors of network services is deferred until the next major section.

The following subsection studies the “first best” network prices that are optimal for social welfare. The next section characterizes the Ramsey network prices that are optimal under a constraint on the size of the profit or loss from the operation of the network.

**First Best Network Prices**

The first best network prices are chosen to maximize

\[
W[c^h(p, \tilde{p}, E, N, m^h + \alpha^h \Pi), c^h(p, m^h + \alpha^h \Pi)] .
\]  

(39)

Here, the arguments of \( W \) are the utility levels of all consumers. The expression in equation (39) is an abbreviation in that it includes the utility of only some representative network subscriber, \( h \), and the utility of only some representative individual, \( k \), who is not a subscriber. The constants \( \alpha^h \) and \( \alpha^k \) denote the shares of the network’s profit, \( \Pi \), that accrue to consumers \( h \) and \( k \), respectively. These shares sum to one over all consumers. Network profit is defined to be the revenues from the sale of network access and network services, minus the cost of network operation defined in equation (1):

\[
\Pi = p \cdot S + E \cdot N - C(N, S) .
\]  

(40)

Here, \( S \) is the vector of total quantities purchased and produced of network services, so that

\[
S = \sum_k x^k .
\]  

(41)

Necessary conditions for equation (39) to be maximized by choices of \( p \) and \( E \) are that its partial derivatives with respect to every component of \( p \) and \( E \) be equated to zero. These necessary conditions will be analyzed by utilizing the results derived earlier to provide economically meaningful characterizations of the first best prices.

**Mechanical differentiation of equation (39) yields**

\[
\begin{align*}
\frac{\partial W}{\partial p_i} &= \frac{\partial W}{\partial \tilde{p}_i} \frac{\partial \tilde{p}_i}{\partial p_i} + \sum_j \frac{\partial W}{\partial N_j} \frac{\partial N_j}{\partial p_i} + \alpha^h \frac{\partial W}{\partial m^h} \frac{\partial m^h}{\partial p_i} + \\
&\quad \sum_j \frac{\partial W}{\partial \tilde{m}^h} \frac{\partial \tilde{m}^h}{\partial p_i} \frac{\partial \Pi}{\partial p_i} .
\end{align*}
\]  

(42)

\[
\begin{align*}
\frac{\partial W}{\partial E_i} &= \frac{\partial W}{\partial \tilde{E}_i} \frac{\partial \tilde{E}_i}{\partial E_i} + \sum_j \frac{\partial W}{\partial N_j} \frac{\partial N_j}{\partial E_i} + \alpha^h \frac{\partial W}{\partial m^h} \frac{\partial m^h}{\partial E_i} + \\
&\quad \sum_j \frac{\partial W}{\partial \tilde{m}^h} \frac{\partial \tilde{m}^h}{\partial E_i} \frac{\partial \Pi}{\partial E_i} .
\end{align*}
\]  

(43)

The term \( (\partial W/\partial \tilde{p}_i) (\partial \tilde{p}_i/\partial m^h) \) that appears in equations (42) and (43) is the rate of change of social welfare with respect to the income of consumer \( h \). If these rates of change were different for the incomes of two different consumers, then the social decision maker would find it preferable to effect an income transfer between them. Consequently, the assumption that the distribution of income is of no concern to the social decision maker implies that the rate of change of social welfare with respect to the income of any consumer is the same. That is, for all \( h \) and \( k \),

\[
\frac{\partial W}{\partial \tilde{p}_h} \frac{\partial \tilde{p}_h}{\partial m^h} = \frac{\partial W}{\partial \tilde{p}_k} \frac{\partial \tilde{p}_k}{\partial m^h} = w .
\]  

(44)

Making use of equation (44) and the fact that the shares of \( \Pi \) sum to one, equations (42) and (43) become

\[
\begin{align*}
\frac{1}{w} \frac{\partial W}{\partial p_i} &= \sum_n \left( \frac{1}{\tilde{n}} \frac{\partial \tilde{c}_n^h}{\partial \tilde{m}^h} \left( \frac{\partial \tilde{m}^h}{\partial \tilde{p}_i} \right) + \sum_j \frac{\partial \tilde{c}_n^h}{\partial N_j} \frac{\partial N_j}{\partial \tilde{p}_i} \right) + \frac{\partial \Pi}{\partial \tilde{p}_i} ;
\end{align*}
\]  

(45)

\[
\begin{align*}
\frac{1}{w} \frac{\partial W}{\partial E_i} &= \sum_n \left( \frac{1}{\tilde{n}} \frac{\partial \tilde{c}_n^h}{\partial \tilde{m}^h} \left( \frac{\partial \tilde{m}^h}{\partial E_i} \right) + \sum_j \frac{\partial \tilde{c}_n^h}{\partial N_j} \frac{\partial N_j}{\partial E_i} \right) + \frac{\partial \Pi}{\partial E_i} .
\end{align*}
\]  

(46)
Recall from equation (56) that \( \frac{\partial c^k}{\partial e_i} = - \frac{\partial c^k}{\partial m^k} \) if price \( E_i \) pertains to consumer \( k \), and that it is zero otherwise. Consequently,

\[
\sum_i \frac{\partial W}{\partial e_i} = N_i,
\]

where \( N_i \) is the number of consumers who purchase network access at the price \( E_i \). Substituting this relationship, equation (37), and equation (41) into equations (45) and (46), and rearranging, yields

\[
\frac{1}{w} \frac{\partial W}{\partial p_t} = \sum_i \frac{\partial N_i}{\partial p_t} \left( p_t - \frac{\partial C}{\partial S_j} \right) + \sum_i \frac{\partial N_i}{\partial p_t} \left( E_i - \frac{\partial C}{\partial N_j} \right),
\]

and

\[
\frac{1}{w} \frac{\partial W}{\partial e_i} = \sum_i \frac{\partial N_i}{\partial e_i} \left( p_t - \frac{\partial C}{\partial S_j} \right) + \sum_i \frac{\partial N_i}{\partial e_i} \left( E_i - \frac{\partial C}{\partial N_j} \right).
\]

Mechanical differentiation of equation (40) gives

\[
\frac{\partial W}{\partial S_j} = N_i + \sum_i \frac{\partial N_i}{\partial S_j} \left( p_t - \frac{\partial C}{\partial S_j} \right) + \sum_i \frac{\partial N_i}{\partial S_j} \left( E_i - \frac{\partial C}{\partial N_j} \right),
\]

and

\[
\frac{\partial W}{\partial e_i} = N_i + \sum_i \frac{\partial N_i}{\partial e_i} \left( p_t - \frac{\partial C}{\partial S_j} \right) + \sum_i \frac{\partial N_i}{\partial e_i} \left( E_i - \frac{\partial C}{\partial N_j} \right).
\]

Substitution of equations (49) and (50) into equations (47) and (48), together with rearrangement, yields these useful expressions:

\[
\frac{1}{w} \frac{\partial W}{\partial p_t} = \sum_i \frac{\partial N_i}{\partial p_t} \left( p_t - \frac{\partial C}{\partial S_j} \right) + \sum_i \frac{\partial N_i}{\partial p_t} \left( E_i - \frac{\partial C}{\partial N_j} \right),
\]

and

\[
\frac{1}{w} \frac{\partial W}{\partial e_i} = \sum_i \frac{\partial N_i}{\partial e_i} \left( p_t - \frac{\partial C}{\partial S_j} \right) + \sum_i \frac{\partial N_i}{\partial e_i} \left( E_i - \frac{\partial C}{\partial N_j} \right).
\]

The necessary conditions for \( p \) and \( E \) to maximize \( W \) are that the expressions in equations (51) and (52) equal zero, for all components. The solution to this system of equations is now easy to state:22

\[
p_t = \frac{\partial C}{\partial S_j},
\]

and

\[
E_i = \frac{\partial C}{\partial N_j} - \sum_k \frac{\partial c^k}{\partial N_j} \left( p_t - \frac{\partial c^k}{\partial m^k} \right).
\]

These are the conditions that characterize the first best network prices. Not unexpectedly, optimal prices for network services are equal to the associated marginal costs of production.

In contrast, each optimal network access price is below the associated marginal production cost by an amount equal to

\[
\sum_k \frac{\partial c^k}{\partial N_j} \left( p_t - \frac{\partial c^k}{\partial m^k} \right).
\]

This term is the marginal network externality effect on the total real income of all individuals. Thus, it can be said that first best network access prices are equal to the associated marginal social costs, and that these are the marginal production costs less the marginal network externalities.

The expression in equation (55) makes it clear that the marginal network externality effects that are relevant for pricing are potentially spread over all consumers with network access. In contrast, the effects of potential uninternalized values of incoming flows discussed earlier were concentrated on one consumer. For this reason, it was argued that such values were indeed likely to be privately internalized. And, conversely, it is because network externalities are likely to be diffuse that network access prices appropriately correct for them.

The marginal network externality term in equation (55) can be related to observables by summing the equality in equation (58) over all network subscribers, \( k \). Thus, by ignoring the limit term in equation (58), it follows that equation (55) is not less than

\[
\int \sum_{l \in \Gamma(w,n)} \frac{\partial S_l(q, p_t, E, N, M)}{\partial N_j} \, dq_l.
\]
Here, $S_i(q_i, p_i, E, N, M)$ is the aggregate demand function for the $i^{th}$ network service, and $M$ represents the entire distribution of disposable income. So equation (50) is the rate of change with respect to $N_i$ of the aggregate consumers' surplus in all network services. And this observable quantity provides a lower bound on the amount by which the marginal production cost exceeds the first best network access price.\(^{22}\)

A heuristic argument based on the model in equation (27) can help to build intuition on the quantitative importance of equation (56). Consider the discrete analogue of the latter, which is the change in aggregate inframarginal consumers’ surplus in network services caused by the departure of one marginal consumer from the network. This would be the aggregate surplus in purchased flows to the marginal consumer, if demand for flows to others were unaffected by his departure. Let $\beta$ be the ratio, at all prices, between the quantity of network flows to the marginal consumer and the quantity of network services demanded by him. Equation (56) is then $\beta$ times the consumer’s surplus in network services of the marginal consumer. And, because he is marginal, this surplus is exactly $E_i$. \(^{31}\) Consequently, in this case, equation (56) is equal to $\beta E_i$, and equation (54) yields this expression for the first best network access price:

$$E_i = \frac{1}{1 + \beta} \frac{\partial C_i}{\partial N_i}. \quad (57)$$

If there were no flows to the marginal consumer at any prices equal to or above market levels, then $\beta = 0$, no one would care about departure from the network, and the optimal level of $E_i$ would be marginal production cost. If, instead, every flow demanded by the marginal consumer were matched by an incoming flow purchased by others, then $\beta$ would be one, and the optimal level of $E_i$ would be one-half the marginal production cost. It is conceivable that $\beta$ would be greater than one, with the consequent first best $E_i$ less than one-half of marginal production cost.

Hence, the previously derived relationships between network externalities and observable demand data make it clear that first best network access prices can plausibly be significantly below marginal production costs. The combination of these relationships with the results derived in this section provides a framework for applying demand and cost information to the determination of first best network prices.

**Ramsey Optimal Network Prices**

With the first best prices of network access below marginal production cost, the question inevitably arises as to their financial impact. The answer depends, in part, on the returns to scale properties of the network technology. If it were to exhibit constant returns to the scale of both network services and the population of subscribers, then total cost would be exactly covered by the revenues obtained from prices equal to marginal production costs. In this case, the first best prices would cause negative profit. The losses would be still greater in the more plausible case of increasing returns to scale because, here, even prices equal to marginal production costs leave a deficit.\(^{33}\)

If the network is to be self-supporting, or if the available funds for subsidization are limited, the first best network prices may not be financially feasible. And, in view of equations (53) and (54), such a financial problem is more likely, the more significant are the network externalities that drive the optimal network access prices below marginal production costs. Consequently, for most applications, the relevant normative viewpoint imposes the constraint on prices that they yield a profit of no less than some predetermined figure, $\Pi^*$. Then, what are called Ramsey optimal network prices, or just Ramsey network prices,\(^{34}\) maximize the social welfare function in equation (39) subject to the profit constraint.

The Kuhn–Tucker necessary conditions for $p$ and $E$ to be optimal for this program are:

$$\frac{\partial W}{\partial p_i} + \gamma \frac{\partial \Pi}{\partial p_i} = 0, \quad \frac{\partial W}{\partial E_i} + \gamma \frac{\partial \Pi}{\partial E_i} = 0; \quad (58)$$

$$\gamma (\Pi - \Pi^*) = 0, \quad \gamma \geq 0, \quad \text{and} \quad \Pi \geq \Pi^*, \quad (59)$$

where $i$ ranges over all components of $p$ and $E$, and where $\gamma$ is the Lagrange multiplier on the constraint. If the first best prices satisfy the profit constraint, then they are also Ramsey optimal, and $\gamma = 0$ at the optimum. If, however, due to network externalities or to economies of scale, the first best prices yield a level of profit below the feasible floor, $\Pi^*$, then the profit constraint is binding, $\gamma$ is positive, and the Ramsey prices deviate from social marginal costs.

The Ramsey optimal network prices are characterized by these combinations and rearrangements of equations (58), (51), (52), (49), and (50):
\[
\sum \frac{\partial S_i}{\partial p_i} \left( p_i - \frac{\partial C}{\partial S_i} \right) + \\
\sum \frac{\partial N_j}{\partial p_i} \left( E_i - \frac{\partial C}{\partial N_j} + \left( \frac{1}{1 + \gamma \omega} \right) \sum_{i} \frac{\partial \rho^a_i}{\partial N_j} \frac{\partial \rho^a_i}{\partial n^a} \right) \\
= - \left( \frac{1}{1 + \omega / \gamma} \right) S_i ,
\]
(60)

\[
\sum \frac{\partial S_i}{\partial E_t} \left( p_i - \frac{\partial C}{\partial S_t} \right) + \\
\sum \frac{\partial N_j}{\partial E_t} \left( E_t - \frac{\partial C}{\partial N_j} + \left( \frac{1}{1 + \gamma \omega} \right) \sum_{i} \frac{\partial \rho^a_i}{\partial N_j} \frac{\partial \rho^a_i}{\partial n^a} \right) \\
= - \left( \frac{1}{1 + \omega / \gamma} \right) N_t .
\]
(61)

Here, the terms \( \frac{\partial S_i}{\partial p_i}, \frac{\partial N_j}{\partial p_i}, \frac{\partial S_i}{\partial E_t} \), and \( \frac{\partial N_j}{\partial p_i} \) are price derivatives of market demand functions for network services and for network access.

The qualitative lessons of equations (60) and (61) are clearest when all cross-elasticities of demand are small enough to be ignored. Then equations (60) and (61) can be rearranged into this inverse elasticity rule, suitably modified to account for network externalities:

\[
\left[ \frac{p_i - \frac{\partial C}{\partial S_i}}{p_j} \right] \left( \frac{\partial S_i}{\partial p_j} \frac{S_j}{p_j} \right) = \frac{1}{1 + \omega / \gamma} = \\

\left[ E_t - \frac{\partial C}{\partial N_t} + \left( \frac{1}{1 + \gamma \omega} \right) \sum_{i} \frac{\partial \rho^a_i}{\partial N_t} \frac{\partial \rho^a_i}{\partial n^a} \right] \left( \frac{\partial N_j}{\partial E_t} \frac{E_t}{N_j} \right) ,
\]
(62)

for all \( i \) and \( j \). This rule can be interpreted to say that across all the network services and all the categories of network access, the proportional deviations between the Ramsey prices and the associated marginal social costs are inversely proportional to the associated own price elasticities of demand. Because all such demand elasticities are positive (as defined in equation [62]), all the Ramsey optimal network prices are above the associated marginal social costs when the profit constraint is binding. And the less sensitive to its own price is the demand for a network service or the demand for a category of access, the larger is the ratio between the Ramsey optimal price and the marginal social cost.

The marginal social cost of a network service is just the marginal cost of production. But the marginal social cost of a category of network access is the marginal production cost, less the marginal network externality term of equation (55) adjusted for the profit constraint. This adjustment is made by multiplying by the factor \((1/(1 + \gamma \omega))\). It is equal to one for a nonbinding profit constraint, and it shrinks slowly from that level as prices are constrained away from their first best values.22 Hence, due to network externalities, the marginal social cost of network access can be substantially less than production marginal cost. Thus, although the Ramsey network access prices are above the corresponding marginal social costs, they may nonetheless fall below the marginal production costs of access.

The principal lessons are the same in the general case with cross-elasticities of demand. In fact, equations (60) and (61) are the standard general optimality conditions for Ramsey prices,23 with one important exception. For the determination of optimal network access prices, the role usually played by marginal production costs is here played instead by the marginal social costs, which are smaller by the adjusted marginal network externalities.

It is important to reemphasize that the Ramsey optimal prices for network services generally deviate from marginal production costs to differing degrees that depend upon the characteristics of demands. While this is not the case for first best prices, such prices would cause negative profits, given constant or increasing returns to scale, due to network externalities. Thus, it can be concluded that the special properties of networks generally force optimal prices of network services to be demand as well as cost differentiated. The implications of this fact for technical network access prices are discussed in the next section, after characterizations of desirable structures of such prices are developed.

Technical Network Access Prices

As discussed in the first two sections, a typical feature of networks is the technological feasibility of interconnection with competing vendors of network services. To produce its services, such a firm must purchase the use of access lines, and/or associated switching facilities, that interconnect its transmission equipment with the local switching centers of the network. And it must purchase local network services as variable
inputs for its production of nonlocal network services. The subject of this section is the public interest pricing of such purchases of technical network access.

It is assumed that the network services potentially or actually offered by competing vendors are identical to those offered by the network. A potential entrant's cost of producing the categories of network services in the set \( A \), at the same levels given by the corresponding component of \( S^e \), is denoted as \( C^e(S^e) \). Here, \( e \) stands for entrant. This cost is additional to the costs of purchasing the requisite technical inputs from the network.

The cost to the network of providing the access facilities needed by an entrant is denoted as \( c_e \), and the corresponding price faced by an entrant is \( q_e \). The vector of local network services that at entrant requires to produce \( S_e^e \) is denoted by \( S^e_e \). Then, for example, \( C(N, S_e^e + S^e f S^e f) + c_e \) is the total cost incurred by the network to provide access for the subscriber population \( N \) and for an entrant, to produce local network services of \( S^e e \) for consumers and \( S^e f \) for the entrant which itself produces \( S^e e \) and to produce the vector of nonlocal network services \( S^e f \). Finally, \( q_e \) is the price charged by the network to an entrant for the local network services that the entrant uses to itself produce a unit of the \( i \)th nonlocal network service. The vector of these prices is denoted by \( a^e \), and the vector of its components that pertain to the subset of services \( A \) is denoted by \( a^e_A \).

This notation allows two dimensions of price discrimination to be represented. First, the prices of local network services to consumers, \( p^e \), may be different than the prices of those same services to firms that utilize them as inputs for the production of nonlocal network services. Second, the price to such a firm of a unit of local usage may depend upon the use to which it is put — the particular nonlocal service that it is utilized to produce.

A study by John Panzar (1979) in an analogous framework shows that the first of these dimensions of price discrimination can be essential for the sustainability of Ramsey optimal prices. Furthermore, it may enable prices to be set that all consumers would prefer to any prices that could be achieved absent such discrimination. Moreover, it can make the difference between a financially viable network and one that would be unable to cover its costs. For these reasons, and because such price discrimination between consumers and competing vendors of network services appears to be practically feasible, the analysis will proceed under the presumption that this kind of price discrimination may be practiced.

In general, price discrimination is feasible only if the seller can, without excessive information collection, distinguish the customers who face the higher prices from those who face the lower ones. The second precondition is that the purchasers who face the lower prices be unable to resell the good or service to the higher price group, and thereby arbitrage away the price differentials. These two necessary conditions seem to hold in the case of sale of local network services at different prices to final consumers and to producers of other network services. However, the satisfaction of the conditions seems far more problematical in the case of the aforementioned second dimension of discrimination based on final use.

On the one hand, the network can, no doubt, inexpensively monitor the time, date, and one of the endpoints of the services into whose production the purchased local network services are an input. On the other hand, monitoring other characteristics of these nonlocal services, such as destination, may be impossible without direct and detailed examination of the operations or records of the customer firms. Thus, while it is clear that the prices of local network services can precisely discriminate among some characteristics of their final uses, it may be more difficult to differentiate prices according to others. In view of this potential difficulty, the analysis will assume that precise discrimination is feasible and provide a framework for assessing the importance of such precision. Mechanisms are suggested to increase the precision of price differentiation in cases when monitoring is difficult.

The analysis of technical access prices rests on the fundamental desiderata that they yield profit incentives for the entry of firms that would lower total industry costs and that they discourage socially undesirable entry. These desiderata can be applied to the potential entry of firms that would supply all the nonlocal services offered by the network, or just some subset of those services.

First, a firm that faced technical network access prices \( a^e \) and \( a_n \), and that incurred costs of producing \( S^e \) equal to \( C^e(S^e) \), would experience a profit incentive to market the full quantities of nonlocal services, \( S^e \), if and only if

\[
p^e S^e - C^e(S^e) - a^e S^e - a_n > 0. \tag{63}
\]

This inequality is the condition that the market prices, \( p^e \), provide revenues in excess of the entrant's total costs — production costs plus the technical network access fees that it would be required to pay. The desidera are that this profit incentive occur if and only if total social cost would be lowered by such entry:
Now, consider a partition of the nonlocal network services into disjoint subsets of services, $A_i$, which have production processes that are incrementally separable from one another in the following sense:

$$
C\left( N, S^S, S^T \right) - C\left( N, S^S + S^T e, 0 \right) = \sum_i \left[ C\left( N, S^S, S^T \right) - C\left( N, S^S + S^T e, S^T - S^T_i \right) \right].
$$

The left-hand side of this expression is the incremental cost to the network of providing $S^T$, exclusive of the local services used as inputs into the production of $S^T$. This exclusion follows from the fact that both $C\left( N, S^S, S^T \right)$ and $C\left( N, S^S + S^T e, 0 \right)$ incorporate the costs of these input local services; the former does so implicitly because the costs are incurred on behalf of the network's own production of $S^T$, while the latter incorporates them explicitly because here they are incurred for services sold to an entrant producing $S^T$. The right side of equation (69) is the sum of the incremental costs to the network of providing each subset of network services, $S^T_i$, with each increment again exclusive of local services used as inputs. Hence, equation (69) stipulates that the incremental costs of the subsets of services sum to the incremental cost of the entire set.

This condition would not obtain if two or more of the subsets of services shared some common production facilities. In this case, none of the incremental costs of the subsets taken one at a time would include the cost of the common facilities. But, since the incremental cost of the entire set would include the common cost, this increment would exceed the sum of the incremental costs of the subsets, in violation of equation (69). It is critical to note, however, that the increments of equation (69) hold constant the population of network subscribers, the local services that they purchase, and the local services required as inputs for the production of $S^T$. Consequently, the common facilities that would cause equation (69) to be violated must be additional to those utilized to produce local services and consumers' access.

Hence, there are bound to be subsets of services with common transmission facilities that would cause equation (69) to be an inequality rather than an equality. The implications of this for technical access prices will be studied below. However, it is also plausible that there be some partition of subsets, perhaps on the basis of geography, which are incrementally separable in production in the sense of equation (69).

If there is at least one such collection of subsets, then application of the desiderata to them implies that $a_i = e_i$. This can be seen by summing...
equation (68) over the subsets in the partition, subtracting equation (65) from the sum, and invoking equation (69). Thus, the desiderata require that the network charge each potential entrant a fee for its hook up to the network that is equal to the cost thereby incurred.

When the requirement that \( a_y = a^p_y \) is combined with equation (65), and with equation (68) for each subset in an incrementally cost separable partition, the following implications of the desiderata emerge:

\[
\begin{align*}
\alpha S^T &= \left[ C \left( N, S^x + S^x T, 0 \right) - C \left( N, S^x, 0 \right) \right] + \\
&\left[ p S^T - \left[ C \left( N, S^x, S^T \right) - C \left( N, S^x, 0 \right) \right] \right] \quad \text{(70)}
\end{align*}
\]

\[
\begin{align*}
\alpha S^T &= \left[ C \left( N, S^y + S^y T, 0 \right) - C \left( N, S^y, 0 \right) \right] + \\
&\left[ q S^T - \left[ C \left( N, S^y, S^T \right) - C \left( N, S^y, 0 \right) \right] \right] \quad \text{(71)}
\end{align*}
\]

In both equations (70) and (71), the charges levied on an entrant for its input local services are equated to the sum of two portions. The first portion is the incremental cost to the network of providing the input local services. The second is the difference between the revenues the network earned on the network services lost to the entrant and the incremental costs those lost services caused. Equation (70) shows that the desiderata require this equality to hold over the full set of nonlocal network services, and equation (71) shows that it must hold for each subset of an incrementally cost separable partition.

The implications of equation (71) are clearest when the subset \( A_i \) contains only network service. Then, rearrangement gives

\[
a^2 = C^a_i + \left( p^2 - C^i \right) \quad \text{(72)}
\]

where \( C^2_i \) is the average incremental cost of the input local services, and \( C^i \) is the average incremental cost to the network of the \( i \)th nonlocal service. Thus, the technical usage price \( a^2 \) exceeds the associated average incremental cost by the unit contribution to network profit from the displaced network service.

The representation given by equation (72) exposes the potential conflicts that can occur, absent price differentiation, between Ramsey pricing of network services and technical access prices that satisfy the desiderata. Suppose that \( a^T \) must equal \( a^f \) due to exorbitant costs of effecting the monitoring needed to distinguish their final use from the \( i^p \). Suppose, too, that the average incremental costs of the input local services are the same for the two final uses, so that \( C^a_i = C^a_f = C^a \). Then, equation (72) shows that the desiderata would be satisfied only if the unit contributions from network services \( i \) and \( j \) were equal. Of course, this will generally not be true for the Ramsey optimal levels of the network service prices.

For simplicity, consider the inverse elasticity rule version of the Ramsey characterization. Then, for each nonlocal network service, the Ramsey prices satisfy this condition:

\[
\frac{\left( p^2 - \frac{\partial C_i}{\partial S^T} \right)}{p^2} \quad \alpha^i = \alpha ,
\]

where \( \alpha^i \) is the price elasticity of demand for the service, and \( \alpha \) is a constant. The corresponding unit contributions are

\[
p^2 - C^i + C^i \left[ 1 - \delta \left( 1 - \frac{\alpha}{\alpha^i} \right) \right] \quad \text{ (73)}
\]

where \( \delta \), the degree of scale economies in service \( i \), is equal to the ratio of the average incremental cost to the marginal cost of that service. This same relationship can also be expressed as

\[
p^2 - C^i = \frac{\partial C_i}{\partial S^T} \left( 1 \frac{1 - \alpha^i}{1 - \alpha} - \delta \right) \quad \text{(74)}
\]

Thus, the Ramsey optimal levels of the unit contributions are smaller for services with smaller marginal costs, with larger price elasticities of demand, and with larger degrees of scale economies. And such welfare efficient variations in unit contributions are inconsistent with the desiderata when technical access prices cannot be precisely differentiated. One way to resolve such conflicts is to give up on the Ramsey variations in unit contributions from the nonlocal network services whose input local service prices cannot be suitably differentiated. The welfare losses from this approach would be larger, the greater were the variation in the corresponding levels of the right-hand side of equation (74), and the larger were the elasticities of demand for the affected services.
Another approach is to differentiate the input local service charges in a manner that is less precise, but that requires less costly monitoring. For example, even if final uses $i$ and $j$ could not be feasibly distinguished, it might still be feasible to monitor the firm’s total revenues obtained from the sale of those services. Then, total charges for input local services could be levied according to the formula

\[
\left( S^i + S^j \right) C^i + r \left( p^i S^i + p^j S^j \right)
\]

This mechanism combines a non-differentiated unit price for the input local services with a proportional levy on revenues. It is equivalent to setting $a_i = C^i + D^i$ and $a_j = C^j + D^j$. Generally, this pair of input local service prices would be imperfectly differentiated by final use. Nonetheless, equations (73) and (72) show that it is consistent with both the desiderata and Ramsey optimality in the special case in which services $i$ and $j$ have the same elasticities of demand and the same degrees of scale economies, albeit different marginal costs.

In general, the less precisely the technical access prices can be differentiated by final use, the more serious will be the welfare losses from the requisite compromises between the desiderata and the Ramsey pricing. The lesson from the above example is that more complex payment mechanisms may be socially desirable when precisely differentiated rates are infeasible.

Thus far, the treatment of specific technical access prices has rested on an incrementally cost separable partition of the nonlocal network services. Unfortunately, the implications of the desiderata are less clear for the input local service prices associated with the individual members of one of the subsets in such a partition. For specificity, suppose that the subset of services $A$ is part of such a partition, so that the desiderata imply that equation (71) holds for $A$. Suppose, further, that $A$ is comprised of services 1 and 2 that are not incrementally cost separable because their production involves a common transmission facility. Thus

\[
C \left( N, S^1, S^2 \right) - C \left( N, S^1 S^2 - S^1 \right) > \left[ C \left( N, S^1, S^2 \right) - C \left( N, S^1 S^2 - S^2 \right) \right] + \left[ C \left( N, S^1, S^2 \right) - C \left( N, S^1 S^2 - S^1 \right) \right]
\]

However, it will be assumed that the local service inputs for these nonlocal services are incrementally cost separable.

Then it follows from equation (75) that for equation (71) to be satisfied for the subset $A$, at least one of these relationships must hold:

\[
a^1_i < C^i_e + \left( p^i_e - C^i \right), \quad a^2_i < C^i_e + \left( p^i_e - C^i \right).
\]

Thus, given that the desiderata are fulfilled over the subset $A$, they cannot be satisfied for both of the individual services that comprise $A$. While the desiderata give the strong guidance embodied in equation (71) for $a^1_i S^1_e + a^2_i S^2_e$, they are as yet silent in this case on the individual levels of $a^1$ and $a^2$.

A persuasive case can be made for avoiding

\[
a^1_i > C^i_e + \left( p^i_e - C^i \right)
\]

Such a high level of an input local service price could exclude from the market an entrant with costs below the incremental costs of the network, to the detriment of consumers. And, even if the network would find it advantageous to encourage such an entrant, once its potential were clear, a technical access price such as that in equation (77) might discourage the technological or marketing efforts needed to expose the potential.

Technical access prices such as those in equation (76) can yield incentives for entry by firms with costs above the network’s incremental costs. This possibility cannot be eliminated for all services, simultaneously, given that some of them share common transmission facilities. The lesson is that responsive flexibility in technical access pricing is required for the preservation of industry cost efficiency. The network may have to realign its prices for input local services in response to potential entrants in order to deter entry that would raise total industry costs.

However, the theory has suggested limits on such responsively flexible prices. First, for every subset of nonlocal network services, the analogue of equation (77) should be avoided. That is, the technical access charges over any subset should not exceed the sum of the network’s incremental costs of providing the input local services and the contribution to the network’s profit obtained from that subset of nonlocal services. Second, this inequality should instead be an equality, as in equation (71), for any subset that is part of an incrementally cost
Network Access Pricing

It should be noted that a network operator who were to experience some gain from additional profits and who were not part of return regulated would have private incentives to set technical access prices that fulfilled the desiderata. It is not surprising that incentives exist to deter industry cost raising entry. However, incentives also exist to encourage the entry of firms with costs below the network's incremental costs. The lost contribution to network profits from the sale of nonlocal network services can be recovered from technical access charges that satisfy the above guidelines. And total profit can be increased if the entrant charges consumers lower prices for its nonlocal network services than the network was able to with its higher costs.

Currently, if the network were rate of return regulated and experienced the Averch-Johnson effect, it would have incentive to deter socially beneficial entry. In its comparison of an entrant's costs with its own incremental costs, the network would have a profit incentive to utilize a shadow price for its capital below its actual market cost of capital. Thus, the network would impute private incremental costs to its nonlocal services below the social incremental costs. Consequently, it would gain by deterring socially desirable entrants that have costs between the private and social incremental costs of the network. And it would have incentives to set technical access prices above the levels given by equations (70) and (71).

There is one institutional framework that may promote the vitality of competition in nonlocal services without necessarily causing harmful side effects. This involves separating the network, at least as far as its financial flows and records, into two divisions. The local division would sell network access and local services to consumers, but it would stay out of markets for nonlocal network services. However, it would sell technical network access services to the other, nonlocal network division and to any firms who cared to enter the markets for nonlocal services. It would be required to offer its technical network access services to the local division and to the potential competitors of that division on the same terms. The nonlocal division and its potential competitors would utilize their purchases of technical network services to produce and sell nonlocal network services.21

It is fruitful to explore the forms taken in this framework by the technical network access prices and their guidelines that were analyzed above. Here, it is the local division that levies the final use differ-entiated technical network access prices $a_s$ and $a_T$. The nonlocal division sees these charges as part of its costs. Thus, this rearrangement of equation (70),

$$p^{ST} = C(N,S^T,S^T) - C(N,S^T + S^{ER},0) - c_r + a^T + a_T,$$  
(78)

can be interpreted to say that the revenues of the nonlocal division are just sufficient to cover its costs. For the interpretation to hold, its costs must be defined as the networkwide incremental costs of $S^T$, exclusive of input local services, plus the charges incurred for the purchase of the requisite technical services from the local division. Furthermore, the generalization of the undesirable condition in equation (77) can be rearranged as follows:

$$a_s^{ST} < C(N,S^T,S^T) - C(N,S^T + S^{ER},S^T - S^T) + a^T_s S^T.$$  
(79)

In this framework its interpretation is a familiar one — that revenues from services $ST$ fail to cover the division's incremental costs of producing $ST$. Such a case is essentially one of cross-subsidization, the dangers of which are well understood.

Finally, if all the network prices together yield revenues that just cover total costs, then equation (78) is equivalent to

$$a_s + N^*E + p^{ST} + a^T + a_T = C(N,S^T,S^T),$$  
(79)

This says that the revenues earned by the local division are equal to the costs of providing network access and local network services for consumers and for the nonlocal division. In particular, the revenues from the sale of technical network access services exactly compensate for the relatively low level of revenues obtained from network access prices that reflect network externalities. It is important to note that the costs imputed to the two network divisions in equations (78) and (79) need not correspond closely to accounting notions of costs. In particular, $C(N,S^T + S^{ER},0)$ are the costs that would be efficiently incurred if no nonlocal services were produced by the network. Hence, if the network were actually producing such services, then this cost figure would be a hypothetical one. Also, the incremental costs in equation (78) are calculated from this same hypothetical base.

However, the imputed costs in equations (78) and (79) of the two divisions do sum to the total cost of network operations, $C(N,S^T,S^T)$. And if the production costs of the network were separable, with
\[
C \left( N, S^t, S^t \right) = C \left( N, S^t + S^{te}, 0 \right) + T \left( S^t \right),
\]

then the imputed division costs might well be directly accountable. In fact, equation (80) would be likely to hold only if the production processes of the divisions were physically separable. In this case, the divisions could actually be separate organizational, financial, and legal entities.

In contrast, economies of scope between nonlocal and local network services and the concomitant physical or human capital utilized jointly by the divisions would imply that equation (80) does not hold. In this case, splitting total network costs between the divisions would require economic and engineering analyses in addition to mechanizable accounting procedures. This effort would be more important, and more difficult, the more significant were the economies of scope and the jointly utilized factors of production. Nonetheless, equations (78) and (79) provide a firm conceptual basis for the task. And the fact that equation (80) would obviate the task suggests that the boundary between local and nonlocal services, and the boundary between divisions, be drawn in a manner that minimizes the capital that straddles it.

Conclusion

Whether or not the network is thought to have two divisions, it is now evident that the economics of network access and local network services are quite different from, and yet inextricably related to, the economics of nonlocal network services. Network externalities are an intrinsic characteristic of networks, and they cause public interest prices of consumer access to be low relative to the associated physical costs. The consequent revenue shortfalls must burden all other network services. Yet, the value of those services to their consumers is enhanced by the size of the population of subscribers enabled by the low access prices. It is these countervailing effects that the optimal network access prices balance.

Interconnection of competing vendors of nonlocal network services is often technologically feasible. Technical network access prices best serve the public interest if they render profitable entry that curbs industry costs and, simultaneously, renders unprofitable any entry that is cost increasing. Prices for technical network services that fulfill these desiderata must exceed the costs caused by the purchased services. The reason is that the prices the network charges the consumers of its local and nonlocal services generally reflect contributions to the costs of the network access that make the network services possible and valuable. Hence, the prices charged potential entrants for technical network services must reflect these same contributions, on as close to a service-by-service basis as possible.

Where such technical access fees are also viewed as internal transfer prices between local and nonlocal divisions of the network, it becomes clear that they afford equal opportunities to the nonlocal division and to its potential competitors in nonlocal services. At the same time, it is these fees that complete the coverage of the vital costs of subscribers’ network access.

Notes

2. This effect has been noted and analyzed from a variety of insightful viewpoints. For example, see Roland Arle and Christian Avermous[1973], Jeffrey Rohlf[1974], Lyn Squire[1973], and Roger Klein and Robert Willig[1977].
3. The clearest example is provided in the telephone industry by the specialized common carrier firms that offer communications services through their own long distance microwave links that are interconnected at both ends with the local facilities of Bell System and independent operating companies. Analogously, the Maigrum service is produced by Western Union through the interconnection of its electronic links with the delivery system of the U.S.Postal Service. Also, much surface freight transportation is accomplished through the interconnection of rail networks owned by different firms and through the interconnection of rail networks with truck routes.
4. See Rohlf[1974].
5. See Robert Willig[1978] for the development of methods to measure the welfare effect of quality changes by means of incremental consumer’s surplus.
7. In this measurement, one must not confuse two conceptually distinct effects of network size on market demand for network services. First, as subscribers join or leave the network, their own demands for network services will augment or be subtracted from market demands. Second, the demands of inframarginal consumers, whose whose network access remains unchanged, may be affected by network size. It is only the latter effect that should be included in the incremental surplus calculation of network externalities.
8. See Hal Varian[1978] for a complete treatment of the standard theory. The analogous theory of firms, in a form amenable to analyzing network externalities that affect them, is available in Willig[1978]. Essentially, all results derived below for consumers hold for firms as well. Throughout, for the sake of relative simplicity, many possibly important features of reality are assumed away. These include consumers’ uncertainties about prices and about their own income and demands, dynamic effects, non-
recurring charges for joining the network, and the availability of less convenient and less costly alternative modes of access to the network. See Willig [1978, 1973] for more detailed expositions and tests of this type of condition.

9. The definitive and most general treatment of this function, when only prices and income are at issue, is due to Leonard Hurwicz and Hirofumi Uzawa [1971]. The extension to such variables as product quality or N is developed in Willig [1978].


11. See Willig [1978] for the proof of this assertion in a different, but mathematically equivalent, context.

12. The insufficiency condition asserts that \( \phi \) remains bounded as \( t \) increases without bound.

13. This solution is derived by straightforward, but tedious, application of the methods developed in Willig [1976b].

14. These approximation techniques are the same ones utilized in Willig [1976a].

15. This follows from the results reported in Willig [1979].

16. The investigation of this case was stimulated by some unpublished work by Rolfs on what he termed "caller-called externalities" in the context of the telephone network.

17. Rolfs argues that this is the case in telephone communications.

18. How to link the magnitudes of these understatements to observable demand data must here remain an intriguing unsolved problem.

19. The precise condition is that \( p \) and \( N \) be separable from \( m \) in the function \( A \). This condition is satisfied when all network services have equal and constant income elasticities of demand. It is also satisfied in many other cases described in Willig [1978].

20. Note that, for analytic convenience, it is assumed that all functions are differentiable. In particular, the discreteness of the set of potential network subscribers is ignored.

21. The solution is guaranteed to be unique if the Jacobian of the demand function \( (\partial F) / (\partial x) \) is everywhere nonsingular.

22. Of course, equations (55) and (56) are equal if all variability in valued incoming flows is internalized, as in equation (27).

23. This follows from equation (27) because the model a consumer's entire value of network access is captured by his surplus in network services. If this value were not equal to the price of network access, the consumer would not be a marginal subscriber.


25. This normative viewpoint was first applied by Frank Ramsey [1927] to the determination of optimal taxes. See Baumol and Bradford [1970] for a history and survey of this approach to pricing, and Willig and Elizabeth Bailey [1977] for an application to the prices of telephone services.

26. For example, the numbers calculated in Willig and Bailey [1979] for long distance telephone services and for postal services would imply in both cases that this adjustment factor were greater than 8.

27. For example, see Peter Diamond and James Mirrlees [1971].

28. This strong assumption is made in the interest of brevity. It can be weakened in a straightforward manner to accommodate quality differences that are equivalent in consumers' views to differences in "full prices." See Willig [1978].

29. For example, in the telephone network, the ENPIA tariffs for local services purchased by competing carriers are higher than the rates faced by final consumers. In the rail network, "joint rates" that apply to interlined freight are typically lower than the rates that apply to local movements along the same route.

30. These desiderata would be unimpeachable in a "first best" context. However, the results of Ronald Braeutigam [1979] imply that if an incumbent is subjected to a binding profit constraint, and if potential entrants are not, then social welfare is not necessarily served by industry cost minimization. Instead, it may be socially preferable for some outputs to be produced by the incumbent at higher costs than would be incurred by the entrant. The reason for this is that, on the margin, the incumbent's profits are more socially valuable than the entrant's due to the incumbent's profit constraint. However, in our context this effect does not arise because, as will be seen, the desiderata require technical access prices to protect the incumbent's profits. As such, the social values of profits do not play a role, and the desiderata are hence consistent with the Ramsey as well as the first best point of view.

31. See Ferran [1979] discusses the same institutional framework in an analogous context.

32. See Gerald Faulhaber [1975].

References
Network Access Pricing

Lawrence Garfinkel

Certain aspects of network access pricing may be viewed from the perspective of various pricing plans in the intrastate arena which have been designed to respond to and anticipate the environmental and regulatory changes that are taking place. They are the subject of this paper. In a major sense it is in the state regulatory sphere where the trauma of change affects the mass of consumers, the regulators, and the industry. Future change will require other responses, but many of the plans which the Bell System and the telephone industry have already undertaken are in accord, as I see it, with many dimensions of that future. In trying to emphasize many of the practicalities and considerations in the implementation change, I feel that at the heart of my remarks are the best interests of the consumer, especially the average consumer who is often unaware of the nature of the change or does not fully understand it. However, he or she will react both in the regulatory arena and in dealing with the telephone companies. That relationship and interaction shape our ability to translate theory into a practical pricing structure.

Two important points must be considered in any future restructuring of exchange rates or when thinking about the concepts inherent in network access pricing. First, pricing of basic local network access is inextricably tied to social goals and public interest concerns. The con-
cept of universal service is a primary legislative, regulatory, and industry objective, and it should remain so. Second, pricing of basic local network access, which is currently defined as included in exchange service, involves considerations which transcend strict economic theory and must recognize the public interest as a dominant concern. These points continue to be important, even though the total revamping of the Communications Act of 1934 is being considered (see Exhibit 1).

Expressions of national purpose that are contained in the Van Deelen Committee rewrite of the Communications Act reinforce the type of benefit which, because of past pricing practices, has historically accrued to basic network access. The level of this flow of contribution from message toll and vertical services to the access line has been established by studies done by the Bell System. The average monthly cost of a residence line and local usage is $16.15, while the average monthly local basic rate is only $9.00; therefore, the shortfall of $7.15 is covered by the contribution flow from other areas of the business (see Exhibit 2). Certainly any reduction in explicit contribution would thrust an increasing burden on basic exchange service, a burden which is already (and will be in the future) compounded by inflationary pressures and the erosion of contribution caused by the entry, which is certain, of competing firms into vertical services and the toll market.

Again, the potential loss of contribution to basic exchange services, and consequent substantial increases in the access line revenue requirements, motivated recent filings for exchange network facilities for interstate access (ENFIA).

Several intercity services currently offered by other common carriers (OCC) are substitutable for message telecommunications service (MTS). These other intercity services require connection to exchange facilities for call completion. However, unlike MTS, these services do not provide a subsidy through a separations process to sustain local service at affordable rates. In recognition of this disparity, the ENFIA tariff was filed in order to establish the principle that all intercity carriers using local exchange facilities must pay for such usage on an equal basis.

I would like to clarify one point which undoubtedly will be raised about ENFIA: The tariff is not applied to foreign exchange (FX) and common control switching (GCSCS) services. This is because a modification of the Separations Manual would be required to extend the ENFIA concept to those services. A manual change can be accomplished legally only through a federal and state joint board proceeding, not by the negotiators who were parties to the ENFIA tariff resolution.

After this background, a few terms need to be defined. This is important, because when a number of parties use the same terms with variant meanings the results are often chaotic.

### Exhibit 1. COMMUNICATIONS ACT OF 1978—H.R. 15015

**ACCESS CHARGE, UNIVERSAL SERVICE COMPENSATION FUND**

Sec. 334. (b) The Commission shall establish and administer a Universal Service Compensation Fund in order to maintain toll telephone service and local exchange telephone service rates at affordable levels and to insure the nationwide availability of basic voice telephone service.

### Exhibit 2. 1975 RESIDENTIAL COST STUDY

Network access pricing (NAP), I would suggest, is best understood as a disaggregation of service. NAP, as a pricing concept over the past
few years, has been variously called, among other things, "unbundling" or "stand alone" pricing. The term *disaggregation of service*, in my definition, means the separation of existing tariff rate structures into individual component rate elements (see Exhibit 3). These rate elements will then be established for those products and those segments of services that can be most meaningfully identified and cost supported. This definition of disaggregation of service is consistent with both legislative and regulatory objectives.

Another controversy surrounds the term "public interest." Too often, it is used as though the term represented a homogeneous or monolithic entity. However, there are often divergent publics. To illustrate very simplistically, in the telecommunications market there are residence and business interests, which may be the same but in many cases are not. Frequently, one party proclaims that an action is "in the public interest" when it is actually in the interest of the business segment to the detriment of the residence consumer, or vice versa.

An abstract from the 1977 AT&T Annual Report about pricing policy is shown in Exhibit 4. Bell's policy is to price discretionary services above relevant costs so that their revenues contribute to the joint and common costs of the firm and provide benefit to other services, particularly residential network access. Our purpose, therefore, is to price these services at a level that provides optimum benefit to the general user of exchange telephone service. In a competitive environment, the ability to price in this manner is constrained, and the rates for these competitive terminal sets must correspond more directly to their costs.

The first major attempt to disaggregate the functions and costs included in traditional, basic local service was the plan to charge for directory assistance. Historically, costs for directory assistance were spread across all rate payers regardless of their individual use of the service. The plan was first implemented by Cincinnati Bell in March 1974, and it is now used in 25 state jurisdictions.

Studies that led to that first directory assistance charge plan in Cincinnati indicated a marked disproportionality in demand: 20 percent of the customers generated 80 percent of the traffic. Furthermore, 43 percent of the consumer body did not use the service at all, and 78 percent of the numbers requested could be found in the existing directory. The volume of directory assistance calling had increased 102 percent from 1962 to 1972, but net customer growth was only 35 percent.

The plan afforded a three-call allowance per month (to cover those instances when numbers might not be in the directory) and a rate of 20 cents per call for each call in excess of that. The allowance level was based on usage studies and was chosen in order to preclude a charge condition for 80 percent of the customers. It was anticipated that there would be a reduction of use by consumers, so that approximately 90 percent of the residence customers would still not pay any additional charge.

Almost five years after implementation in Cincinnati, the results of the plan have more than supported this prediction. Calls to local directory assistance were reduced by approximately 80 percent. Today, 10.2 percent of business customers and about 5.2 percent of residence subscribers actually pay for directory assistance calls, which means that only 5.4 percent of all customers make more than three such calls per month. For the customer who receives a bill for overcalls, the average charge is $1.16, a figure which has been remarkably constant since the plan was introduced.

Studies in other parts of the country revealed an impressive consistency in market use of directory assistance prior to introduction of a charge plan. The highly skewed nature of the demand for directory

---

**Exhibit 3. DISAGGREGATION OF SERVICES**

The separation of existing tariff rate structures into stand-alone component rate elements. These rate elements will be established for products and those segments of services that can be most meaningfully identified and cost-supported.

**Exhibit 4. BELL SYSTEM PRICING POLICY**

(ABSTRACT FROM 1977 ANNUAL REPORT)

It is the Bell System's policy to establish rates that will bring telecommunications services within the economic reach of more and more people, thereby enhancing their usefulness to everybody. To this end certain services — mostly services for business and discretionary services for residence customers — are priced so that their revenues contribute to the joint and common costs of other services, particularly residential exchange service. Our purpose has been to price each of these services at a level that will provide the optimum benefit to the general user of telephone service. To the degree that competition forces us to relate our rates for these services more directly to the costs involved, local exchange rates will rise, thereby jeopardizing the historic trend that has brought telephone service to 95 percent of American households.
assistance is widespread. Therefore, it is not surprising that consumer response to charging for excess calls has been similar in the other 25 state jurisdictions in which a rate for discrete directory assistance use has become effective. Examples of the range of response are detailed in Exhibit 5.

<table>
<thead>
<tr>
<th>Exhibit 5. DIRECTORY ASSISTANCE RATE PLAN — STATE RESULTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>State</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>Overall Charge</td>
</tr>
<tr>
<td>Repression</td>
</tr>
<tr>
<td>Customers Billed</td>
</tr>
</tbody>
</table>

Local directory assistance call reduction among the random and typical states represented ranges from a low of 45 percent to a high of 80 percent. The number of customers billed ranges from 5.3 to 8.2 percent, and the average bill ranges from $0.89 to $1.47 per month. In the 26 jurisdictions where the plan is in effect, labor and other cost savings are estimated to be $115.2 million annually. The customer may choose between looking up the number or paying for the directory assistance call when his allowance has been exceeded. The discretionary choice he makes is a function of the value he attaches to the service in view of its price.

Not only the charge for directory assistance, but also the particular charge per call seems to have had an effect on consumer behavior. This is reflected in the lower use of the service and in the billed revenues. In one state where the charge per call is 10 cents, the overall reduction in local calling to directory assistance is 45 percent compared to the 60-80 percent rates in other states.

Yet there are 29 jurisdictions in the Bell System where this plan, despite the beneficial consequences to the vast majority of the consuming public, has not become a reality. In many instances, no reasons were given by commissions for rejecting the proposal, although the emotionally charged atmosphere in the regulatory proceedings was probably responsible. In fact, a bill has been enacted by the Iowa legislature which prohibits implementation of any such charging structure.

The lessons learned from directory assistance illustrate many of the differences between the theoretical justification of a concept and the practicalities of bringing it into reality. In the process of change, an efficient and theoretically logical proposal often encounters emotionalism, misinformation, and self-interested opposition. Dealing with the public in such a charged environment is difficult. An educational process is required, and persistent efforts must be undertaken. Putting good theory into practice cannot be done simplistically or overnight.

Work is under way throughout the telecommunications industry to disaggregate basic service in another manner. The project is known as measured service. The basic concept is to separate network access from usage and to create a two-part tariff pricing structure with a fixed fee for network access and a variable charge which is a function of individual usage, thus allowing the customer to have more control over his monthly bill by making adjustments in his usage. The fixed network access rate can then be kept at a price to the customer which is consistent with the legislative and regulatory purpose of universal service. The measured residence access line to which contribution flows from other services is the benefitted service.

In this instance, network access means a network telephone service that provides the capability of information transfer via access to and from the local and toll network, including appropriate maintenance. This definition of network access recognizes the difference between, on the one hand, the necessity of having access to the network and being accessed from the network and, on the other hand, making calls once network availability is obtained. Usage of the network would be provided at a rate level that makes a contribution to the access line. In other words, recognition will be given to different demand functions between having access to telephone service at the lowest reasonable rate in order to be a member of the larger community, and the opportunity to make one or many calls of varying duration. Looking to the future, when the impact of competition will reduce the contribution of residence access, local usage will then provide another revenue source to keep residence access line rates at affordable levels.

Rising exchange rates, of course, are not solely attributable to contribution loss through competitive entry. Inflation, in a capital-intensive industry such as telecommunications, is compelling. A disaggregated two-part tariff affords a means of generating additional revenue growth concurrent with usage costs and increased volumes. But there must be a valid basis for creating such a tariff derived from appropriate analysis and quantification based on data about individual customer use of the local network.

In the Bell System, the average calling rate per main station over the past two decades has been increasing at the rate of 1.5 percent per year.
But this growth and the distribution of calling between the residence and business customer bodies represents wide ranges of use. Exhibit 6 is based on a recent study of 79 offices scattered throughout the country and shows the functional form of that use which is common to all offices. On the high use side, 20 percent of the residence customers make approximately 45 percent of the calls; on the low side, 20 percent of the customers make about 4.5 percent of the calls.

Call duration, or holding time, for completed messages also has some unusual characteristics. Generally, 50 percent of all residence customers have average holding times of less than four minutes, while 7 percent average less than two minutes per call. However, some calls have lasted as long as 11.25 hours, and the most remarkable call that was recorded and charged for as an unlimited single message unit was between two points in Manhattan which lasted for 43.5 days. Holding times are shown in Exhibit 7.

By disaggregating usage from the access line, the tariff structure can reflect charges for call frequency, duration, distance, and time of day. Thus, measured service not only draws a closer revenue-cost relationship, but also sends proper pricing signals for discretionary calls to consumers. One option is a threshold or low use access line rate designed to sustain universal service level. Relating the distribution of charges to usage characteristics also creates a pricing structure that is more fair and equitable for all customers (see Exhibit 8).

Exhibit 6. MEASURED SERVICE RATE STRUCTURE
- Charge For Access
- Charges For Usage
  - Frequency
  - Distance
  - Duration
  - Time-Of-Day
- Customer Options

Given divergent market demand, it would seem plausible that a reasonable tariff could be developed and rapidly implemented. After all, such a means of charging would be consistent with standard economic theory and certainly more equitable to the customers. However,
of the residence consumers would have received the same or a lower bill under the proposed tariff.

Revenue redistribution implicit in usage related pricing can favor a majority of consumers. Results in a local measured service trial currently being conducted by General Telephone and Electric and other exchanges in Illinois show that after a year of measured service rates, 66 percent of the residence consumers have experienced no rate increase or a reduction, compared to rates charged in comparable exchanges in 1977.

The psychological barriers of customer suspicion and apprehension, fear of change, and constraints on personal freedom are real. There is a need for education and implementation processes which facilitate the transition from theory to practice of a sound plan and one which is in the public interest. Plans for the introduction of local measured service, at a minimum, must demonstrate the need for the service, portray distributive effects on various socioeconomic segments, and involve consumers and regulators in the process of transition, from planning to implementation.

Another dimension of network access pricing and disaggregation involves the manner of charging and accounting for the customer premises work (including wire and labor) needed to provide network access in the manner the customer desires. At the present time, certain costs related to inward movement, including the establishment of network access and provision of inside wiring, are capitalized to Station Connections Account 232 in the Uniform System of Accounts. This FCC accounting treatment was originally consistent with the objective of maintaining low entry rates in order not to inhibit universal service development.

In recent years, both the FCC and the Bell System have become concerned over the growth in this account and the associated burden on the general body of rate payers, largely due to increased customer mobility (for example, reconnection of service) and the provision of discretionary services (such as extension of plumbing). For these reasons, AT&T submitted a proposal to the FCC in 1977 which would, for future additions, change the accounting from capitalizing all of these costs to expensing approximately 75 percent of them in the year incurred. The proposed phase-in of the changes is shown in Exhibit 9.

However, as Exhibit 9 shows, not all the costs related to station connections would be expensed. In order to maintain a minimum working service on the customer premises at an affordable rate, the costs of wiring from the pole to an initial station outlet on the premises would still be capitalized and recovered in the basic monthly rate for service. Again, this would ensure that the residence "access line" price would be consistent with universal service objectives, because the full access line would be the benefited service. The customer can choose the provider of the station, although we continue to believe that it is in the overall public interest for the telephone companies to provide at least one telephone instrument to the customer. The question of who provides the set is not pertinent to the proposal filed for expensing station connection costs or the pricing of network access in accordance with the proposed new capital expense split.

Exhibit 9. CHANGES IN ACCOUNTING

<table>
<thead>
<tr>
<th>Year</th>
<th>Phase</th>
<th>Elements To Be Expensed</th>
<th>% OF Acct. 232</th>
</tr>
</thead>
<tbody>
<tr>
<td>1979</td>
<td>1</td>
<td>Residence Extension</td>
<td>18.6</td>
</tr>
<tr>
<td>1980</td>
<td>2</td>
<td>Assignment and Test</td>
<td>15.4</td>
</tr>
<tr>
<td>1981</td>
<td>3</td>
<td>Business Extension</td>
<td>22.4</td>
</tr>
<tr>
<td>1982</td>
<td>4</td>
<td>Res. and Bus. Main</td>
<td>18.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reconnect and Reinstall</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Apparatus Handling</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>74.5</td>
</tr>
</tbody>
</table>

| Residence and Business Main (Initial Installation) | 25.5 |
|                                                   | 100.0 |

There are other aspects to the continued capitalization of the access line termination. (1) Nearly 45 percent of Bell customers (and a greater proportion of the consumers served by the Independent Companies) have only one working telephone. For these customers, the accounting proposed by the Bell System would be the most economical plan and support the concept of sustaining universal service. (2) The installation of a dedicated access line termination in the home would be available to all new customers moving to a location. The life characteristic of this investment creates a continuing asset which is capitalized only once through its usable life. (3) The proposed expensing plan is consistent with the definition of basic service, in that the capitalized portion
provides the capability for the customer to place and receive calls over the network. (4) In the conversion to the expensing of Account 252, the short-run revenue requirements or earnings problems encountered by the industry would be reduced by continuing to capitalize the network access line termination rather than expensing the entire account, including the network access line termination. In the peak year of the plan, the revenue requirements due to the change from capital to expense as proposed would alone amount to approximately $1.6 billion when compared to revenue requirements under present accounting.

As a result of expensing, not only would the rate base be reduced in the long run, but also the change would facilitate a more cost causative rate structure; that is, customers causing the costs would pay for them.

Moving beyond the main station connection, at the present time the provision of extension service is aggregated to include two rate structure elements: the station and the inside wiring. In the future these services will be disaggregated. In conjunction with the expensing plan, when these two elements are priced, the future expenses costs of placing extension inside wire could be recovered by a nonrecurring charge at a compensatory level for that activity. Thus the future revenue recovery would match the cost in the same accounting period.

A dilemma occurs in reconciling this future pricing plan for new extension wiring, that is, a nonrecurring charge, to the existing or embedded cost of wire which can and will often be used and reused by existing and future customers. In the case of the Bell System, the embedded base, including inside wiring, represents estimated to be a large portion of this sum is associated with the vertical (nonbasic) equipment investment. While it is planned that this investment would be amortized and ultimately removed from the company books during the transitional period, there is a need to continue to recover the cost of this installed capital and earn on it. At the same time, with the alternative of obtaining telephone sets from outside the Bell System, the customer will want to use this installed wiring to have additional locations from which to access the network. Since we rely on the customer to report intended use of the installed wire, there will be increasing customer unwillingness or reluctance to report because of the recurring rate application. There is an investment still associated with each previously installed outlet, and while it is logical to earn on this investment, the ability to secure adequate revenue recovery is a contentious issue.

Several possibilities can be suggested for coping with the resultant revenue deficiency from insufficient customer reporting. First, in a few states where this issue has surfaced, the revenue requirement on inside wiring has been considered to be a common earnings requirement shared by the general customer body. In these situations, the recurring rate for inside wiring associated with extension service has been eliminated and that revenue burden shifted to the monthly rate for basic service. For example, when this treatment was filed by General Telephone and Electric of the Northwest in the state of Washington, the basic monthly rate for network access was increased 25 cents. This approach poses several problems. While this figure may seem small, amounting to an increment of $3.00 a year, it is only one pound increase on the price of the network access line, that is, basic service. In addition, there is the matter of consumer equity. If the extension wire charge is eliminated completely, then a customer, for example, who currently is paying for three extensions, may receive a net decrease in excess of any increase for the basic monthly service. On the other hand, a customer who had no extensions receives no benefit from the increase in basic monthly service; instead, he must pay an increase to support another customer's optional use of extensions.

Second, the Bell System does not propose to dispose of its interest in inside wiring. Yet, some have suggested that it might be appropriate to sell the wiring to existing customers or create a one-time payment for its use. The pitfalls here are the customer's reluctance to pay a one-time charge for something presently in use and being paid for, and the telephone company's inability to recover adequately the embedded investment through a one-time charge or sale plan. This latter point was recognized by the New York Public Service Commission official: "It appears that the telephone industry would have difficulty in selling existing inside wiring at or above net book cost." Use of this quotation is not meant to imply that the New York commission has taken a position opposed to sale of inside wire; this matter is still the subject of hearings. The point is made to raise the difficulty of using this approach as a means to recover costs adequately.

Third, the position of the Bell System is to apply the recurring rate for extension wiring and collect it. We recognize this may be difficult, but it is the preferred alternative in view of cost causation. After expensing of the station connection account begins, the problem will be eased somewhat; when the amortization of the embedded base for inside wiring is complete, we will no longer be able to substitute the recurring wire rate.

Interestingly, this is one practical reason for introducing network access pricing, and there is ample theory to guide us. When extension service is disaggregated, there will have to be a major review of the
pricing of all telephone sets, of whatever style — 500, 2500, Princess, Trimline, and so forth. These prices will be derived by incremental analysis costing methods used in the Bell System, which relate market demand and relevant direct resource costs to determine the appropriate rate level.

However, the repricing effort is complicated by the fact that, in the past, discretionary terminal equipment, when used for extension station service, was priced on the basis that the telephone company would provide the complete extension service and receive the extension service rate. The incremental difference in investment was the base for computation of the filed Princess, Trimline, and multibutton set rates. Since under network access pricing (disaggregation) there will no longer be an identifiable extension service, all terminals must be repriced to eliminate the differential consideration and to relate directly the level to the actual costs of each instrument.

These are certainly not all the implications of network access pricing, but they do highlight many of the pricing plans and practical problems that the Bell System and the whole telecommunications industry have been working with for some time. Directory assistance charge plans, measured service for local network use, expending currently capitalized costs, disaggregation of extension service, and "product only" pricing for telephone sets do have significant consumer implications. While sound economic theory serves as a general guide in structuring objectives, the interplay of consumer attitudes, regulatory proofs, marketplace competition, and environmental and/or social interests creates many practical concerns in implementation of theory. At times, it seems that some people apply theories in a manner that is indifferent or insensitive to the problems and concerns of consumers. Theory must stay in touch with the perceptions and understanding of the consumer or it will be irrelevant.

Public Policy Issues of Network Access Pricing

Andrew Margeson

It is helpful to recall the public policy debate which gave rise to the discussion of network access pricing. In 1959, the Federal Communications Commission and the courts began issuing a series of decisions opening the interstate communications market to competition. Today, competition extends to virtually every corner of what was once considered a legally protected market. A principal telephone industry argument against this policy has been the alleged adverse impact of competition on local rate payers, particularly those living in rural areas. The telephone companies assert that MTS and WATS (long distance telephone service) are priced approximately 30 percent above the direct cost of providing them in order to make a substantial contribution to the joint and common costs of the local telephone network — basically nontraffic sensitive investment and expenses associated with giving the customer access to the network. In addition, the telephone companies point out that MTS rates are averaged, that is, they vary only with distance and not with the cost of service on particular routes. The industry believes these two results, sizable toll contributions to the joint costs of the local telephone network and toll rate averaging, are mandated by public policy. The argument concludes that competition will destroy the achievement of these social goals because the cross-
elastic impact of the competitive services will eliminate the ability to price MTS above its direct cost and will, in addition, force toll rate deaveraging. AT&T, for example, has said that 80 percent increases in residential telephone rates could result.

The industry’s response was to propose the Consumer Communications Reform bill, the basic purpose of which was to contain competition within “tolerable” limits. In the hearing which followed introduction of the legislation, the issue of the adverse impact of competition dominated the discussion. I think most would agree that competition was discussed much less on its own merits than on the basis of its impact on residential telephone rates.

Although this issue has been debated extensively before the FCC and the Congress, no consensus has emerged as to whether the impact will or will not materialize. A logical question, therefore, is whether this link between the achievement of social goals and restraints on competition is inevitable, or whether there is some means by which the link can be broken so that each can be discussed and decided on its own merits. Parenthetically, I might note that a reading of economic history seems to indicate that the conflict between social goals associated with income redistribution through utility rate regulation and the introduction of competition into those regulated industries arises almost inevitably. During the same period I am discussing, congressional subcommittees considering increased competition in air service were told that it would force the abandonment of service to smaller communities, and those involved with trucking deregulation were told that unrestricted competition would cause poorer service and high rates in rural areas.

In the telephone case, there seem to be several options. (1) Assume that competition would not make significant inroads and that, therefore, the industry could maintain its current arrangements for toll contributions to the joint costs of the local network. (2) Make the achievement of social goals an explicit governmental function akin to any tax and transfer process, with the funds derived from general treasury revenues or some kind of special excise tax on communications. (3) Eliminate the contribution process and allow local telephone companies to charge for access to the local telephone network when used as part of a through toll service (for those steeped in industry lore, return to something like board-to-board separations and settlements). (4) Include all services in the contribution process by requiring some payment into the industry’s revenue pool. (5) Devise some variant or combination of these.

A number of these alternatives were identified in the staff Options Papers prepared in support of the House Communications Subcommit-tee’s efforts to rewrite the 1954 Communications Act. At subsequent hearings, representatives of Southern Pacific and Microwave Communications Inc., two of the major specialized common carriers, agreed that under certain circumstances they would agree to make a contribution to the joint costs of the local network analogous to those made by MTS.

Following those hearings, the subcommittee staff drafted a rewrite of the Communications Act of 1934 which was introduced last January. That legislation provided for a “universal service contribution fund” derived from access charges placed on all services connected to the local telephone network. The new Communications Regulatory Commission is given authority to use the fund “to maintain toll telephone service and local exchange telephone service rates at affordable levels and to ensure the nationwide availability of basic voice telephone service” (section 354(b)).

Meanwhile, the Execuant case had cleared the way for offering direct MTS substitutes by specialized common carriers. Both Southern Pacific and MCI currently offer switched voice services from any touch-tone telephone to any telephone in the cities they serve. In response, AT&T filed its Exchange Network Facilities for Interstate Access (ENFIA) tariff, the stated purpose of which was to charge for toll substitute access on the same basis as access for MTS and WATS. The ENFIA tariff was composed of three elements: a charge for connection to the central office, a charge for local switching and metering, and a contribution to the costs of joint nontraffic sensitive local plant. Combined, the charges would have amounted to $14.25 per line plus 11.2 cents per minute, which is significant compared to an average charge of $0.26 per minute for Southern Pacific’s Sprint Service.

Even though ENFIA claimed to put toll substitute services on a parity with MTS and WATS, this tariff filling caught the FCC at a bad time. In response to the Execuant decision, the commission had established a major proceeding to set long-range policy on intercity competition, including as an integral part of that proceeding the issue of the form and level of contributions to the joint costs of the local telephone network. The tariff review process is quite cumbersome and lengthy, and many participants feared that five or more years would elapse before the commission could establish even an interim charge for toll substitute access to the local network. Accordingly, Assistant Secretary of Commerce Henry Geller suggested a negotiated interim access charge to be applied pending a longer term policy decision. The commission and AT&T agreed, and the ENFIA tariff was voluntarily postponed.
Those negotiations have recently resulted in a three-year agreement between AT&T and the specialized carriers. Its terms are as follows. The charge for a voice grade connection to the central office will be the current tariffed rate of, on average, $4.50 per month. The charge for local exchange switching and trunking will be $37.97, less a message unit credit of $18.14, or $19.83 per line per month. The charges for jointly used subscriber plant are slightly more complicated and are shown below.

<table>
<thead>
<tr>
<th>Step</th>
<th>Combined OCC</th>
<th>Percentage of MTS/WATS cost to be paid by specialized carriers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MTS/WATS equivalent cost</td>
</tr>
<tr>
<td></td>
<td></td>
<td>per month*</td>
</tr>
<tr>
<td>1</td>
<td>6 - 110 million dollars per year</td>
<td>5.5%</td>
</tr>
<tr>
<td>2</td>
<td>111 - 250 million dollars per year</td>
<td>5.5%</td>
</tr>
<tr>
<td>3</td>
<td>Above 251 million dollars per year</td>
<td>5.5%</td>
</tr>
</tbody>
</table>

*This amount is growing. The calculation will be based on actual results for the period.

Rochester Telephone did not concur in the outcome of these negotiations for a variety of reasons, one of which I will mention here. The purpose of the ENFIA negotiations was to promote fair competition and guard against the erosion of support by intercity services for the joint local telephone plant. This was to be accomplished by applying access charges to toll substitutes. Unfortunately, the negotiators decided to confine the talks to ExecuNet-like toll substitutes offered by specialized common carriers. None of the numerous other toll substitutes, such as foreign exchange, common control switching arrangements, and so forth, were included, even though the latter are, for the present at least, more significant competitors with MTS and WATS than ExecuNet or Sprint. In Rochester’s case, for example, we have an order for ten Sprint access lines, but we have many times that number of toll substitute access lines which will not be covered by the interim agreement.

Rochester argued the point very strongly during negotiations, and the following was included in the agreement:

The parties have agreed that there are interstate services in addition to MTS/WATS, ExecuNet-type and Sprint-type services, like FX, CCX-ONALS, value-added data and facsimile services which also utilize the local telephone distribution plant. This agreement establishes charges only for the use of that plant by ExecuNet-type and Sprint-type services. The parties recognize that an overall interim solution to the problem of access to local telephone distribution plant by services not included in this agreement is needed while Docket 78-72 is pending.

Nevertheless, we have some doubt that this commitment will result in timely and equitable access charges for all toll substitutes. We are attempting, independent of these negotiations, to establish that principle for our own company at least.

I want to turn now to the public policy justification for network access pricing. I will use Rochester Telephone as an example. Table 1 depicts some relevant financial information. A major portion of Rochester’s investment and expense occurs in supplying network access to its customers. Over 60 percent of total investment is tied up in terminal equipment and nontraffic-sensitive local distribution plant.

Rochester’s problem, of course, is to recover its total costs. Under current procedures, 26 percent of the joint investment is allocated to toll, and $4.87 per month per access line is recovered in settlements revenues for terminal equipment and local distribution plant alone. After a complete allocation of investment and expenses to toll, Rochester recovers the residual in the form of local rates, which in 1977 amounted to an average of $18.92 per access line per month.

This explains why MTS is unfairly vulnerable to competition in the absence of an access charge. Rochester makes no similar allocation of nontraffic-sensitive access costs to toll substitutes regardless of who provides them.

An obvious approach to resolving this competitive inequity is to fit the toll substitutes into the current process. This is essentially the tack taken in the ENFIA negotiations. Specialized carriers will pay into the toll pool, but as a short-cut they will pay a percentage of the MTS/ WATS contribution per minute of use instead of going through a separate settlement study.

In my view this is unacceptable as more than a “rough justice” interim policy. Even if the agreement were extended to encompass all toll substitute services, as Rochester proposed, I believe the very concept of separations and settlements needs to be revised. The historical process was devised with little concern for achieving economic efficiency. As a result, across-the-board allocations to toll are made without regard to alternative arrangements that might be much more suitable in many cases. Indeed, the fundamental premises of separations procedures is just plain wrong — “relative use” is not the proper basis for allocating costs of nontraffic sensitive plant from the perspective of
either realizing economic efficiency or achieving social goals. The former requires that demand elasticities be taken into account, and the latter should be based on consideration of a need for support on a case-by-case basis. Arbitrarily weighted relative use allocations approximate neither of these.

I do appreciate the critical historical role that relative use separations and settlements have played in achieving universal telephone service. However, one example will illustrate how these procedures distort telephone industry incentives. A basic economic fact of the industry is that plant sits idle most of the time. Rochester's average access line, for example, is in use just 46 minutes a day. Of course, even this figure masks great differences between business and residence use. Suppose some young whippersnapper devises a totally new service—an electronic newspaper, for example. The marketing people consult their black box and find that the new service would generate revenues exceeding costs. Will it be offered? Maybe and maybe not. The decision maker will have to account for the loss in toll revenues resulting from the impact of increased local traffic on the relative use allocation factors. Without belittling the very important historical contribution of these procedures, the time has arrived when we must pay a good deal more attention to efficiency considerations as well.

Before I suggest an alternative to separations and settlements, let me clear up what I see as a conceptual problem. The "joint" costs of network access are joint only because of the way we define services. Toll is viewed as an end-to-end service, and so some portion of local access facilities is allocable. This is similar to saying that when I ride the bus from my home to the office, it is local, but when I ride that same bus to the airport, it is intercity, and therefore my airline ticket should be priced above cost in order to make a contribution to joint bus service costs. Substantial arbitrary cost allocation problems are introduced. I prefer to think of toll as a service that extends only from one toll facility to another or, to return to my previous example, from one airport to another. If this view is accepted, then I see the toll network access point as only one of many local points to which customers have access over the local network. The fact that they will proceed to another city from there is immaterial to me. I am engaged in providing a local service. If the reader accepts this, then he has been tricked into turning the network access pricing problem into the familiar problem of pricing local telephone service. The usage sensitive pricing debate, self-selecting two-part tariffs—all the discussions heard so often—immediately become relevant.

I envision that, under this approach, there would be self-selecting
multipart "local" tariffs which would be applicable to intra- and intercity calling. A simple example illustrates: I might pay a flat rate for network access and a usage charge applicable to all calling; the total cost of a toll call would be the usage charge for local distribution on each end plus the intercity charge.

I believe this approach would have many advantages over present practices, not the least of which would be dramatic savings in administrative costs. The separations and settlements process is exceedingly expensive, particularly for the smaller telephone company. In addition, the policing problem inherent in placing an access charge on "intercity" facilities alone could be avoided. Customers will not be given an incentive to switch from direct to indirect means of access to the local network, which is a serious problem with the ENFIA approach. This suggestion would, I believe, clarify jurisdictional issues that are an increasingly serious problem in the station-to-station separations environment. State regulators would be in a much stronger position to retain their jurisdiction over the local network by accepting total responsibility for its revenue requirement. Finally, such an approach would be much easier to apply in a way that is obviously equitable to all intercity competitors.

One could object that this approach will not always be sufficient to achieve social goals of universal service and affordable toll rates in rural areas. The fact that some independent telephone companies currently retain more than 100 percent of their billed toll revenues supports the magnitude of the problem. Indeed, the argument might be that separations and settlements in their current form were devised precisely because of the need to transfer revenues to companies for which network access pricing would be inadequate.

While I would concede that this may be the case in some situations, I think the appropriate starting place is a careful delineation of what our social goals are. I can accept the idea of subsidizing access to the telephone network when it is necessary to bring it down toward—but not to one-half—the average, as is sometimes the case today. I cannot accept the idea that it is wise public policy to subsidize extensive usage, or directory assistance, or extended area service, or a host of other services beyond basic network access. As for intercity toll rates, I am not convinced that deaveraging will reach unacceptable levels, but I will leave further discussion of this to others.

Assuming that there will be situations in which, by my criteria, additional support will be required, I would suggest a separate fund administered by the federal government and targeted to those areas which require assistance. For theoretical and practical reasons, and following the precedent of the REA program, I would prefer to fund this support from general tax revenues. Some variant of the universal service contribution fund contained in the House version of the Stetson Communications Act might be a good candidate. Alternatively, the REA program might be expanded. In general, however, I believe this social support function should be kept entirely separate from any regulatory responsibilities so that the two do not become confused, as they have in the past.

But is such a proposal at all realistic? I believe so. Basic to the success of such an initiative is the adoption of new forms of local service pricing. State regulators are becoming more and more accustomed to the idea of usage-sensitive pricing, and the industry is gradually acquiring the ability to provide measured rate service. The state regulator is in a key, perhaps the crucial, position. The social repercussions associated with the kinds of change I am suggesting are considerable, and it is the state regulator who will be on the front line, both in making the decisions and in facing unhappy customers. Completing the successful transition to a new environment must involve a willingness on everyone's part to forgo traditional arrangements for extended area service, free directory assistance, optional flat rates, and so forth. We are coming face to face with economic reality. The broader our social goals become, the more expensive and difficult they will be to achieve in a reasonable way. These costs have been hidden in the past, and so have been more palatable, but they certainly were no less real.

One starting place might be to freeze the nontraffic sensitive plant allocation factors which otherwise will continue to grow. Regulators also might attempt to identify those cases in which a company-by-companv application of network access pricing would be incompatible with public policy goals and devise means of dealing with them.

In addition, there are significant opportunities offered communications subcommittees in the Senate and the House. I believe Senator Hollings and Congressman Van Deerlin are quite serious in their intention to revise or rewrite the Communications Act of 1934. In fact, I will go out on a limb and say that I am convinced they will succeed. The prospect of whether we make the transition easy or difficult, however, the forces for dramatic change are already in motion without realistic prospect for reversal. For this reason, if for no other, I am optimistic that an approach similar to the one I am suggesting will ultimately be adopted.
Comments

Ronald R. Braeutigam

Robert Willig has written an interesting and informative paper on a rather difficult topic. For many years regulators have struggled with a resolution of the conditions under which vendors of network services will be allowed access to a telephone network. Two basic kinds of issues have emerged over time. The first category includes technological problems (not addressed by Willig), such as the compatibility of interconnected systems, the possible need for network protective devices, and the division of maintenance and troubleshooting responsibilities among companies involved in an interconnected system. The second set of issues (and the one of concern here) has at its heart the pricing for access to the network. Among other things, access prices will affect the profitability of the existing network as well as the interconnecting companies, the prices consumers pay for services, the extent to which one service of a network might be able to subsidize another, the number of subscribers to the network, the efficiency with which telecommunications resources are allocated, and the variety of services available to users of telecommunications services. These concerns emphasize the timeliness of Willig's paper.

Network Externalities

The author correctly points out that "network externalities and interconnection are basic features of networks that create special complexities in the analyses of prices for network access." He has developed his analysis using a simplified and stylized network that succeeds in representing these major features. In his model a consumer tied in to a local switching center has access to nonlocal network services. A competing vendor of nonlocal network services may also connect with a local switching center, with access to the rest of the network. Technical feasibility of such interconnection is not a problem in the formulation; rather, the central issue is the price charged for this access.

Among the most important contributions of the paper is a formal treatment of network externalities. At the most basic level, network externalities are modeled as a shift in the demand schedule for service for the i-th consumer as the population of network subscribers changes. In the usual case, a consumer's demand schedule will shift out (or increase) as more subscribers join the network. As Willig indicates, the effect of such an externality can be represented using the notion of consumers' and producers' surplus, and it can be measured using observable market demand data. Thus, he succeeds in one of his major objectives by demonstrating "that network externalities need not be considered metaphysical phenomena and that they are, instead, quantifiable and amenable to scientific analysis."

Two additional points of interest emerge in the derivation. First, network externalities become more important as inframarginal consumers (those who would pay prices higher than existing market prices to get service) are more sensitive to network size. Second, the analysis is greatly simplified if incoming network flows have no value to the recipient consumer. Network externalities can still exist in this case because a consumer places value on his ability to call more subscribers. However, if consumers place a positive value on incoming network flows, then Willig shows that the incremental consumers' surplus methodology will provide a lower bound on consumers' gain from an expansion of the network. By a symmetrical argument, the paper implicitly suggests that to the extent that recipients place a negative value on incoming network flows (for example, "nuisance" calls), then the methodology might require some downward adjustment in the measurement of consumers' gain from expansion. 1

First and Second Best Access Prices

With a methodology for measuring the welfare effects of network externalities in hand, the author proceeds with a natural characteriza-
tion of first and second best prices of two types: a price charged for access to each network service and a price for each unit of each type of network service purchased. A straightforward derivation shows that the first best price for each network service is equal to the marginal cost of production of that service, not a surprising result. However, the first best access price for each service is below the marginal production cost of providing access.

The latter result is quite significant. It suggests that if externalities are present, even a firm operating with constant returns to scale will not break even at first best prices. Thus, a second best set of prices (one that maximizes total surplus subject to a break-even constraint for the firm) may be of interest even if there are no economies of scale. Moreover, second best prices become even more interesting if there are economies of scale in production, since both the economies of scale and externalities have effects that make it impossible to break even when the price of access equals the marginal production cost of providing access.

**Access Pricing and Entry**

Willig correctly notes that the execution of a second best pricing scheme may require price discrimination both for access prices and for each unit of service purchased. He further points out that the maintenance of such a scheme may be jeopardized (if not made impossible altogether) if other suppliers are allowed to provide some of the services already produced by the existing regulated firm.

The fifth section of the paper characterizes "fundamental desiderata for technical access prices." The author has assumed that "the network services potentially or actually offered by competing vendors are identical to those offered by the network," and he rests the analysis on the fundamental desiderata that "technical access prices yield profit incentives for the entry of firms that would lower total industry costs, and that they discourage socially undesirable entry." The primary form of entry of concern here is in the supply of some or all of the nonlocal services offered by the network.

The analysis emphasizes several interesting points. First, it may be very difficult to approach second best prices if it is difficult to distinguish among the types of users of local network service. Second, if final uses cannot be distinguished, then more complex payment mechanisms might be used to allocate resources more efficiently. Third, Willig indicates why an extant regulated firm might need some price flexibility in order to discourage entry that would raise industry costs. At the same time, he characterizes conditions that should serve to delineate the boundaries of such pricing flexibility so that the extant firm does not deter entry that will lower industry costs.

This line of investigation is quite interesting. It should be noted that it rests rather strongly on the premise that the industry structure that is least costly is best. There are two reasons to be careful about leapfrogging too quickly from the analysis to the "real world." First, the premise leaves open the possibility that entry today might never be viewed as socially beneficial if industry costs are therefore higher today, even if tomorrow’s industry costs are lower as a result of the experience gained by entry today. Of course, this sort of dynamic activity is difficult to analyze without some means of relating expected future technologies to the present. The issue is not central to Willig’s theoretical analysis, but is relevant to a policy decision as to whether entry should be allowed.

A second reason to be cautious in applying the theory is that entrants may provide services that are differentiated from those of the extant firm. In that case, the welfare maximizing industry structure may not be the cost minimizing structure, since consumers may be willing to sustain higher costs in order to have available a variety of products.

**Conclusion**

Perhaps the most important point of Willig’s paper is that the measurement of the welfare effects of network externalities need not be left entirely to those able to utter mystical incantations. The theoretical underpinnings for scientific measurement have been laid out clearly. The paper succeeds in showing how that theory can be useful over a broad range of problems likely to be encountered in an application designed to determine efficient network access prices.

**Notes**

1. Equation 30 in Willig’s text provides the basis for this statement. In note 19, the author notes that the methodology for assessing the magnitudes of error resulting from the use of observable demand data when incoming network calls have a nonzero value to a recipient remains an unsolved problem.

2. See Willig’s equations 53 and 54 on this point.
Comments

James R. Nelson

The first contribution I would like to discuss is that by Lawrence Garfinkel. Once I recovered from the shock of discovering that the inside wiring provided by the Bell System has cost $8 billion, I decided to organize my comments into two sets: those which are functional from the economist’s standpoint, and those which are functional from the Bell System’s standpoint.

An Economist’s View

Early in this paper, Garfinkel writes: "Pricing of basic local network access which is currently defined as included in exchange service involves considerations which transcend strict economic theory and must consider the public interest as a dominant issue."

This reminds me of a remark attributed to an anonymous congressman: “Every once in a while, I find that I must rise above principle and do the right thing.” As an economist who spends much of his time on matters pertaining to railroads, I find “the right thing” advanced for consideration by Garfinkel in the passage quoted to be the equivalent of sewing oneself in a sack and throwing oneself overboard.

From the beginning, railroads have been expected to respond to "regulation," on the one hand, and to "competition," on the other. As the competition moved off the statute books and onto the highways, one casualty had to be consideration of the public interest "as a dominant issue." Unfortunately, this concept did not expire the very day that plans for the Interstate Highway System were announced. The concept is still with us, in a languishing and moribund condition. If satellite communications eventually do for the telephone companies what the internal combustion engine did for the railroads, at least part of the blame must be attributed to ideas of "the public interest" which exclude "strict economic theory," and therefore, in the present state of economic theory at least, tend to exclude the idea of competition and provide maximum play for regulatory decisions which are guided by some otherworldly idea called the "public interest."

Later in his paper, Garfinkel states: "It is Bell System policy that discretionary services are priced above relevant costs so that their revenues contribute to the joint and common costs of the firm and provide benefit to other services, particularly residential network access. The purpose, then, is to price these services at a level that provides optimum benefit to the general user of exchange telephone service."

In terms of common usage, there surely can be no objection to this passage. And that is precisely the basis for my objection. It is haggling to point out that "joint costs" are not "common costs" and that a rate structure which is designed to respond to the former may provide a completely inappropriate response to the latter. A familiar example of a joint cost situation is a telephone which jangles all day during business hours and, without highly promotional rates, sits forlornly by itself after 11 P.M. An equally familiar example of a common cost arises when John Q. Latecomer, just up the street from me, finally abandons his string and tomato can and decides to install a telephone. His advent may increase the value of my service (a possibility, by the way, which escapes all rigorous cost analysis based on the implicit assumption of a homogeneous product), but it does not automatically give me more service or more service possibility. For telephone service, as for electricity supply, joint costs (and the appropriate pricing measures relating thereto), tend to be related to time, whereas common costs tend to be related to space. The area of common costs becomes, when more carefully examined, the area of decreasing costs. It is precisely in this area, where marginal costs are lower than average costs, that (1) the case for at least a certain amount of price discrimination must rest; (2) the case for regulatory intervention with respect to what would otherwise be unbridled exploitation of differential demand elasticities must also
rest, and (3) the case for expecting competition to operate as an economically acceptable substitute for regulated monopoly is weakest.

From the Telephone Technology Standpoint

The underlying theme of Garfinkel's paper seems to be the increased need for disaggregation, or unbundling, in an increasingly competitive supply environment with respect both to long distance telephonic communications and users' installations. Quite apart from the relative economic positions of the industries, it is obvious that communications companies have more opportunities for unbundling than railroads have been able to find. Also, apart from technical considerations, present control of a high percentage of the U.S. telephone industry by one company makes further progress in the direction of separations and segregation easier, for institutional reasons, in the telephone case. Fortunately, a bystander need not worry about the Pandora's box that may be opened as unbundling proceeds. As a bystander, I would like to offer several comments.

First, a very useful preliminary to disaggregation of rates is the disaggregation of thoughts. This is not a plea for schizophrenia. Instead, it is an attempt to underscore a point which emerges from time to time in Garfinkel's paper: Recent developments, especially greater permissiveness with respect to the introduction of competition, compel greater attention both to the structure of costs and the structure of what is being demanded. Here, again, telephone companies appear fortunate in comparison to railroads. So far, the new competition is mainly based on the same technical principles and presumably reflects more or less the same cost structure and relative appeal to different types of consumers as do the innovations being brought to the public by established companies. Therefore, these companies have not yet had to face the double railroad problem of severe new competition from a different mode with both a different cost structure and different bases for appealing to customer demand.

Second, an economist might try to reassemble the different parts which are torn asunder for examination in Garfinkel's paper along these lines.

Interior wiring is and probably must continue to be a matter of customer tailoring. Any implicit rate subsidy in aid of interior wiring therefore has an open end with respect to the trend of future hourly labor costs, just as it fortunately has a closed end, or nearly so, with respect to single lines for existing businesses and residences. These brief comments bring up two sides of an issue which should be of concern to both regulators and regulated: Regardless of the present size of an internal subsidy, how large can it be expected to become, and how soon?

Proceeding from the extreme of the retail case, with interior wiring, to the extreme of the wholesale case, with mass transmission of messages between population centers, one encounters a clear form of the problem of differential density (marginal cost below average cost to a greater or lesser degree). In the railroad industry, extreme forms of the low-density problem have already been recognized as the root of economic difficulties with branch lines. But this kind of approach to a solution was reached only after the possibilities of cross-subsidization at the expense of security holders had been thoroughly exploited. A major difficulty here is political: Who is going to raise rates against faraway places, with strange-sounding names? But one economic point should be stressed at the outset: Geographical averaging, in the form of mixing high and low densities and awarding both the same rate, is inapplicable not only with the introduction of meaningful competition, but also with the introduction of meaningful economic analysis. The term marginal is customarily employed with reference to units of output of a good or service (although this usage is by no means self-explanatory when applied to telephones). But marginal can also refer, in effect, to where the service is used.

We are drifting into a microeconomic, industry-by-industry case for equalization funds and government subsidization — for those who live in less densely populated areas and for those who live in more densely populated areas. There must be some unfortunate souls in the United States who live in towns too small for mass transportation and too large to produce high cost telephone service. Why these individuals should be victims of geographical blight is a question no one public utility management could hope to answer.

Turning to the paper by Andrew Margeson, I found that it increased my knowledge in a number of ways, which I will attempt to specify briefly. It also may have made a contribution to the case of the existence of local independent telephone companies amid the ocean of Ma Bell and the Great Lakes of General Telephone, Continental, and so on. When the issue is relative profitability of long lines versus local service, it may be difficult for a regulatory body to appreciate the economic significance of taking from one pocket to put into another. When Peter and Paul are unrelated except to the extent required to operate long lines services and divide the revenues therefrom, it is in the interest of both to analyze the underlying economic issues as carefully as possible.
Beyond this institutional or ex officio aspect, I found Margeson’s paper interesting in a number of ways — so many, in fact, that I am forced to use what he has to say as a point of departure without doing full justice to the text itself.

First, Margeson provides specific description of the general decreasing cost dilemma in the form of what he calls “nontraffic-sensitive access costs.” I completely agree with his statement that these are most definitely not to be confused with “joint costs.” What I would appreciate, again as the proverbial innocent bystander, is some further idea as to whether nontraffic-sensitive access costs include all access costs on subscribers’ premises, or only some portion thereof. When it comes to divisions of revenue between Rochester Telephone and AT&T Long Lines, I would suppose that the Rochester case would be built on the basis of all access costs. But if the issue involves prospective growth in long lines revenues, in the face of both mounting competition and of new technologies for duplicating the function of Long Lines without the lines themselves, then I can understand why a company in Rochester’s situation would be interested not only in the size of its present slice of pie, but also in the principles of pie-slicing as they will be applied in the future. Finally, to ruin the metaphor, there is the question as to the extent to which different principles of pie-slicing may affect the total size of future pies. If I may revert to plain English: Both Rochester and AT&T must now develop an interest in the other’s rate structure, which was unnecessary in the good, old, relaxed, monopoly days.

I must express my gratitude to Margeson for providing me with a handy reference list of the possible options that might emerge from “unrestrained competition.” I note that two of his four effective options involve explicit redirection of revenues — as a result of the taxation and expenditure process, for option two, or as a result of pooling and redistribution, for option four. Perhaps it is fair to add that both of these options not only are possible results of unrestricted competition, but also could be possible causes of unrestricted competition in a new dimension. Although the new kind of competition which might be produced by these two options might have much in common with the haggling and bargaining which now goes on in connection with segregation and separations, it could also carry intercompany relationships even closer than they are now to the edge of interneceine warfare. At first glance, option three looks superior: “Eliminate the contribution process and allow local telephone companies to charge for access to the local telephone network when used as part of a through toll service.” At second glance, this option seems to plunge matters further into the interest in each other’s rate structure which has already been mentioned. The same telephone and the same interior wiring are used for both local and long distance calls. If the charge for the former is made entirely on a subscription basis, and for the latter on a toll-per-unit-of-use basis, then the attractive rule that every tub should stand on its own bottom may be vitiated, in the face of future competitive pressures, if one of the tubs turns out to be a tumbler. Reliance on charging for message units even for local calls was an early feature of the industry in at least some localities. Repugnant though the whole idea may be to devotees of short-run analysis, who are inclined to stress high fixed costs and low marginal costs, there may be a strong pragmatic case in an economic as well as a political sense for relying less on a fixed monthly charge for local service, and more on toll charges levied on a per-call basis, or even (and perhaps preferably) on a total holding time basis — during business hours, at least.

This brings me to what I like best about wandering in foreign fields: the encounter with what appears to be a staggering fact but which, for all practical purposes, be a mirage. My Staggering Fact, obtained from Margeson’s paper, is that “Rochester’s average access line . . . is in use just forty-six minutes a day.” With my head full of Hopkinson two-part rates for electricity, plus my memory of the magic wand of Alfred Kahn at the Civil Aeronautics Board, I say: This looks like a ridiculous load factor. What can be done to improve it, for the benefit of all? Having asked myself this, I can perceive one obvious general answer, and another answer which may be specific to the telephone industry.

The general answer, which is familiar to costing experts in the electrical industry, to statisticians, and I am sure to virtually everyone in the telephone business, is: “How sweet are the uses of diversity.” Moreover, in the absence of any knowledge on my part of the local rate structure in Rochester, I can testify that toll-free local calls in Amherst, Massachusetts, do not cause us to use our home telephone all the time. In fact, I cannot imagine how much I would have to be paid to interest me in improving my local call load factor.

The specific answer seems to me more interesting. It is this: Telephone service, whether local or long distance, is two-way service by its very nature. Thus, the word access is not adequate to describe a situation which involves both the outgoing, or exit, and (to borrow a mining term) the incoming, or adit. For anyone frustrated by a busy signal while trying to put through a call, exit, or outward access, might better be termed “excess.” The type of congestion theory which has been developed in the field of transportation should find a useful applica-
tion here: The net costs of providing a telephone installation which is, in fact, being used consist of the objective costs (the telephone set, interior wiring, and so forth) which can be charged directly to the subscriber, plus the possible subjective costs to others who are unable to obtain access to the subscriber because his telephone is in use. Recasting local telephone rates on the basis indicated by this argument not only might have several consequences, but also, and perhaps more important, such recasting might produce a more rational justification for what is going to happen anyway. The following are two obvious examples: The unlisted number must be the source of many fruitless calls for directory assistance, and businesses presumably have much more interest in adit relative to exit than do most residential subscribers. To charge someone for directory assistance when he cannot obtain assistance, because the number is unlisted, scarcely seems appropriate; instead, anyone with an unlisted number should be made to pay a nomadic fee to cover the subjective costs to others of telephone calls forgone. As for businesses, one would suppose that, in Rochester, their telephones are in use for more than forty-seven minutes a day; therefore, a local toll system might help to capture the extra revenue which is presumably available from them. But, in principle, there would seem to be nothing wrong with charging businesses both for exit and (doubtless at a lower rate) for adit.

Comments

William H. Melody

Interest in network access pricing has been spawned by the opening and expansion of competitive opportunities in the supply of interexchange (intercity) telecommunication services. Since an essential component of these services is the local connection supplied by monopoly telephone companies, the level and structure of charges paid by new competitors for access to local exchanges will be a principal factor in determining the extent of competition that develops.

Lawrence Garfinkel describes how competition erodes the "contribution" made by the basic long distance message telephone and vertical services to the coverage of joint and common costs, requiring local rate increases and the introduction of new telephone company charges. Andrew Margeoson points out that competition may require that some social goals reflected in current ratemaking be sacrificed, that cost separations and revenue settlement methods be modified, and that independent telephone companies may not fare well as a result of the ENFIA negotiated revenue settlement between the specialized common carriers and AT&T. Robert Willig examines the theoretical implications of network externalities within a framework of static welfare theory, concluding that if a beneficent philosopher king were to set prices for all services in terms of the inverse elasticity
rule (called Ramsey pricing), some network access prices could be below marginal product costs. He then concludes that because the prices charged by telephone companies for existing services "generally reflect contributions to the costs of network access . . . , the prices charged potential entrants for technical network services must reflect the same contributions."

The analytical framework common to these three papers is striking. Each implicitly assumes that the "network" is a fixed resource that presently is being provided most efficiently and is achieving policy objectives optimally. In static analysis, with these assumptions, competition can provide no real benefit, but it can create serious problems as the system attempts to adjust to the new competitive circumstances and still achieve the same efficiency and social objectives. In reality, the network is not a fixed resource, but a rapidly growing one; it is not provided at maximum efficiency; it is not achieving policy objectives optimally. Moreover, the primary benefit of competition is not static efficiency. It is efficiency in the dynamic process of adjustment to new technology, new production methods, new demands, new services, and new ideas. This has been documented on many occasions, perhaps most thoroughly in FCC Docket 20003.1 To constrain one's analysis to considering only assumed short-run static implications is to miss the essence of the issue.

Perhaps the most significant problem that must be addressed in any examination of the development of network access pricing for competitive suppliers is that of monopoly power. For the vast majority of the nation, the major supplier of the competitive local services (AT&T) has a monopoly on the local exchanges to which the new competitors must have reasonable access. AT&T has a record of seeking to block, resist, or destroy competition by all possible means. There is no reason to believe that behavior will change. Indeed, AT&T's anticompetitive behavior simply reflects its vested economic interest. Yet, Garfinkel and Margeson view the network and local exchange monopolist as benign; Willig's characterization is "beneficent."

Garfinkel recites the now familiar, but thoroughly discredited, AT&T horror story: Local residence service costs $16.15 per main station, but realizes only $3.00 in revenue, benefiting from a $7.15 contribution from interchange services that are now being opened to competition.3 A parallel study treating local service as the incremental service would show it is making a substantial contribution to the interchange services. All this means is that there is a high proportion of joint and common costs in the system. As a result of competition, these costs will grow, and the new competitors will provide reve-

ues that also contribute to the coverage of joint and common costs. The dynamic effects of competition are more likely to reduce local service costs than to increase them. Regulatory policy can ensure that it does.

If competition forces telephone companies to take a closer look at their detailed cost characteristics and to propose price adjustments to reflect costs, it can only mean a move toward increased efficiency. This is a change that many regulators have been attempting to accomplish over the last decade, with only a few delayed successes.8 But a simple telephone company announcement that directory assistance charges and measured service rates are being proposed is no reason for regulatory agencies not to require proper cost justification and evaluate the implications in terms of discrimination, equal access to information, and social criteria. Finally, rules for expensing or capitalizing inside wiring costs, or any others, should be based on standards of proper accounting, not on the need for additional revenues or the desire to facilitate pricing responses to competition.

Margeson's critique of traditional cost separations and revenue settlements is well taken. The process is designed to achieve only an aggregate sharing of total costs and total revenues among telephone companies and regulatory jurisdictions.4 Historically, the detailed methodology has been manipulated beyond belief to achieve that objective. The underlying calculation formulas make no sense in terms of any other criteria. Current separations and settlements procedures are a method of sharing costs and revenues in an environment almost absolute power in the manager of the network revenue pool, AT&T. The procedures are certain incompatible with competition. If effective competition is to develop, a system of network access prices will have to replace separations and settlements.

Margeson's suggestion of adopting a single multipart tariff for local network access for all users is sound in principle. But his proposal to freeze the nontraffic sensitive plant allocation factors is not. Intercity competition can take place at any level of allocation of local exchange costs to interexchange services. All competitors must access the local exchange. The important issue for competition is that the network access prices be applied equally to all. It may well be that a greater allocation of the joint and common local exchange costs to intercity services will be necessary if local exchanges are to have sufficient incentive to upgrade local plant in order to accommodate the new and more sophisticated intercity services.

The ENFIA settlement arrangement is an expedient compromise that does little to facilitate the development of competition or the
change to network access pricing. That it is necessary is an indictment of the ability of the regulatory process to deal with such an important issue in a timely fashion. ENFIA provides an interim solution in a process that ultimately is part of the larger agenda of solutions. When considered against the backdrop of recent FCC regulatory history, it simply indicates that the FCC has all but abandoned its role as independent public policy maker on the issue in favor of the role of mediator. The specialized carriers must now negotiate their way into the industry. Given the enormous disparity in economic power between AT&T and the specialized carriers, one can hardly expect a balanced result in the settlement. That the specialized common carriers (SCCs) have agreed to the ENFIA settlement is evidence only that the alternatives were worse, not that the result was fair or in the public interest. It also increases the risk that they will soon lose their independence and be forced to settle into a cartelize market in a niche similar to that of independent telephone companies.

Revenue settlements among carriers traditionally have applied to the basic MTS and WATS network services. AT&T's private-line services, including TELPAK, FX, CCSA, and other potentially competitive specialized services, have been exempted. Some of these provide a greatly reduced revenue contribution for local exchange access; some make no contribution. As a result, these specialized services have been provided at a discriminatory competitive advantage over the MTS and WATS network services. AT&T has been cream-skimming its own basic services.

The SCCs compete most directly with AT&T's specialized services for business communication. Their revenues are a small fraction of AT&T revenues for these services. AT&T and the other telephone companies are concerned about the fact that the SCCs might drain toll service revenues from MTS and WATS. The ENFIA settlement is the result. Incredibly, that settlement does not apply to FX, CCSA, facsimile, or value-added services. The SCCs are restrained from skimming the cream off the MTS and WATS services, but AT&T may continue. It appears that the FCC's competition policy is going timid in the implementation phase.

Competition does not mean that the traditional social goals incorporated into telecommunications costing and pricing must be for-gone. Certain methods of implementation of social policy may not be compatible with active competition, for example, the current separations and settlements process, but this merely requires that the method of implementation be changed, not that the social goal be abandoned. In virtually all cases in which changes in the method of implementation of social policy are required by competition, it will mean that the social goals must be specified and implemented more precisely and directly, rather than simply leaving them in the hands of the industry. Fair competition simply requires that the costs of subsidies be imposed uniformly upon all competitors. Here, also, changes of this kind would seem to be beneficial, whether or not competition requires them.

Willig provides an impressive tour de force of mathematical analysis demonstrating that the obvious can be proved in 80 equations as well as 8, and that elaborate analytical structures can be built on quacksand simply by assuming away the problem. He observes that if network externalities are positive, the social marginal costs of network services by new firms will be less than the marginal production costs of supplying the local facilities. Therefore, pricing at social marginal cost would yield financial losses to the local network. If the inverse elasticity rule is applied, a system of "optimal" network access price discrimination could be established by the "social decision maker" that would yield the desired level of total revenue. All prices would be above social marginal cost, but some prices may still be less than the marginal production cost.

But, Willig argues, new competitors should not be encouraged to enter the market unless they pay an access price that covers the marginal production cost. Optimization under this constraint will require more refined price discrimination between final consumers and competing vendors. Final consumers can be charged prices below marginal production cost, but competing vendors cannot. A new entrant should pay a network access price that will cover the marginal production cost of providing the local facilities plus the net revenues (revenues minus incremental costs) of the network services lost by the established company. Lessons from the exercise include: Complex payment mechanisms may be socially desirable if maximum price discrimination is infeasible; responsive flexibility in network access pricing is required to preserve efficiency.

Willig briefly mentions the advantages of separating the local exchange carrier from the competitive intercity carriers, but quickly backs off because of the possibility of "economies of scope" between local and intercity services. He apparently has forgotten that network economies can be realized, as they are now, by the industry without requiring singular ownership.

AT&T would love Willig's system. AT&T is the beneficent social decision maker responsible for implementing public interest network
access pricing. That pricing may require setting some final consumer prices below marginal cost and must require new entrants to pay access prices that will compensate AT&T fully for lost revenues. Maximum price discrimination is necessary to maximize consumer welfare, and maximum pricing flexibility is also required. AT&T would decide the level of competition that it judged to be efficient. Under these arrangements, there would be no competition.

It would seem that the obvious solution implied by Willig’s analysis is that the local and intercity carriers should be separated and all intercity carriers charged the same marginal cost. This would maximize consumer welfare. However, it would do so only because the analysis fails to recognize the trade-off in externality effects. The addition of intercity network services will provide externality benefits to the total system, but the maintenance of low prices for basic local exchange service attracts customers who would not otherwise be on the network, thereby providing externality benefits also. Because of the joint and common cost characteristics of local facility supply, the pricing policies of local service and network access are interdependent. Once prices in one area are set, the prices in the other area must be established to cover the residential costs. Thus, a trade-off must be made between the externality benefits of low local service prices and low network access prices. If the former are more significant, the externality problem addressed by Willig does not arise as long as carriers must set prices to cover total costs. However, since externality benefits are in both the economy that intercity network access services, the Willig analysis does provide justification for pricing below marginal production costs in both areas to achieve optimum social efficiency, and for the adoption of internal or external subsidies to finance it.

In reality, of course, neither AT&T nor other carriers set their prices on the basis of social marginal cost, marginal production cost, or the inverse elasticity rule. Nor should they. The entire analysis is based upon the assumptions of static optimization of a fixed resource by a benevolent social decision maker (that is, monopolist) with foreknowledge. There is no way competition can improve the situation — by definition of the problem. It is indeed ironic that an analysis purportedly addressed to economic efficiency and consumer welfare in a system of competitive private markets is premised on an assumption that the incentive of the private monopolist is not self-interest, but the optimization of social welfare as measured in terms of Pareto optimality and Ramsey prices.

But we cannot blame all this on Willig. He is simply building an elegant extension to neoclassical welfare theory, as refined by W. J. Baumol and others.² If one wishes to optimize under a system of Ramsey prices, it will be done by establishing an absolute monopoly and by creating conditions that will permit the maximum possible barriers to resale and price discrimination — to be benefitted applied, of course. But in competitive market theory, which states that firms act in their own economic interests, the power to discriminate on the basis of differential price elasticities of demand is the very essence of monopoly pricing. It is inefficient. In competitive markets, optimal efficiency requires that all price discrimination be precluded by competitive substitutes and the maximization of opportunities for resale. Therefore, a specific objective of regulation, in reality, is to prevent monopoly price discrimination, not to encourage it.

By simply changing a tacit assumption, from the monopolist acts in its own best interest, to the monopolist is beneficent, an entirely new area of analytical game playing is opened up, with ample opportunity for theoretical technicians to demonstrate their skills. As the recent economic literature illustrates, the Princeton crowd is strutting its stuff. Unfortunately, it so flies in the face of reality as to be rendered totally irrelevant. AT&T may be beneficent to Princeton, but in the marketplace it has the same incentives as the other firms.

Obviously, if competition is to be efficient, discrimination in network access pricing will have to be prevented. According to the criterion of Pareto optimality in welfare theory, these prices should be set equal to marginal cost. But marginal cost prices will only be optimal if all prices in the economy are simultaneously set equal to their respective marginal costs. This cannot be done even for other telecommunications services offered by AT&T. Thus, since total cost must be recovered, we are once again back to average cost pricing as the most efficient standard.

Finally, it must be observed that the criterion of optimality in welfare theory has nothing to do with reality. Optimality is achieved when the firm does its best from among a set of known plausible alternatives by comparing the myriad of possible marginal costs. If the decision turns out to be a disaster in reality, no matter. The decision was still optimal. That is another important reason why neoclassical welfare theory and marginal cost are simply irrelevant for problems concerned about efficiency in reality.

Notes
1. See FGC, Report in Docket 20003.
2. Ibid.
3. For example, the FCC spent more than a decade resolving the principles to
be employed to measure service-by-service costs for AT&T in Docket 18128. It has yet to specify, or approve, a methodology for implementation.


Part Four

Costing Methodologies for Time Differentiated Rates
Time-of-Use Rates: Marginal Costing and the System Planner

Leo T. Mahoney, Jr.

My purpose is threefold: first, to outline the rationale for using the system planner's work in deriving time-differentiated marginal costs (TDMC) for ratemaking; second, to assess briefly where we now stand regarding the great rate debate; and third, to discuss what has been done to implement marginal cost principles in ratemaking.

The rationale for using the system planner's work product in deriving TDMC for use in ratemaking rests on the assumption that the complexities of the real world make a simple specification of the technological opportunities available to all firms virtually impossible. It recognizes that considerations of life-cycle costs, lumpy investments, maintenance, and uncertainty prevent a simple static model from reflecting all of the nuances of a firm's investment and production decisions. It also recognizes that there are likely to be important differences among firms which cannot readily be distinguished by a generalized production function and a simple static optimization program. What is needed is a dynamic optimization specific to each company, and this is precisely what system planners and engineers do. It is reasonable to assume that they are better able to provide the optimiza-
tion results than is the economic analyst using a general econometric model and a static optimization program.

An important feature of the system planning process is that the existing system is taken into account. Although economic theory states that sunk costs are sunk and that decisions on the margin should reflect only "avoidable" or opportunity costs, this does not necessarily imply that the existing system should be entirely ignored. The capital investment in plant and equipment is indeed a sunk cost, but the equipment is already in place and often will have a long useful lifetime. Thus, these plants continue to have opportunity costs associated with them, since the firm can still decide whether or not, and how much, to use them. Any sensible program to provide electricity at minimum cost must consider the value or opportunity cost of these plants when making supply decisions. These supply alternatives, although unique to a particular firm, must be reflected in supply decisions, the costs of production, and the prices charged for service.

Making appropriate use of the existing system also has attractive properties in terms of efficient pricing. Any utility system is likely to be "nonoptimal" at a given moment if viewed from hindsight. Efficient prices may require a departure from strict long-run marginal cost prices in favor of a pricing system that reflects short-run marginal costs. Data about the operation of the existing system and how it will evolve over time are critical if rates based more closely on short-run marginal cost considerations are to be developed.

Since electrical generating equipment is normally built to last 30 to 40 years, investment decisions must be based on the expected "life-cycle" costs of the plant. This means that current prices for fuel and maintenance do not necessarily give a complete picture of the long-run economics of generating alternatives. The prices of some fuels may increase faster than others, and plants built to operate as baseload equipment today may serve as cycling or even peaking equipment sometime in the future. Incorporating these considerations requires a detailed dynamic economic model. Such models are routinely used by system planners in electric utilities and are generally available for analyzing problems of this sort. Modern electric generating facilities must have planned downtime for maintenance and, in the case of nuclear plants, refueling. These requirements play a critical role in the planning of any utility system and in its costs of operation. Our experience indicates that maintenance scheduling is a unique and difficult engineering and economic decision. We have found that the system planner takes these considerations into account in developing the least-cost plan.

Uncertainty about demand and capacity availability is an important reality in the production of electricity. It is for this reason that utilities carry reserve margins. The cost of shortages plays a critical role in determining the optimal amount and mix of capacity and in estimating the relevant costs of production, especially the marginal costs. Those who insist on referring to the use of loss-of-load probabilities and the concept of shortage costs as being ad hoc and inconsistent with neoclassical economic theory are incorrect. French economists recognized the importance of uncertainty and outage costs 20 years ago and began to factor them into their ratemaking for EDF at that time. While there are differences of opinion as to how uncertainty should be factored into the analysis, there are apparently very few who believe it should be ignored completely. Again, we find that the system planner has done the job well and has included considerations of uncertainty in the planning process.

Given all of the foregoing attributes of the system planning process and the fact that the objective function of the planning process is to minimize costs, we have concluded that marginal costs derived from the system plan come closer to the criteria set forth by the economic theoreticians than do those derived in other ways. This is not to say that the system planning process is perfect but, rather, that it is the preferred tool for use in the ratemaking process.

The second purpose of my presentation is a brief assessment of the current situation regarding the great rate debate.

The passage by Congress of the Public Utility Regulatory Policy Act of 1978 (PURPA) appears to me and my colleagues at National Economic Research Associates (NERA) to have brought the great electric utility rate debate to an end. It is our contention that the language of the act itself compels one to conclude that marginal cost data must be gathered and given consideration in the ratemaking process. The act inaugurates, among other things, "a program providing for . . . increased efficiency in the use of facilities and resources by electric utilities." Furthermore, at Section 115 we find:

In prescribing such methods, such State regulatory authority or nonregulated electric utility shall take into account the extent to which total costs to an electric utility are likely to change if —

(A) additional capacity is added to meet peak demand relative to base demand; and

(B) additional kilowatt-hours of electric energy are delivered to electric consumers.

Indeed, the foregoing passages, in language uncluttered by academic jargon, require a movement toward economic efficiency in terms of
resource use and also set forth a simple definition of marginal costs (although the term itself is not used). Even if the act does not emphasize marginal costs, it is our opinion that a carefully reasoned approach to the implementation of the act demands that data on such costs be gathered and given material weight in the ratemaking process.

There does, however, remain some question as to the differences among the various approaches to marginal costing that have been offered during the course of the debate. It is our sense that the differences between the approaches on a conceptual level are not as great as first meet the eye. An examination of several indicates that each of these approaches takes, as its starting point, a conceptual model which attempts to meet a load duration curve at minimum cost, given prevailing technological opportunities. Each takes as its theoretical starting point the choice of a mix of techniques which meets the loads of various durations at minimum cost. The conceptual model underlying all of them essentially involves building and operating a system to meet the load at minimum cost, given current technology and factor prices. The differences that do exist lie primarily in the area of implementation rather than in basic conceptualization. Finally, for the benefit of those who will continue to oppose the use of TDNC in the electric utility ratemaking process on various grounds, such as administrative infeasibility, lack of account precision, consumer comprehension and so forth, I would like to review briefly some real world events regarding utility ratemaking. A number of jurisdictions (New York, Oregon, Wisconsin, California, and Massachusetts) have adopted or are reviewing as a prime consideration in ratemaking. In a number of others (New Jersey, Pennsylvania, Virginia, and New Mexico), rate proposals based on marginal costs have been adopted even though no explicit requirement regarding marginal costs exists. TDNC-based rates have also passed a legal test in New York, where the Court of Appeals has overturned an Appellate Division finding that the Long Island Lighting Company’s TDNC-based rates for its largest commercial customers were discriminatory.

As Exhibit 1 indicates, using recent events in Wisconsin as an illustration, not only have TDNC-based rates been instituted, but also an evolutionary process with regard to rate design has begun. In order to examine this evolutionary process, a comparison of the Madison Gas and Electric Company’s (MG&E) 1976 TDNC-based rates with its 1976 proposal illustrates the effects on large industrial customers. The data are given in Exhibit 2.

The 1976 rate has, for each of two seasons, a peak period demand charge and a seasonally differentiated energy charge for diurnal peak and off-peak periods. The energy charges consist of the system marginal running costs. In contrast, MG&E’s 1978 rate proposal recognizes that local distribution facilities must be sized for the consumer’s maximum demand regardless of the time of occurrence (the 15-minute demand charge); measures on-peak demand charges over a longer integrated period (two hours); and rolls some of the demand-related costs into the energy charge.

The TDMC-based rate proposed by Arkansas Power and Light (AP&L) for residential consumers on a voluntary basis is shown in Exhibit 3. This rate features a simple kilowatt-hour charge differentiated between seasonal and diurnal periods. The customer charge includes a component representing the marginal cost of those local distribution facilities related most closely to the consumer’s own maximum demand.

In Ohio and New Mexico, TDMC-based rates closely mimicking the French residential tariff have been filed. These are shown in Exhibit 4. The tariff filed by Dayton Power and Light Company (DP&L) levies the local facilities charge on the basis of kilowatt-hour consumption, while the tariff filed by Public Service Company of New Mexico levies this charge on either connected load or service entrance size.

For those concerned about consumer comprehension, an examina-
tion of what Virginia Electric and Power Company (VePco) is doing may be instructive. In June 1977, VePco submitted a two-part proposal concerning implementation of TDNC-based rates. The first part recommended mandatory implementation for all customers with a peak summer month usage in excess of 3,500 kilowatt-hours. Details of the program included: (1) a consumer education program; (2) installation of time-of-use (TOU) meters and issuance of comparative bills for one year (consumer pays traditional rate during trial period); (3) after one year, comparative billing would cease and customers would be charged according to TOU rates; and (4) also after one year, an evaluation of the costs and benefits of the program would be made and recommendations formulated as to whether or not the program should be made mandatory for small consumers. Mandatory implementation for large customers was selected because it is the only way to gather statistically valid information to assess the costs and benefits of TOU rates and determine to what size customer the rates should be extended.

The second part of the proposal suggested voluntary implementation for all customers with peak summer month consumption between 720 and 3,500 kilowatt-hours. VePco proposed to: (1) solicit a maximum of 1,000 volunteers from within the kilowatt-hour usage zone; (2) set up a consumer education program; (3) install TOU meters and issue bills based on the TOU rate schedule; and (4) collect and analyze data concerning the consumption patterns of volunteers.

EXHIBIT 5. Sample TDNC Rates Designed with Emphasis on Consumer Comprehension, Arkansas Power and Light Company, Filed October 1978

<table>
<thead>
<tr>
<th>Customer charge</th>
<th>$9.10 per month</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kilowatt-hour charges (€/kWh)</td>
<td></td>
</tr>
<tr>
<td>On-peak period</td>
<td>Off-peak period</td>
</tr>
<tr>
<td>Summer</td>
<td>Winter</td>
</tr>
<tr>
<td>6.42</td>
<td>2.20</td>
</tr>
<tr>
<td>Peak period: summer, Monday—Friday, 11:00 A.M.—9:00 P.M.; winter, Monday—Friday, 7:00 A.M.—8:00 P.M.</td>
<td></td>
</tr>
</tbody>
</table>

Finally, NERA conducted a survey of state regulatory commissions concerning the various costing methodologies sanctioned for use by the states' public utilities. The results are shown in Exhibit 5. Significant findings were:

1. State commissions in 11 states have prescribed or specifically approved one or more methods for determining costs associated with
### EXHIBIT 5. Costing

<table>
<thead>
<tr>
<th>Number of states</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commission specifies method(s) for determining cost of service by class</td>
<td>11</td>
<td>38</td>
</tr>
<tr>
<td>Required method measures TOD, seasonal cost differences</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>Required method measures marginal or incremental costs</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Utilities measure cost of serving each customer class</td>
<td>40</td>
<td>9</td>
</tr>
<tr>
<td>Utilities measure TOD, seasonal cost differences for each class</td>
<td>22</td>
<td>18</td>
</tr>
<tr>
<td>Utilities measure marginal or incremental costs for each class</td>
<td>14</td>
<td>25</td>
</tr>
<tr>
<td>Commission requires cost-based rate structures</td>
<td>27</td>
<td>21</td>
</tr>
</tbody>
</table>

providing electric service to classes of electric consumers.

2. In 40 states, utilities do measure the costs of providing service to each class of customers for rate cases or other purposes.

3. In 22 states, utilities also measure differences by TOU and/or seasons.

4. Commissions in 27 states require that the structure or design of rates for each class of customers be based on the cost of providing service to each class.

In conclusion, I believe that we have come a long way since the beginning of the electric utility rate study. The techniques used have been sharpened, and time-differentiated pricing of electricity services is on the increase. The passage of the Public Utility Regulatory Policy Act of 1978 is now history. It is time to call a halt to the at times acrimonious debate and begin implementing the act’s ratemaking provisions.

**Notes**

2. Ibid., p. 10.
Windmills of Costing Methods — 3; Poor Don (Rate Managers) — 0; Sancho Hasn’t Helped Much

William J. Leininger

The title I have selected is intended to raise a wider range of issues than just the contestings among the participants to the great rate debate. At a minimum, the remarks in this forum are a continuation and elaboration of the different views expressed at the Montreal Incremental Cost meeting in 1978, the Southeast Electric Exchange in 1978, in the recent Texas generic hearing, and, finally, in the comparison and critique of those of us who have been labeled marginalists by Temple, Barker, and Sloan in their EPRI-sponsored review.

I will discuss four topics. The first outlines my Don Quixote perception of the great rate debate. The second summarizes some overriding concerns that must be dealt with almost immediately if the utility industry is to survive as an industry rendering efficient utility service that is competently and effectively regulated at the state level, where the most valid and complete understanding of the problem is found.

My third topic outlines Ernst & Ernst’s position regarding the inherent ability of any costing method to support time-of-use pricing. It also outlines the main elements of our approach and briefly comments on the main criticisms it has received. The fourth topic presents some action steps that we think would make Sancho a little more helpful.

The Windmills of Ratemaking

At this point in the great rate debate, it appears that whatever value marginal cost/time of use had at the outset, the debate has become like Don Quixote’s quest. The new ratemaking concepts that have appeared are like the windmills, and they are activated by the winds of change. It also seems that rate managers and regulators have been the unlucky, and in some instances unwilling, challengers of these windmills. The Sanchos of the world (consultants, zealots, and even some parts of some regulatory bodies) seem to have been mainly responsible for creating the many-armed windmills, for sending the companies and the regulators into the churning blades, and for then handing them another lance (in the form of a new idea) and putting them back on their horses to charge again.

Thus, it appears to be a no-win situation. The windmill is continuously equipped with repaired old blades and more new blades (call them refinements of the old-new rate philosophies and rate structures, and totally new concepts). And hurricane force wind power in each rate hearing is supplied by the various parties who are now able and willing to intervene. On the other side, the challenger of the windmill is provided by his Sancho with lots of enthusiastic support (both of the old school — what do all these new people know, we will get them with the old tried and true methods — and the new school — let’s attack with this concept, or let’s try out that idea); the wrong weapons (an historical cost-of-service study when the issue is marginal costs, and vice versa); and weak or incomplete weapons (arguments about the new concepts that are inconsistent or clearly ad hoc). After putting the company or the commission on its rate case steed, Sancho stands aside and observes. The game score is at least Windmills — 3, Rate Managers and Commissions — 0.

Assuming this is a reasonably realistic picture of the situation today, some serious concerns are raised.

Overriding Concerns

While the quest remained at the state level, and then only in a few states at one time, it did not, in my opinion, represent a major threat to the survival of the industry. It was painful, expensive, and time com-
suming to companies, sometimes to commissions, and where actually implemented, it was many times painful to rate payers. But at the state level, the quest remained under the control of the participants; it was always possible to dismantle the windmill, as has been done in some states, and everyone was learning about the game.

In our opinion, however, the game has now changed dramatically. With the National Energy Act, we have come upon an enormous windmill with many more blades (federally constructed with a potentially inexhaustible variety of approaches to time-of-use pricing) that will be driven by the biggest hurricane ever seen (federal resources funding federal intervention, backed by congressional interest in moving the regulation from the state to the federal level).

At the NARUC meeting in Las Vegas, Nevada, the message of John Dingell and Dave Bardin was very clear. It is the federal belief, as I understood it, that the following are the new rules of the game: (1) Rate structures are to be slanted to accomplish the new national energy policy (however defined); (2) the policy goals are efficiency, equity, and conservation (however defined, and in an unknown order if they conflict); (3) there are eleven specific guidelines to assist state regulation to support national policy; (4) the ERA will have as one of its operating objectives producing consistent implementation of the guidelines; (5) conservation apparently means a conscious effort to shift to natural and other resources and away from oil; and (6) if the states do not conform, new laws will be drafted to ensure these rules are carried out at the federal level.

In this game the field of action is now nationwide, with favoritism toward time-of-use pricing and toward some concept called "marginal" or "incremental" cost. It seems that the industry is about to lose the luxury of being able to experiment and learn in the quest. The industry has come upon a windmill that must be mastered, else it can, and perhaps will (in my opinion), administer a beating from which the industry may be very long in recovering.

Continuing my analogy, it is time for Don Quixote rate managers and commissions to tell their Sanchos to do their job and provide the tools to deal with this new challenge. First, in the electric area, the tools must be fully explained so they can be understood and used properly; in the gas area we must stop calling the tools by the wrong name. Otherwise, in both areas jousters will be undone when they use the new ideas, concepts, approaches, and so forth, as the name implies. In the gas area, for example, I am referring to the present discussion of marginal or incremental costing in the industry which immediately calls to mind the classical definition of the change in total cost from a one-unit change in output. As I understand the term as it is now being used, marginal or incremental costing does not refer to a change in cost due to a one-unit change in output, but to the concept of pricing that charges specific customers or customer classes at all times the cost of the highest priced gas, regardless of whether they are responsible for that gas (for example, interruptible industrials). This seems to me to be penalty or subsidy pricing. The electric analogy would seem to be to charge certain customers or classes with the highest cost electricity, no matter when they take service.

In summary, I believe the time for questing is over since the potential for nationally mandated and regulated rate structures is now very real. I believe such action could well result in at least higher costs and might lead to other significant changes that would be undesirable from the rate payers’ and investors’ viewpoints. Consequently, resolution of the great rate debate has now become, in my mind, a major concern that must be settled immediately.

Ernst & Ernst’s Position

Our position in the great rate debate is very straightforward. We do not believe that there is an “economic welfare theory” argument that shows that one costing method is superior to another. Note that I am not saying that we cannot produce cost estimates, and resulting rates, that are fully consistent with engineering or economic theory. I think this is possible.

We do not believe that either approach can be selected on a theoretical basis for the following reasons. First, as I understand it, the declining block rate forms and the newer historical cost-based time-of-use rates are justified as reasonable approximations to engineering reality, not as springing from any engineering theory. Second, the new economic approach to time-of-use ratemaking based on the marginal concept also faces difficulties. There are significant theoretical problems in the form of the second-best arguments, which reduce theoretical support to an industry level. Even within the industry, it is not possible to use the theoretically consistent rates. What can be done is to preserve those theoretical features that are most desirable. In particular, we think it is important to preserve the peak/off-peak energy differential which leads to proper consumer responses through elasticity/cross-elasticity values. What we have seen is that applying different approaches to the same company often yields answers that can be, and have been, quite different.

Now, let us discuss whether the economic approach is supported. If there is a supporting argument, we think it comes from the fundamental purpose of regulation. Our view of that purpose is that the regu-
latory body should create a situation which enables the utility’s financial results to look as much as possible like what would happen in a competitive market. If this view is correct, then there is a value in deriving prices on the basis of the economists’ marginal and average cost so that we can see how these rate structures compare to rate structures derived from other costs bases (for example, those of the engineer, the system planner, or the regulator).

Starting from the position I have just stated, we arrive at several conclusions. First, there is no “right” method for developing the cost basis for time-of-use rates. It seems to us that the selection of a preferred method becomes more a function of adequately meeting the generally accepted rate objectives and the specific utility situation (for example, a situation of excess capacity). Consequently, the economists’ method is chosen as the way to estimate the costs to be used for determining time-of-use rates, then one can choose from among the economists’ average cost, short-run marginal cost, or long-run marginal cost. The discipline, of course, offers some guidance in choosing, depending on the kind of market situation one thinks one wants (that is, if one wants regulation to approximate the competitive marketplace). Third, if the economics approach, or that of any other discipline, is chosen, and it is claimed that the approach is valuable because one gets a number of benefits as shown by the theory of the discipline, then the approach must stay true to the theory to claim those benefits. Fourth, if the economists’ marginal approach is selected as the basis for time-of-use pricing, then we favor the short-run marginal cost basis over the long run because we do not have the world of perfect competition to reflect the trade-off between heat rate and capital costs. Hence, the utility and the rate payer are the same (that is, revenue consequences to the utility are the same as the cost consequences to customers); and income stability is preserved. This preference may be tempered by the capacity position of the utility.

Given our position and our conclusions, what is our approach, and what are some of the comments that have been made about it? As has everyone else in this field, we have progressed through a number of steps. We began, as most researchers do, with a simplified approximation to electric generation, assuming that the system uses a single fuel conversion technology. We subsequently modified the analysis for generation to include consideration of multiple fuel conversion technologies, maximum/minimum unit sizes, specific cycling capabilities, and heat rate—capital cost trade-offs. The system planners with whom we have worked indicated this was a proper and meaningful range of real world considerations. For analysis of transmission and distribution, we have used a relatively simple approach similar to that of other practitioners. Similar to theirs, our marginal cost estimate is an aggregate composed of the generation, transmission, distribution, and “other” marginal costs.

Turning to specifics, as one might expect, we have developed the economists’ approach in a way we think very closely corresponds to the theoretical requirements. Therefore, we think that if one wants to use the economists’ approach in estimating cost for time-of-use rates, ours is the correct way to proceed.

The fundamental elements and analytic steps comprising our long-run marginal cost estimation method are based on two simple economic definitions: (1) Marginal cost is the change in total cost due to a one-unit change in output from the production process. The long-run marginal cost approach assumes all inputs (plant, fuel, and labor) can be changed, while the short-run approach assumes only some inputs (perhaps fuel and labor) can be changed; (2) the long run is that period of time within which all inputs can be changed, and changed so that they produce the output at least cost.

We have developed our analysis so that these two definitions have been almost completely preserved. The main elements of our approach are noted below:

1. Examine multiple fuel conversion technologies available to the utility, including coal-steam, indoor boiler oil-steam, outdoor boiler oil-steam, combustion turbines, diesel, and boiling water reactor nuclear.
2. Utilize the economists’ production function concept—the equation that illustrates the relationship between inputs and outputs—such as capacity factor and cost. Subsequently, use an optimization algorithm to reflect the trade-off between different kinds of machines (different fuel conversion technologies) to meet load. (3) Utilize the theorems of economics to arrive at what the optimum system generation configuration would look like for each fuel conversion technology, given certain real world constraints on available technology. (Note that this optimum situation is a fundamental theory requirement in order to arrive at the economists’ comparative statics long-run marginal cost). (4) Determine separate cost equations for each technology based on the best heat rate—capacity size combination using the utility’s own cost (fuel and capital) information and load shape. (5) Analyze the best method for meeting load duration curve based on minimizing total cost, while respecting unit size and cycling constraints. (6) Analyze the specific reserve requirement to account for scheduled and unscheduled outages with effects on both unit size and type. For example, units require
certain maintenance time, but if enough days cannot be found to spare a big machine, a smaller machine should be used. (7) Develop rating periods based explicitly on cost homogeneity, rather than select rating periods based on some other criteria and then compute costs. (8) Adjust rates to authorized revenue requirement so as to maintain peak/off-peak differential so that cross-price elasticity (if it exists) can operate to alter consumption patterns. The adjustment respects the necessity to cover operating costs. (9) Provide an easy ability to examine a range of “what if” questions about changes in technology or capital cost, fuel cost, load duration curve shape, government regulations, and so forth, for a single year or several years. (10) Provide all the elements necessary to produce estimates of short-run marginal cost from the long-run method.

I believe it is apparent that in some areas our approach is different from that of other practitioners. This has led to several kinds of criticisms, and I will briefly summarize the main criticisms we have received and our response to each.

**Difficulty, Cost, and Reproducibility**

There are those who maintain that the Ernst & Ernst method is difficult, costly, and may not be reproducible. The “difficult and costly” misconception apparently arose from the reviewer receiving input from other commentators. We believe that since we use company data and real-world constraints, plus well-defined economic theory, the analysis is not particularly difficult to apply. As best we can tell, our cost for the same company is about the same as others, or perhaps a little lower. With respect to “difficult and reproducible,” we believe that because the theory logic and arithmetic processes have been automated, the method should be simple to perform. Automation should also facilitate replication of results because it locks in the economic theory (which does not change) and the arithmetic computations, which also do not change. The company-specific data and operating conditions do, of course, change with each company.

**A Production Function Approach**

Some critics claim that the use of a production function approach is weak because of data problems and problems with selecting a functional form. The data we use are those available when the system is being designed (ex ante data). We use machine capability data that are not distorted by how the machine has been run (when actual loads differ from design loads). To identify the characteristics of the existing technology, we have collected this information on units across the country. We explicitly consider fuel prices the company must pay, as well as unit availability, in our process.

With respect to choosing the Cobb-Douglas production function form, it permits heat-rate—capacity trade-offs, but imposes constant elasticity of output for each input. We did test other forms, but they performed no better. The test statistics for the Cobb-Douglas met all the usual econometric requirements. Finally, it appears from the data that the portion of the production surface with which we are working for each technology is relatively well behaved, so that no matter what functional form is used, the results will be about the same. We are fully ready to use another functional form when it is shown, rather than postulated, that the new function is superior to the one we are using.

**Capacity and Energy Costs**

Some have maintained that the Ernst & Ernst method does not produce capacity and energy costs. In fact, the analysis does produce costs which include both capacity and energy, and these components can be separately identified. All our results have been provided with the capacity and energy charges separately identified on the basis of the mix of capacity and fuel used to meet the load duration curve.

**Handling Reserves**

It has been claimed that our method for handling reserves, both scheduled and unscheduled, is ad hoc. The criticism apparently arose from our not being explicit enough. Our method is, in fact, quite detailed, explicit, and (we think) complete and accurate.

**The Approach Ignores “Reality”**

If one defines reality as the existing system, then our approach ignores reality. One must ignore the existing system if one uses economists’ long-run analysis. However, even when we wave our magic wand over the plant, we require it to meet the current load, and we constrain the plant configuration to meet all existing or expected legal and regulatory constraints, as well as the engineering constraints applicable to each machine. We always consult the system planner and ask if he could meet the load with this “optimal” system. In every case he has said yes. We have further found a close agreement between system plans and our optimal system, to the degree that the amount of capacity for each generation method is within ±10 percent of the system plan. Also, as I pointed out before, the method uses company load data, company-specific fuel and capital costs, and available engineering
technology. Finally, we do use the existing plant when computing economists’ short-run marginal cost.

In summary, the “errors” attributed to our approach were what we had previously identified as limitations; the “improvements” identified have mostly been based on speculation or postulation, rather than a demonstrated alternative. Finally, we think the theory has not been shown to be incorrect.

How Sancho Can Be More Helpful

I believe the current situation is critical and requires Sancho to take some dramatic action to demonstrate that he should be retained as a servant. The giant windmill of potential federal intervention is entirely too serious to permit us to continue in the old ways. I think certain actions can go far to dispose of the great rate debate. The industry can then move forward to the next stage of careful empirical analysis of whether time-of-use rates will, or should be expected to, do much other than more precisely match in time the cost to serve with the rate paid for service. These actions are: (1) Carefully define each costing method, that is, show how the method derives from engineering or economics; define what each term means; explain exactly what data we use and what analysis and computations are performed; explain what advantages will be gained from using the method and the basis for claiming these advantages; show what kind of results are produced when a system has excess capacity, just the right amount of capacity, or too little capacity. Temple, Barker and Sloan, Inc., has made an excellent start in this kind of comparison. (2) Based on the generally accepted ratemaking criteria, the specific conditions within a state or region, and the clear understanding of each method made possible by the first action, agree on a method, or methods, that is most suitable for use in a particular jurisdiction. (3) Initiate discussions with ERA to work out that they understand individual positions and reasons herefor so that a commonality of interest and understanding can be created that permits continued rational regulation.

Appendix

Description of the Ernst & Ernst Method for Estimating Long-Run Marginal Cost

Three main tasks comprise the Ernst & Ernst method for estimating long-run marginal cost (LRMC). These are presented in the following three sections. Also discussed are the supplemental data request items that we like to have for estimating LRMC and the corresponding rates.

Analysis of the Generation Function

The first step in estimating LRMC is construction of the optimal economic system. From economic theory, we use an important and extremely useful tool: the La Grange method. A La Grange function is used to minimize by mathematical techniques the cost of meeting output (load), given how one can combine inputs to meet the output (production function) and the prices one must pay for the inputs (the costs of capital input and fuel by production type). An intuitive explanation of this type of optimization is that if one must meet a certain load shape with some type of technology, and if one faces a particular array of input prices, there is one set of input combinations which minimizes the cost of production. This we refer to as “within technology optimization.”

This is only the first type of optimization which we use. The second is “between technology optimization.” This latter type of optimization configures these optimal machines into an optimal mix, given a particular load pattern. It further recognizes that the optimal machine to produce loads of long duration may be fossil-fired steam, while short duration loads are best met with low efficiency (measured in heat rate) machines with low capital costs.

This discussion leads directly to the data requirements for the first step of the estimation of long-run marginal costs. This step is initiated to develop the economically optimal long-run system with thermal resources. These data are: (1) the load duration curve (LDC), because this is what the utility optimizes against (plans for), and it must meet this load; (2) the price of fuel in cents per million Btu, because this is an important parameter in deciding not only what technology to buy but also how much to pay for fixed capital stock; (3) the price (levelized fixed charge rate) of capital as a percent, because this, in conjunction with the price of fuel, leads to the optimal system mix. While we prefer this item to be company supplied, we can and, if necessary, do calculate a levelized fixed charge rate of capital. It is calculated as the normalized present value of a revenue requirement stream resulting from a $1,000 investment. For this calculation, we need the book life of the investment, the composite income tax rate, the equity return rate weighted by the percentage of equity in the capital structure, the interest rate weighted by the percentage of debt in the capital structure, the property tax rate, and the Asset Depreciation Rate guideline tax life of the investment. Of course, these items are optional if the levelized fixed charge rate of capital is supplied by the company; (4) the cost of scrubbers on coal units, because this may alter both the optimization within and between technologies and in turn the optimal system. We
prefer this cost to be expressed as a percentage increase over the base cost of an unscrubbed coal unit.

Hydraulic, geothermal, and/or solar power, where feasible, are now considered available long-run options in the optimization process. We need to know what (5) developable and existing hydraulic/geothermal/solar resources, because these may be cost effective and would thus be included in the economically optimal system. Within this category, we specifically need to know about proposed/existing locations, the year the unit came/will come on line, the existing and/or initial cost, the capacity and energy constraints on any such resources, and, if the source is on line, both the hours connected to load in the most recent calendar year and the normal O & M expenses per year.

The final set of data requirements for an analysis of the generation function focuses on the exogenous political, institutional, and operating constraints which utilities face. For example, a state commission may allow companies to use no more than a certain percentage of oil-fired generation. Or, the current political milieu in which utilities operate may effectively prohibit the installation of nuclear capacity. Operating constraints refer to the need for reserves: operating reserves, capacity (installed) reserves, and reserves for required maintenance. The first two types are reserves for forced outages, while the third is a reserve for planned outages.

In our experience, the best sources for these data are direct talks with representatives of the companies which we are analyzing and the system agreements. In outline form, these data requirements are: the political and institutional constraints which limit the selection of equipment in the system planning process, because this may have a large impact on the economically optimal but still feasible system; (7) the system agreement, because this explains in detail the extent of reserves which must be carried and further affect the economically optimal system; and (8) general maintenance requirements by each type of technology, because these can and do change the "marginal" machine.

The Ernst & Ernst method for generation analysis results in an optimally met load duration curve (LDC) which is, we have been told and believe to be, the basis for long-run planning used in the electric utility industry. As a final step, the marginal costs are computed for each point on the LDC (load levels). This is done in accordance with the theoretical definition of long-run marginal cost by examining the change in costs of the economically optimal system as loads change.

Note that, at this point, no mention has been made of rating periods or when to charge the various long-run marginal costs which flow from our analysis. This reflects our belief that load levels with similar marginal costs should define the rating periods. To determine rating periods before the costs of those rating periods are defined is, in our opinion, a backward way of approaching the problem.

The Analysis of Nongeneration Functions

The production function approach which we use for the analysis of generation does not readily lend itself to nongeneration applications. This is primarily due to the numerous specialized factors such as terrain, size of service territory, voltage levels, and so forth, which would need to be considered. To incorporate all of these factors would render the analysis analytically intractable. For this reason, a simplified approach is used.

The simplified approach used for the analysis of nongeneration cost components involves regression analysis to relate an annual, price adjusted series of costs of nongeneration functions to those factors which cause the costs. The annual costs associated with plant investment as well as price adjusted operations and maintenance expenses are determined for a twenty-year sample period. The causal factors are sales in kwh, number of customers, and system coincident peak. The resulting total cost equation is then used to derive marginal costs. Transmission and distribution O&M expenses as well as sales expenses, administrative and general expenses, and customer accounting expenses are indexed easily with the use of the Handy-Whitman index.

To begin this analysis, the following data are required: (1) historical data, annually, for the past twenty years, on transmission O&M expenses; transmission plant investment, retirements, and transfers; distribution O&M expenses; distribution plant investment, retirements, and transfers; sales expenses; administrative and general expenses; customer accounting expense; number of customers by class; annual sales of energy by class; and annual system peak. The first seven items, after price adjusting, serve as our independent variables; the last three as our independent variables.

The greatest difficulty in our analysis is, obviously, reacounting for plant investment in transmission and distribution facilities to arrive at a price leveled, twenty-year historical series.

The data and assumptions which we use for this procedure for both transmission and distribution plant are: (2) the beginning plant balances in the first year of observation, because this serves as our basis for the reconstruction of the plant investment series, and (3) the average life of the investment category in the latest year, because we assume that the average life remains constant throughout the sample period.
Windmills of Costing Methods

We further require (4) the accumulated depreciation for each category in the first year of observation, because of the need to reconstruct the reserves in the indexing process. With these data, we can perform our indexing.

The indexing calculations used vary by type, that is, additions, retirements, and transfers. The first step is to index the beginning balance and accumulated depreciation, assuming they are both composed of equipment that is one-half the composite average service life. Then, for each year, these figures are updated using current dollar amounts of additions, retirements, and transfers and the assumed average depreciation rate. With these indexed values, a current dollar net plant investment is calculated.

Since additions are purchased in the year when they are reported, the basis for their index is that year. This amount for each year is added to the previous year’s current dollar plant value. One-half of the indexed plant additions are used in computing each year’s depreciation expense.

Since the vintage of the retirements is unknown but is generally the oldest equipment, the retirements are removed from the beginning balance. That is, all retirements in the sample are assumed to have been of one-half average service life during the first year of the sample. These amounts are subtracted from current cost plant balance, and one-half their value is used in the annual depreciation computations.

Transfer amounts between transmission and distribution are handled by assuming that, when of opposite sign, the lesser quantity is transferred entirely to the other category. Thus, there are two components in the analysis of transfers, offsetting amounts and excess amounts. Each is indexed differently.

The offsetting amounts are proportioned out over the past sample years on the basis of original cost. This procedure attempts to recognize that some portion of the transfer may have come from the most recent additions, while another portion is original plant. This resulting stream of transfer components, which, of course, must sum up to the transfer amount, is then indexed to current dollars using as a basis each applicable intermediate year. This amount is then transferred between the two categories, and accumulated depreciation is adjusted accordingly. The excess amount is indexed as if it were one-half average service life. The result is then transferred, and accumulated depreciation is adjusted.

We then have two investment streams, one for transmission and one for distribution. To derive the annual cost, we multiply each series by the relevant fixed charge rate. For the fixed charge rate calculation we need the same data as listed above in item 3 of generation for transmission and distribution plant. Recall that the expense items have been indexed. Thus, when combined with the price indexed cost of plant for transmission and distribution, we have five twenty-year, price indexed, historical series of costs for each nongeneration cost component. These are then regressed in the relevant dependent variable to arrive at a long-run total cost function, the derivative is taken with respect to the relevant output variable, and this becomes our marginal cost estimate.

Development of LRMC-Based Rates

Based on our generation analysis, we now know the costs associated with the various load levels. We must, therefore, determine the times when these load levels are achieved. The first data item necessary for this determination is (1) hourly system loads, because we need to determine periods of high demand, medium demand, or low demand, that is, peak, shoulder, and off-peak periods, respectively. We can then, by using a linear program, determine what the peak, shoulder, and off-peak rating periods should be.

The next step is to put the generation and nongeneration functions on a common basis. The nongeneration costs are expressed on a sales basis, while generation costs are on an energy generated (or busbar) basis. To do this, we need (2) the average system energy and demand loss factors, because these will enable us to express all costs on a common basis. Because of the need to reconcile the full LRMC-based rates to the revenue requirement, the next item we need is (3) the revenue requirement by function (generation, transmission, distribution, and so forth). Asking for the revenue requirement by function reflects our philosophy that utilities should recover the costs as defined by the revenue requirement, but that collections from one cost component should not subsidize another. We feel this to be required for the development of true cost-based rates.

Following this reconciliation, we need only do one more adjustment. This is an adjustment for line losses. Because of the line loss phenomenon, a utility may need to produce more electricity (kilowatt-hours) at a higher rate (kilowatts) at the busbar or transmission level than at the secondary distribution level, for example. Therefore, consumers at the end points of the transmission and distribution network should pay more (reflecting the greater losses) than those at intermediate stages. The data necessary for this adjustment are (4) cumulative energy and demand loss factors at each consumption point.
of the transmission and distribution network. Implicit in the above discussion is that, for purposes of rate design, a consumer at the transmission level should not pay any distribution costs.

**Supplemental Data Request Items**

This final section of the philosophy of and data for the Ernst & Ernst LRMC methodology is intended to list data which we like to compile, but which are not absolutely necessary for the estimation of LRMC and the corresponding rates. The Ernst & Ernst method relies on neoclassical production functions, and the more current our data, the better our estimates of the production functions. Accordingly, the better will be our estimates of long-run marginal cost. In an attempt to strengthen these production functions and replace older units in the sample with newer units, for each thermal generating unit in the system or scheduled to come on line in the next three years, we need to know its type, for example, steam, combustion turbine, and so forth; net continuous generation capability; use, for example, base load, cycling, peaking; year the unit went on line or is scheduled to come on line; original cost of the unit and associated equipment; annual or projected average heat rate; normal or anticipated O&M expense per year; and test and design heat rate curves.

---

**Another Approach to**

**Time-of-Use Rates**

*Alan Schoor*

As is becoming increasingly clear, the supply and demand for electricity are intertwined. Recent events illustrate that not only electricity demand but also all forms of energy consumption are price responsive, to some degree. Changes in future levels of electricity prices will undoubtedly affect the level of future demand. Similarly, a move to time-of-use rates will most likely result in modifications of system load patterns. These patterns may, in turn, affect the amount and type of system capacity (that is, generation, transmission, and distribution facilities) required in the future. Consequently, the decision as to the proper costs upon which to base price signals to consumers will affect both the utility and its customers.

Over the past several years, the question of which costs to use in rate making has received much attention. To this debate on rate structure reform must be added another factor, load management. Utilities are being requested, and in some areas mandated, to encourage load management. Many believe that the areas of utility planning (including load forecasting, rate design, and load management) are interrelated and must be dealt with on a consistent basis.
Load management has many different definitions, but defining it as any deliberate attempt to modify or reshape a utility’s load curve encompasses all the mechanisms generally considered as load management. These range from time-of-use rates to mandatory controls on storage water heaters. To a limited degree, load management can be regarded as an alternative to system expansion. It will not eliminate the need for expansion, but it may reduce it. However, load management is appropriate only if it is cost effective. It is this aspect that links the areas of planning, rate design, and load management. Exhibit 1 illustrates the interrelationships between various utility processes and load management. It underscores the fact that all these different activities are intertwined and thus must be conducted on a consistent basis.

Increasingly, utilities are being called upon to guide customers in the determination of appropriate load management equipment. A system planner asked to evaluate the cost effectiveness of a load management option undoubtedly would reply that an analysis is required of the costs that can be saved by the utility compared to the additional costs a customer will incur (including both equipment and socioeconomic costs). Questioned further, the planner would say that the relevant utility costs are the incremental costs the firm could avoid as a result of the customer’s load changes.

A rate structure should be designed in such a way to signal customers the cost of the scarce resource that would be used, or saved, should their consumption increase or decrease (that is, marginal costs). Furthermore, these marginal costs must reflect the costs of the probable “real world” response of the system, not a theoretical abstraction. In this way, consumers’ decisions (for example, an analysis of load management options) and the utility’s planning process will be consistent.

The first step is to understand the utility’s objectives for projected load changes and its constraints. This methodology seeks to simulate the results of the utility’s actual goals. In this way, an estimate can be obtained of how costs will change at the margin when loads change marginally.

After the marginal costs for the coming year and a number of subsequent years have been computed, rates based on marginal costs can be developed. Although one major purpose of such rates is to charge for marginal consumption at marginal cost, the overall rate design must also reconcile those revenues produced under marginal cost pricing with those allowable by the jurisdiction involved. In addition, consideration must be taken of equity, rate continuity, consumer comprehensibility, and revenue stability. Finally, the degree to which
time-of-use cost variations are directly reflected in such rates for each customer class must be supported by a cost-benefit analysis which shows that benefits outweigh the additional metering, billing, and other costs necessary to institute such rates.

**Signaling Cost Information to Customers through Rates**

As mentioned earlier, the purpose of marginal cost-based rates is to signal to customers the cost of providing for additional consumption or the savings resulting from decreased consumption. From the point of view of economic efficiency, customers should be charged only for the avoidable costs which their marginal consumption causes to be incurred. On this basis, past costs are irrelevant, and the only genuine costs are those caused by present consumption (for example, increased maintenance costs and shortened life of equipment).

The marginal cost of providing service at any time, which is the avoidable cost of resources thereby utilized, is called short-run marginal cost. It consists of the sum of the marginal running (or energy) cost and a marginal charge for maintaining optimal system reliability.

The marginal running cost is the additional operating cost incurred by the unit dispatched to pick up the marginal load. To the extent that this load results in additional wear and tear on any part of the system, the corresponding cost of maintenance or earlier replacement must be included in the running cost.

If reliability of power supply were of no concern, economic efficiency would require that rates be set at marginal running costs. However, since customers assign a value to the reliability of supply in addition to the value they place on the electric energy they receive, economic efficiency also requires that a marginal demand charge must be added to the marginal running cost to constrain demand at a level at which the "shortage" cost per marginal kw of demand just equals the marginal demand charge. When this equilibrium occurs, the level of reliability is considered optimal. Naturally, the demand charge needed to achieve this equilibrium will be lower when ample capacity is available in relation to demand and higher when the reverse is true.

Although economic efficiency requires that customers be provided with price signals reflecting short-run marginal costs since these represent the costs of the actual resources that will or will not be utilized, it is also necessary to provide an estimate of the future trend in such costs. When customers are making investment decisions, they thus can account for the probable cost of electricity over the life of the investment.

Large industrial and commercial customers can be expected to comprehend and respond to information on short-run marginal costs and their likely trend. For smaller customers, however, it may be necessary to base price signals on an average of current and forecast marginal costs (expressed in terms of today's dollars).

In addition to estimating marginal costs in the short run, one can estimate, from a planner's perspective, the additional costs, as perceived now, that would probably be incurred in satisfying a marginal increment of load some years in the future. In the case of generation, this would involve estimating what changes in capacity additions and retirements, and in the scheduling of generating units, would result from a perceived future change in load. If one is looking three years into the future, the only generating capacity additions that may be feasible, over and above those to which the utility already is committed, are combustion turbine units.

Farther in the future, other types become feasible, for example, coal and nuclear units. Because of such differences in feasible capacity rearrangements, the marginal costs of meeting prospective loads will depend upon how many years into the future one is looking. How should one designate such marginal costs? It is not worthwhile to engage in semantic arguments on this point. Suffice it to say that the long-run equilibrium envisaged by the theorist not only assumes that all factors of production are variable, but also ignores existing facilities in determining the least-cost optimal equilibrium. Ignoring existing facilities and postulating an idealized situation with an optimal mix of facilities is an abstraction of little value in estimating real world marginal costs.

In this paper, the term "short-run marginal costs" is used to denote such costs in the immediate present and projections of them in future periods. The determination of short-run marginal costs for any future period involves computing two different (but related) components. The first is the net cost of augmenting or maintaining in service that capacity needed to serve marginal demand in the period while meeting a specified reliability criterion. The second is the marginal running (or energy) cost after such capacity changes take place.

**Methodology**

The diagram in Exhibit 2 shows the basic components of the computerized generation optimization model developed by Gordian Associates, Inc. As can be seen, the model is provided with load forecast data for the years to be considered (for example, ten to twenty years hence). In addition, it is provided with complete data on the operating characteristics of existing units; investment costs and operating charac-
teristics of potential units; trends in all running cost elements, such as fuel, labor, and maintenance; reliability criteria for the span of years being modeled; and the incremental cost of capital to the utility.

To obtain an optimal solution to the capacity addition problem, the optimization goal ("the objective function") must be defined. This is normally the minimization of the present value of all future utility system costs and, hence, of future revenue requirements. After an optimal solution has been obtained, capital outlays and operating costs are transferred to a financial model which develops pro forma financial statements for a span of years starting from base-year financial data. To the extent that a financial constraint is violated, the capacity addition plan is reoptimized subject to the constraints.

After an optimal expansion and operating plan has been developed which satisfies financial constraints, the hourly marginal costs of generation are computed. These costs, along with the marginal costs of transmission and distribution, can then be used to develop marginal cost-based rates. This step requires definition of the revenue adjustment mechanisms to be used in designing rates which conform to the allowable revenues. In addition, considerations of equity, continuity, and stability must be introduced at this point.

As indicated in Exhibit 1, a projection of the rates must be provided to the load forecasters to determine in what way, if any, the original load forecasts need adjustment. Current data on the price responsiveness of demand are generally inadequate, particularly for estimating usage between different rating periods; however, some estimates of these price effects should be included, along with estimates of the effects of device-specific load management, before the rate design is completed.

The result of the iterative process is rates based on marginal cost which represent an equilibrium between the price charged and the quantity demanded.

The first step in marginal cost pricing is to estimate hourly marginal costs for the coming year and a number of subsequent years. These hourly marginal costs can be differentiated as follows: production cost (including running or energy cost and demand cost), transmission cost, distribution cost, and customer cost.

Marginal Production Costs

Marginal production costs are those associated with generation. These can be segregated into two components: running costs and demand related costs.
Marginal Running Costs

Marginal running costs are the additional operating costs incurred by the unit dispatched to pick up marginal load. Specifically, they are comprised of the cost of fuel burned by the "marginal unit" plus other variable production costs, such as variable maintenance, water treatment, and so forth.

The Gordian approach develops marginal running costs from an hourly dispatch simulation model. The units available for dispatch in each year being studied are obtained from the solution to the linear programming optimizer. It should be noted that the marginal cost of providing an additional kilowatt-hour need not correspond to the running cost of the most expensive unit on line, but of the unit that is incrementally picking up the load.

The dispatch simulation schedules units, hourly, in order of increasing incremental running costs, subject to constraints on the cycling of different types of units. This is achieved by restricting the minimum operating capacity of a unit that will be required to pick up load within a specified number of hours. The dispatch simulation optimally schedules the operation of pumped or other energy storage units within the constraints of energy storage capacity, input-output capacity, and input-output energy efficiency.

Accordingly, the model produces marginal running costs for (1) current or early forecast years when no generation capacity addition is possible. The program dispatches available units in order of increasing costs; and (2) future years, when generation capacity additions become possible. The generation mix in any year is obtained from the output of the linear programming optimizer. The dispatch simulation is then rerun with this generation expansion plan and the forecast daily load curves.

During any period, the utility imports or exports power through interchange with neighboring utilities, the marginal running cost will be the cost of marginal energy imported or the realizations from marginal energy exported, respectively. To the extent that the dispatch simulation is provided with the timing and pricing of interchange opportunities, the marginal running costs will automatically account for such interchange. This approach is feasible when the utility is not a member of a centrally dispatched power pool and when limited exchange occurs on an ad hoc basis. When the utility is a member of a centrally dispatched power pool, it will probably be necessary to represent the entire power pool in order to obtain meaningful marginal running costs. The Gordian methodology has been used successfully for this purpose.

It should be noted that the marginal running cost to the specific utility being studied will not necessarily be the system lambda that the central dispatcher sees. Rather, it will be the cost as computed in conformance with the interchange agreement (for example, halfway between the marginal energy cost of the unit providing the energy and that of the receiving unit's unit which otherwise would have provided the energy).

Marginal Demand Costs

In addition to the value of the electric energy they use, customers assign a value to the reliability of supply. Since reliability preferences differ, the ideal situation, from an economic efficiency viewpoint, would be for each customer to select his own level of reliability based on a trade-off between the value to him of additional reliability and the additional cost he would have to pay. The use of interruptible rates is an approach which would achieve this for large customers. Another example is the recent suggestion of the California Energy Commission that residential customers be allowed to select the extent of cycling of air conditioning that suits them, with a larger incentive being offered for the acceptance of longer periods of cycling each hour.

In general, such tailored reliability is not feasible. At present, data on customers' reliability preferences are inadequate to permit an exploration of the determination of the optimal level of reliability. However, the utility must provide "firm" power at some optimal level of reliability. Based on historical experience, utilities, power pools, and/or reliability councils have established reliability criteria which implicitly reflect the preferences of customers.

Determining the net marginal costs of capacity requires the use of a reliability criterion which conforms to that used by system planners in determining the level of additional capacity needed. In many cases, loss of load probability is used, and this is the present basis for the Gordian reliability program. However, since this procedure has certain drawbacks, the Gordian model is designed to handle other types of reliability measurements as long as they can be specified mathematically.

When Targeted Reliability is Achieved. In computing marginal demand charges, it must be recognized that for the coming year, and for the next several years, construction lead times will limit capacity additions to those units already under construction and scheduled for start-up during that period. The utility will be unable to adjust capacity to match the targeted reliability criterion, and reliability may either be lower or higher than the target.

Starting with the year when it becomes feasible for the utility to
adjust the capacity mix, the linear programming optimizer will do so if such capacity is needed to meet the optimal reliability criterion. As discussed earlier, if we assume that the specified reliability criterion is optimal, then, by definition, the marginal cost of capacity will equal the marginal cost of curtailment when the targeted reliability criterion is met. It follows that the marginal demand charge needed to constrain demand to a level which results in optimal reliability will equal the marginal capacity cost.

The linear programming optimizer will select the type, size, and timing for the addition of new generating capacity, as well as the mode of operation for all capacity. This minimizes the present value of avoidable costs while satisfying forecast system loads and a specified reliability criterion. The optimizer will also provide, for each year being considered, the marginal capacity cost. This is the net cost of adding one kw of additional generating capacity and equals the economic carrying charge for that capacity less the running cost savings which result from this capacity displacing other capacity with higher "running costs." This net capacity cost recognizes the effect of future relative escalation rates in the various types of fuels and plant equipment. Therefore, the net marginal capacity cost of generation which results from the model reflects the capacity—fuel cost trade-off inherent in the generation planning process.

The marginal demand charge is expressed in terms of a marginal kw of demand in any hour; hence, the relationship between marginal demand and marginal capacity must be established. This is done through the generation reliability program, which indicates how much additional capacity must be added to the reserve to maintain the reliability criterion when demand increases. The magnitude of this capacity is a function of the characteristics of the capacity being added, as well as the existing mix. Thus, the net marginal capacity cost produced by the model reflects the reserve margins necessary to maintain the targeted reliability level on a unit-by-unit basis, rather than on an average system basis.

OVER-OR UNDERCAPACITY AND THE TARGETED RELIABILITY CRITERION. If generation overcapacity presently exists, then the current computed reliability level, and possibly the reliability level for a number of succeeding years, will be higher than the targeted reliability criterion. On the other hand, if undercapacity presently exists, the reverse situation will obtain. Under these circumstances, how should the marginal demand cost for firm power be computed in the short run?

The Gordian method first determines what the marginal demand cost would be if available capacity exactly matched that needed to meet the targeted reliability criterion (for example, one day every ten years). This can be done by analyzing the net marginal capacity cost produced from the model for several years in the future. This cost can be interpreted as the value to the average customer of having the utility satisfy his last kw of demand at the targeted reliability level. From a different perspective, it can be viewed as the cost of probable interruptions which are avoided through the utility adding capacity to meet the customer's marginal demand at the targeted reliability level.

The next step is to determine the utility's anticipated reliability level for the next several years as compared to its targeted level. If, for example, the system is more reliable than the planned level, then the value of the last increment of reliability to consumers will be less than the marginal cost produced from the model (which is at the desired reliability level). Similarly, if the system is less reliable than the desired level, the value of the last increment of reliability to consumers will be greater than the marginal cost at optimum. As discussed earlier, there is no way presently to measure consumers' shortage cost. However, two approximate approaches can be used.

First, if some information on customer demand responsiveness is available, one could estimate the price that would be required to bring demand into balance with supply (since, in the short run, capacity adjustment is not feasible). Second, if sufficient data are not available, a simplifying assumption could be made to estimate the relationship between the cost of interruptions to customers and the reliability level. For example, one might assume that the cost of interruptions varies linearly with the relative reliability level so that the more reliable the system (relative to the targeted reliability level), the less will be the value of the last increment in reliability to customers (the cost of shortage will be less).

However, in no case would the demand-related component of the short-run marginal cost be less than the cost to the utility of providing the last increment of reliability. In the case of excess capacity, this would be the cost incurred in maintaining the least economical unit in reserve rather than retiring it. The decision as to which capacity cost (the pure short-run cost or a longer term cost averaged over several years) to use in developing actual marginal cost-based rates is discussed in more detail later.

Marginal Transmission Costs

Marginal transmission costs are related to the planning of the transmission system. While the voltage and size of such systems depend
upon utility size, geographic area covered, degree of interconnection, and reliability criteria, among others, the basic concept used in determining the marginal cost of transmission is the same: What would be the effect (cost) of adding an additional customer, kilowatt, or kilowatt-hour to the transmission system?

The transmission system is comprised of several subsystems (more for large utilities or those heavily interconnected) that may be affected to different degrees by the factors noted above. These subsystems may include a bulk power EHV grid, an intermediate level regional transmission system, as well as a more localized subtransmission system, with typical voltages of 345-765 kV, 115-230 kV, and 23.7-89 kV, respectively. In developing marginal costs for a particular utility, these distinctions must be understood in order to identify the cause-and-effect relationships. For example, the expansion of the EHV system may be more sensitive to additions to the generation system than the regional 115 kV transmission system, and the analyst must recognize this in his study.

The way to determine the marginal cost for transmission is to study the effects, on each segment of the transmission system, of additional customers, peak demand, and kilowatt-hours. In theory, this should be done by reoptimizing the transmission system for each of these changes, as is done in the development of marginal generation costs. However, as any transmission planner will quickly point out, the nature of transmission planning is such that this kind of analysis, while possible, would be very difficult and time consuming. This is because the transmission system is comprised of many components, such as transformers, breakers, phase shifters, and so forth—each of which is affected differently by growth depending upon when and where it occurs. A detailed process analysis would require analyzing numerous transmission load flow cases and evaluating numerous alternative expansion plans, from both technical and economic aspects. While transmission optimization models do exist, their application is somewhat limited at this time.

It is generally agreed that the major factor affecting the amount of transmission investment is the peak load of the system. However, the transmission system may be comprised of different components which may be more related, in terms of cost causation, to factors other than peak load. These factors include: (1) generator location, capacity, and stability requirements; (2) interconnection capability for reliability and economic dispatch; and (3) loss minimization. Most planners would agree that local transmission (115-230 kV) is related to the need to maintain service at peak load conditions during outages. However, at higher voltages (345 kV and above), the purposes and causal relationships become less clear-cut. In fact, most EHV lines serve several utility purposes simultaneously, including the wheeling of power for others, which further complicates the issue. Finally, the peak demand that is used to analyze transmission expenditures actually should be called peak planning demand since it is not necessarily the system peak demand. For example, the local or subtransmission system may be planned on the basis of the area peak, the 115 kV system on the basis of the regional peak, and only the EHV system on the basis of the overall coincident system peak.

The Gordonian approach analyzes historical and projected future expenditures for transmission. Investments should be segregated by voltage levels so that the various segments of the transmission system can be analyzed separately with respect to the appropriate load growth. It is important to investigate a period of sufficient length to smooth out unusual variations in investment which may be especially pronounced in the case of higher voltages. Furthermore, care must be exercised to ensure that the results (especially at higher voltage levels) are not distorted due to shifts in the timing of investments associated with generation additions.

It is always useful to cross-check the results of this type of analysis with those from an engineering approach. This can assist the analyst in determining whether the period studied was atypical because of a lack of load growth, a change in planning criteria, or a curtailment in construction due to financial constraints.

Although transmission investment usually is related to peak planning demand, this does not imply that all costs are associated with the peak only. The correct way to examine these costs is by using tech-
upon utility size, geographic area covered, degree of interconnection, and reliability criteria, among others, the basic concept used in determining the marginal cost of transmission is the same: What would be the effect (cost) of adding an additional customer, kilowatt, or kilowatt-hour to the transmission expenditures actually should be called peak planning demand since it is not necessarily the system peak demand. For example, the local or subtransmission system may be planned on the basis of the area peak, the 115 kV system on the basis of the regional peak, and only the EHV system on the basis of the overall coincident system peak.

The Gordon approach analyzes historical and projected future expenditures for transmission. Investments should be segregated by voltage levels so that the various segments of the transmission system can be analyzed separately with respect to the appropriate load growth. It is important to investigate a period of sufficient length to smooth out unusual variations in investment which may be especially pronounced in the case of higher voltages. Furthermore, care must be exercised to ensure that the results (especially at higher voltage levels) are not distorted due to shifts in the timing of investments associated with generation additions.

It is always useful to cross-check the results of this type of analysis with those from an engineering approach. This can assist the analyst in determining whether the period studied was atypical because of a lack of load growth, a change in planning criteria, or a curtailment in construction due to financial constraints.

The problem associated with aggregate time-trend analysis is the inability to separate energy-related investments. While the need and timing of a line may be determined by peak planning demand and reliability criteria, the actual size of the line may be related to projected average transmission loading or the average load factor. The planner determines the cost of losses against the capital cost of lines for the expected load growth to determine the economic size of the line. Accordingly, the determination of the investment in transmission facilities should be considered to be energy related.

In a similar manner, other transmission investments which are related to factors other than load growth should be analyzed separately. While this is generally difficult to do on an historic basis because of the way in which records are kept, it is usually possible to perform this more detailed analysis for the forecast period.

Although transmission investment usually is related to peak planning demand, this does not imply that all costs are associated with the peak only. The correct way to examine these costs is by using tech-
niques similar to those employed in determining the marginal cost of generating capacity by time of use. This method develops hourly demand costs as a proxy for the marginal charge for the customer's reliability preference. While transmission equipment is also subject to forced outages and maintenance requirements, the level of these is significantly less than for generation equipment. If it is assumed that a fixed amount of capacity is available year-round, then the need for additional capacity is related to the load on the system. Expressed in reliability terms, the need for additional capacity on the transmission system is related to the probability that load will exceed capacity. In reality, the simplifying assumption concerning fixed capacity availability is not appropriate because of capacity variations due to ambient temperature changes, high seasonal forced outage rates due to seasonal weather conditions, and so forth. The analysis of seasonal and/or daily risk patterns must take these effects into account.

Simply stated, the method we suggest for developing time differentiated cost responsibility factors is based on the relationship between the expected load level and the expected capacity available. This requires obtaining hourly load patterns and capacity availabilities for transmission equipment. While hourly loads are not generally available for transmission lines (except, perhaps, for EHV lines and interconnections), this type of data may be available for transmission transformers. The effort required can be minimized by selecting a random sample of transformers at each voltage level. Applying statistical techniques to this information, the probability that a load increase in any hour will result in the need for additional capacity can be determined.

It should be pointed out that, unlike energy availability, transmission and distribution capacity availability is location specific. While one sector of the system may be heavily loaded in summer and another in winter, the overall transmission system reliability results may not recognize this and indicate equal risk in the summer and winter. However, attention should be paid to these seasonal risk patterns. Since the extent of differences and the administrative costs of developing and maintaining area-specific cost data are the ultimate determinants of the appropriateness of corresponding rate differentiation, it would be improper to ignore such differences in the marginal cost analysis.

Marginal Distribution Costs

The expansion of the distribution system is necessitated by growth in demand and new customers. Much has been written regarding segregation of customer and demand costs, ranging in opinion from whose, like James Bonbright, oppose the use of a hypothetical system of minimum capacity as a customer cost, to those who favor a minimum system based on today's construction standards. It is impossible to prove that one method is better than another, but most practitioners and commissioners have accepted the use of a minimum system customer cost. Furthermore, since the customer cost is a function of system density and local terrain, such differences should be recognized.

Expressed simply, distribution costs can be represented as a function of both demand and customers: distribution cost = f (number of customers, peak distribution demand).

As others have pointed out, it is almost impossible, statistically, to separate the effects of customer and demand growth over time. A possible solution is a cross-sectional analysis using data for different utilities, but this has a very critical shortcoming: Many factors which significantly affect distribution costs are not amenable to simple quantification and/or are not readily available. Two examples of such factors are differences in geographical characteristics (hilly versus flat terrain, rocky versus porous soil) and reliability standards (double versus single contingency planning). Consequently, the method we recommend requires a separate analysis of customer and demand costs. The customer cost component (marginal hook-up cost) is defined as the cost for a distribution system capable of serving a hypothetical minimum-load customer. If, as discussed earlier, significantly different costs exist in different areas served by the utility, different customer costs should be computed. Similarly, different customer costs should be computed for overhead and underground service when appropriate.

After these components have been defined, the demand-related elements of distribution investment can be determined. The first step is to select a period for analysis. As suggested in the discussion on transmission, the period should be sufficiently long to smooth any temporary aberrations and ideally should encompass both an historic and prospective time span. Unfortunately, most companies do not have distribution investment forecasts in adequate detail for many years into the future, so the analysis is generally limited to an historical period. After selecting a suitable time frame, the demand-related component is obtained by subtracting the increase in the customer cost component from the total increase in investment in distribution costs.

Care must be exercised in relating investment to load growth. As in traditional cost-of-service studies, it is necessary to recognize that growth in local distribution facilities is probably most closely related to the sum of the customers' individual maximum noncoincident demand, while growth in primary voltage level investment is related to the sum of the individual class peak demand. The actual demands used
for analysis will depend upon the specific utility's planning process. This is particularly important in analyzing situations in which system load growth in the last several years has been relatively low, while the growth in customer noncoincident and peak demands may have been close to historical levels.

The demand-related investment costs should be disaggregated to determine the sensitivity of such costs to the explanatory variables. These components should at least consist of primary and secondary investments and, if possible, should be further disaggregated.

While the marginal demand-related distribution costs are expressed per kilowatt of peak planning demand, this is not to imply that these costs are all peak related. As discussed earlier with respect to transmission, an analysis of the loading patterns on substations at all voltage levels should be undertaken to develop the responsibility factors for demand-related distribution costs by time of year and time of day.

Other Marginal Costs of Power Supply

There are, of course, other costs of power supply which must be examined to ascertain whether any portion of them can be causally related to customers' consumption decisions. These include administrative and general expenses, customer accounts and sales expenses, and common and general plant costs. All of these, although real costs to the utility, may not be marginal costs. For example, the franchise cost in certain areas is not dependent upon utility size but is a fixed cost per year. In such a case this should not be regarded as a marginal cost.

Loses must be analyzed to determine their effect on marginal energy and demand-related costs at different voltage levels.

Operation and maintenance expenses also must be analyzed to determine what portion of them can be causally related to demand, energy, and number of customers.

Marginal Cost-Based Time-of-Use Rates

To devise rates using marginal costs, several steps are required: (1) The revenues which would be produced by charging rates equal to marginal costs must be computed for a test year (or years), on a class-by-class basis; (2) the class revenue targets must be established for use in designing actual rates; and (3) an actual rate structure must be designed using marginal costs which attempts to provide the consumer with the correct price signal while yielding class (and company) revenues allowed by regulators. Exhibit 3 contains a simplified flowchart of these steps.
This step is necessary in order to determine whether there is a revenue excess (or deficiency) for the company as a whole and to give guidance in the determination of class revenue targets.

One consumption unit for each class for a test period must be determined. These should include the kilowatt-hours and kilowatts by time of day, season, and class. Such data, which are normally not available for any but the largest customers, should be extrapolated from the company's load research program. If they are unavailable, suitable estimates must be made using available information and data from other sources or companies.

The consumption data are then multiplied by the relevant marginal costs to yield class revenues for the test year. In theory, this should be done on an hourly basis using actual marginal costs. However, this would require a significant computational effort, both to obtain the required load data and to compute the revenues. Therefore, the practical approach is to group periods of similar costs before making computations.

Many regulatory jurisdictions establish not only the total revenues a company may receive but also the revenue level for each class based on average embedded costs. However, it is also possible to use marginal costs to set class revenue targets. It is highly unlikely that the revenues produced by charging marginal costs will be equal to either the total revenue requirement of the company or the revenue requirement for each class. The next step, therefore, is to select a method that will adjust those revenues produced under direct use of marginal cost levels. There are two broad categories: those approaches which reproduce existing class revenue targets as well as the overall company target, and those which are constrained to produce only the overall company target.

In theory, the revenue level for a particular class should not affect the design of time-of-use rates. An adjustment procedure could be designed (commonly referred to as a lump-sum transfer) which would result in price always equaling marginal cost, the necessary and sufficient condition for economic efficiency. In practice, such adjustments are not possible, and the amount of adjustment required to reconcile marginal costs with revenue level will affect the prices seen by customers.

It has been argued that in order to minimize the required adjustment to marginal costs (and hence minimize distortions to economic efficiency), marginal cost-based revenues should be used as the starting point in specifying class revenue targets. While it is not possible to prove that this necessarily minimizes distortions to economic efficiency, it should minimize the required adjustments in marginal costs. Several different methods are available, but Gordian believes that the class revenues which would result from use of marginal costs should at least be examined before deciding on class revenue targets. In general, Gordian favors the uniform proration approach to determine class targets since it treats all customer classes equally, thereby avoiding charges of discrimination.

In deciding upon the appropriate revenue target for a class, however, many factors must be considered in addition to marginal cost revenue levels: continuity, equity, revenue stability, and so forth. Obviously, it is impossible to make any broad generalization as to which method is necessarily best in a particular instance.

Another solution is to use the existing class revenue targets (however determined) and design time-of-use rates to match them, ignoring marginal costs for setting rate levels. However, this poses two potential problems. First, the utility would have to perform both marginal cost and embedded cost studies, thereby increasing administrative costs. Second, if the class revenue distribution differs markedly between the embedded cost and marginal cost studies, it may not be feasible to design meaningful time-of-use rates that reflect marginal costs. The degree to which the two revenue distributions differ is a function of several factors, the most important of which are: (1) the present method used to set class revenue targets; (2) the relationships between the embedded cost-of-service allocation factors and seasonal marginal costs; and (3) the method used to compute marginal costs.

Rate Design

In designing rates, there are numerous aspects to consider, but three are particularly worthy of discussion here. These are selection of the rating periods and actual construction of time-of-use and interruptible rates.

Rating Periods

If transferring price information to consumers were costless, then as much price delineation could be undertaken as was warranted by variations in marginal costs. In practice, the cost of communicating such information is not insignificant. Consequently, determining how much price delineation is warranted requires a cost-benefit trade-off between the value of communicating additional pricing data and the cost of transmitting them.

The more price signals conveyed to customers, the higher the costs
of metering (and perhaps the less customers will understand). Thus, it is generally necessary to group costs into rating periods. These are defined as predetermined sets of hours during which a uniform rate for electric service is charged.

Rating periods should be selected after marginal costs are computed, and the variation in costs should be the primary determinant in choosing the periods. Since there are two basic types of marginal costs — demand and energy — and these vary across time, both costs should be considered. Furthermore, the marginal costs that should be used in selecting rating periods are those costs after response to time-of-use rates has been taken into account. The degree to which system marginal costs can reflect future responses to time-of-use rates will depend upon the availability of demand elasticity data; at the present time, these are limited. Accordingly, future responses may have to be estimated using the enlightened judgment of the analyst.

Selecting rating periods can be divided into two steps. The first would be the selection of cost periods. These would be groupings of hours of similar marginal costs and can be as numerous as is warranted by cost variations. All other factors are not recognized at this point. The second step would be to select rating periods from the cost periods which take into account other economic and social factors to be considered. These rating periods may be different for individual classes because of other considerations: consumer comprehension, consumer responsiveness, revenue stability, data availability, stability in rating periods over time, and implementation and administrative costs.

The importance of these factors will vary depending on the class being analyzed and the viewpoint of the utility or class served. Some are not quantifiable and will require judgment. The cost-benefit study must weigh the advantages to consumers of time-of-use rates (in the form of lower costs of providing electric energy and maintaining reliability) against the costs of implementing them (including metering, administration, consumer equipment, and the socioeconomic costs associated with shifts in usage patterns). The decision as to exactly which rating periods to choose for a particular utility or class should be based on detailed information on the type of industry in the service territory and existing hours of shift work, the daily working hours of the commercial work force, business hours, residential appliance stock, and alternative metering costs, to name only a few factors.

**Time-of-Use Rate Design**

Once the revenue targets have been established by class and the rating periods selected, it is necessary to design the actual rates. Since class revenue targets generally will differ from the revenue produced by charging marginal costs, it is necessary to adjust the marginal costs to produce the targeted class revenue levels.

While there may be strong justification for use of a particular rate design from an economic efficiency perspective, regulatory authorities and the courts have long recognized the concepts of continuity and equity as fundamental tenets in rate making. Customers have made decisions to purchase capital goods (appliances in the case of residential customers, process equipment in the case of industrial customers) and have established usage based on historical rate structures. As a result, they have a right to expect that modifications in the rate structure will not be made in such a way as to render these decisions un-economic in short order. The design of an acceptable time-of-use rate based on marginal costs must, therefore, satisfy three basic goals: (1) signal to consumers the additional cost that is incurred or saved as the result of a decision to increase or decrease usage (marginal cost); (2) provide the utility with the allowed revenue levels; and (3) meet the regulatory and legal precepts of rate making.

Satisfying these three objectives simultaneously makes the design of time-of-use rates very difficult. However, these objectives are the necessary and sufficient conditions in order for a marginal cost-based rate design to be acceptable, although the emphasis placed on each may vary depending upon the analyst's viewpoint.

There are several methods of adjusting marginal costs to the desired class revenue targets. Some of these are: (1) lump-sum transfer; (2) inverse elasticity rule (both theoretical and practical applications); (3) proportional adjustment; and (4) Ontario Hydro's large user pricing rule. Gordian believes there is no one approach that satisfies all the considerations (some of which are conflicting) involved in rate making. The particular method chosen depends upon such factors as sophistication of the customers in the class, size of the adjustment required, and existing rate structure.

Therefore, it is not possible to recommend any one method for all applications. Each situation must be analyzed separately. For example, when dealing with the residential class in the case of a relatively small revenue excess, it might be sufficient merely to reduce the customer cost partially in order to match the class revenue target. This might be the desirable approach because the effect on consumption decisions and utility earnings stability probably would be small. However, in the case of the industrial class, such an approach might be meaningless because of the relative insignificance of the customer cost revenues, and another method would be necessary.
Interruptible Rates

The design of interruptible rates is especially difficult. The customers most likely to be interested in such rates are large industrial users with different plant processes and different abilities to withstand interruptions of varying duration and frequency. We believe the way to maximize the use of interruptible service is to allow such customers to select the mode of interruption that most closely fulfills their needs. A matrix of credits should be developed which reflects the savings that the utility would achieve through alternative interruption schedules. This matrix reflects both a variation in the frequency (number of interruptions per year) and duration (length in hours) of each interruption.

As discussed earlier, it is highly unlikely that the reliability in the coming year and in succeeding years will exactly match the targeted optimal criterion. When designing interruptible rates in such a situation, actual savings that are likely to result from being able to interrupt customers must be taken into account.

When capacity is short, the probability of curtailment of the load to firm customers is high; hence, marginal cost pricing would indicate that a significant discount should be offered to encourage customers to shift part of their demand to an interruptible schedule. When excess capacity exists, the value of the marginal improvements to reliability that result from interruption will be low, and only a much smaller discount would be justified. The magnitude of the discount that can be offered can be estimated through determining the actual reliability level and relating it to the targeted reliability level.

Since long-term interruptible contracts are desirable, and since it is also desirable to inform existing and potential users of these rates of the future trend in the interruptible discount, these estimates should also be computed.

The credits that are calculated in this way represent the savings that the utility can achieve by not building capacity. These figures can be used in two different ways in designing an interruptible rate. The first, a long-term contract, would provide the customer with annual credits for accepting a lower quality of service (service of lower reliability). Furthermore, since the customer has allowed the utility to modify its long-run expansion plans, based on the currently perceived load forecast and plans, the customer should receive these credits regardless of what occurs in a particular year (that is, whether the system exceeds or falls short of its reliability level). While this may seem difficult to accept, it should be remembered that the reliability target is usually met in the long run, even though the level may vary in a particular year. At the same time, the customer must be prepared to agree to accept lower reliability over the long run, and if there is a change in the customer's desired reliability level, sufficient lead time must be provided for the utility to adjust its expansion plans.

As an alternative, a short-run interruptible contract can be designed. The credits received in a particular year will reflect the actual reliability level of the utility in that year. Assuming that long-term planning predictions are correct, the average credit in both cases will be the same, but in the second case the utility may have more flexibility to correct for unexpected developments. However, it may be more difficult to attract customers to such a rate since they will have no guarantee beforehand about the magnitude of the credit they will receive beyond the current year for modifying their reliability preference.

Notes
1. The "shortage" or "curtailment" cost represents the expected cost to customers of probable interruptions. It is equivalent to the concept of a cost of congestion in telephone service.
2. The economic carrying charge need not be the same as the levelized carrying charge. It represents the decline in economic value of the unit as a result of deterioration and obsolescence.
Time-of-Day Rates: Past, Present, and Future

J. R. Crespo

I am convinced that the time has come to stop debating the pros and cons of different time-of-day rate methods; I believe we must move forward in an orderly and logical manner to achieve the ultimate goal — properly designed time-of-day rates for those companies and jurisdictions where they make sense. I say this very emphatically because we now have a mandate from Congress under the Public Utility Regulatory Policies Act, Title One, Section Three (d): “The rates charged by any electric utility for providing electric service to each class of electric consumers shall be on a time-of-day basis which reflects the costs of providing electric service to such a class of electric consumers at different times of the day unless such rates are not cost-effective with respect to such class, as determined under Section 115 (b).”

Of course, Section 115 (b) states that time-of-day rates should only be implemented in cases in which the benefits to be derived are likely to exceed the metering and other costs associated with the use of such rates. In essence, what I would like to do is discuss some ideas and issues which I think must be addressed realistically in order to carry out this mandate. I will make brief reference to what we have done and what we are doing. It would not be accurate to say I will discuss what we should all be doing in the future. What I intend is an exchange of ideas which may provide some insight into how to implement time-of-day rates.

The Public Utility Regulatory Policies Act, in its present format, is subject to some interpretation. However, two things are clear: The act mandates neither marginal nor embedded cost. Furthermore, the act is not specific as to how rates should be designed in order to achieve time differentiation.

In my recent travels, two things have left a deep impression on me. One is a little plaque I saw in the office of a rate analyst of a midwestern utility: “There’s No Limit To What Can Be Achieved If We Do Not Care Who Gets Credit For It.” I think it has great pertinence and significance to the issues we are facing as well as the message I would like to convey. The other thing that has impressed me is the great uncertainty, doubt, and frustration felt not only by regulators, but also by those who they regulate. Without doubt we are at a crucial crossroad in the entire regulatory process. I believe that consumers, regulators, and utilities are all weary of having different pet theories advocated as if they were the ultimate answer.

The idea of peak pricing and costing, and time-of-day pricing in particular, is getting to be an old story. So are the arguments on how to do it. We have been reminded on several occasions that some of the ideas date back to 1926, and even as early as the 1890s. Believe it or not, some of them still apply today. A review of history shows the efforts to specify the true objectives of time-of-day rates along with other pricing methods, and yet no one has specified a better list of objectives than James Bonbright. Today, as we deal with the implementation of the Public Utility Regulatory Policies Act, the list is somewhat shorter. The objectives of efficiency, equity, and conservation mirror the primary three of Bonbright’s eight criteria. But the list of alternative definitions is as long as ever, and the industry is still trying to define those objectives.

The fact is, we are still seeking to derive the “best” costing and pricing methods to meet these objectives. That is especially true for time-of-day rates. There are many approaches and theories from varied disciplines, and they all have advantages and disadvantages.

We must be prepared to deal with two very pragmatic issues: (1) How do we include in rate base the large and expensive nuclear or fossil power plants that are coming on line within the next five years? (2) How do we resolve the political forces that come into play in the area of rate design? I think these are two of the main issues that are troubling those who must work in the area of regulation today. There are no simple
answers, there are no clear road maps. However, we must strive to
achieve workable and intelligent solutions. One lesson from the past
should be clear and should guide us in the future: The real answers to
these issues will not be developed in New York, Washington, or Palo
Alto! They will evolve from the utilities themselves and at the local
regulatory jurisdiction level.

Let us take a slight detour and allow our imaginations to run wild.
Let us imagine that the Battlestar Galactica, which is rapidly
approaching its final destination — Earth — has sent out a scout. We might
believe in the content of this report. Earth is locked in some kind
of religious war between three opposing sects: accountants, engineers,
and economists. These sects have somehow managed to subdivide
themselves into two theologies known as marginal and embedded.
They direct broadsides at one another: Embedded costing is an
approach which is hopelessly unrealistic. It was born in the Stone Age and
is inapplicable to the problem at hand; all an embedded study can do is
allocate. It is an accounting tool and useless in setting a price; it is time
to replace the rate engineer with the rate economist, or at least an
engineering economist; there are as many ways to allocate cost as there
are stars in the universe. The scout could compound confusion by also
saying that there are nearly as many ways to do marginal studies as
there are allocation methods, and they all generate the wrong answer;
choosing from among the marginal approaches is like choosing be
tween liquor amends and sin; and the newest part about economics then is
that it is not all cluttered up with facts and numbers, so one can use it to
support any position necessary or beneficial.

Finally, our visitor could further report that, contrary to the brief-
ing he received before leaving the Battlestar, not all accountants wear
green eye shades and count beans, not all engineers wear white socks
and have calculators hanging from their belts, and not all economists
have beards, wear jeans, hang beads around their necks, and think
weird thoughts. As a matter of fact, his observation might well be
that all three groups are talented and have something to offer. Neverthe-
less, his recommendation surely would be that one function the galactic-
cians would not want to assume upon landing on earth is that of the
regulators.

Perhaps, with a bit of introspection, we might agree on the follow-
ing principles: (1) Only the most rigid adherents to accounting princi-
pies would deny the need to look at more than a single year of historical
data; (2) only the most sublime of engineers would contest that just and
reasonable rate design does not end with the specifications of units and
unit costs calculated to achieve what he thinks is desirable; and (3) only
the most innocent of economists would deny that not all problems are
solved by inspecting a few heat rates, plugging into some computer
model, and leaving town. These approaches must be reconciled if the
regulator is to make the correct decision. And the reconciliation should
be as straightforward as possible.

The 9 November 1978 issue of the New York Times carried an
interview with Herbert A. Simon of Carnegie-Mellon University, Si-
mon, winner of the Nobel Prize in Economics, asserted that "the eco-

J. R. Crespo

474
fore, the peak is served from a mix of both peaking capacity and base load capacity. But the peak alone did not cause the investment in base capacity, so it should not be costed using only the cost of a peaker.

To continue with this logic, the peak responsibility demand model could be said to be an inadequate specification of cost causation. The proper specification would be that total production costs are a function of both peak demand and load factor. The model could be specified by using the cost of a gas-turbine for the peak and considering all the rest to be energy related.

A closer examination of the two approaches reveals that neither is very satisfactory. If, by some magic, we were able to specify an all-inclusive model that explained all the variations in the cost of production, the independent variables might include load, size, power, capital cost, swings, regulatory constraints, physical characteristics of generation, and a very substantial "dummy" variable for expected future conditions and management judgment. One example of the latter is precisely the specification of peaker cost. There is such a thing as a management decision to avoid construction of peakers for good and valid reasons which are unrelated to load.

Nevertheless, I think that the introduction of a load factor into the model may help. It at least answers this question: What is the cost of meeting a peak load for one hour of the year, and turning off the switch for all other hours? The answer the argument gives is the capital cost of the peaker plus the running cost for one hour. However, rate payers in systems which avoid the simple expedient of peaker construction must apparently pay through the energy charge for this decision — whether the reason is fuel diversification, reliability, economies of scale, or load factor.

So the argument comes down to the introduction of the minimum system concept into the process of classification and allocation of capacity costs. But, as in the minimum system approach to distribution costing, there remains the problem of specification. The same approach in distribution can lead to many answers, including the "wet-string" answer which associates almost all distribution cost with either demand or energy use. But if that is used as the cost causation factor for investment in distribution, there is a problem: All kilowatt-hours are assessed with the opportunity cost of high voltage distribution, but perhaps the only displaced kilowatt-hour which can realize the opportunity cost is the last one — the peak.

The production question is similar, but more complicated. The two-variable (peak and energy) concept outlined above is a simple production function. It would apply to a simple load duration curve like the one shown in Figure 2.

Figures 1 and 2 together show why the argument is itself too simple for cost allocation. We are not facing a base versus peaker situation, or a flat duration curve. What we are facing is a constrained production function. In other words, we confront the same framework as in marginal costing. Must we use marginal costing at this point? If we do, we still have the same problems we have always had. First, most marginal cost answers submerge capital costs into energy charges. Capital cost by itself is important to us. Rate of return measures by class may be more essential, not less essential, in the future. Second, the
revenue and earnings instability problem is still with us, because the concept of demand is submerged as well. When the time comes to make a rate, it is sometimes hard enough to incorporate demand costs in the energy charge. It is asking a good deal of the regulator to order the reverse, or the alternative, which is to ratchet energy use. Third, as always, there is the excess revenue problem.

![Diagram](image)

**Figure 2.**

Some advocates address these problems by simply allocating the embedded costs, demand and energy, following the marginal cost answers. But the results can cause other problems, such as unit energy costs below variable costs. Why not maintain the accounting separation of fixed and variable costs and allocate them separately? At the expense of having some skeptics accuse me of beating an old horse, I must say that the so-called base, intermediate, and peak (BIP) method presented during the course of the earlier EPRI studies would also solve the problem.

Where does all this lead? First, let us recognize that the objective is to design time-of-day rates that will stand the test of stability, reliability, and ease of administration. Furthermore, unless we can find a way to bring together and reconcile economic and accounting concepts, we will never be able to achieve this goal. Let us also recognize that both marginal and embedded costs are not exact and are not perfect, but that each of them may offer some significant guideline or signal useful in achieving our goal in the rate design area. Therefore, I would like to offer several steps or suggestions for consideration. (1) Do not underestimate the role each person can play and the contribution each can make; let us stop the rhetoric and get to work! (2) Define and outline those future costs which will have a significant impact in each company's rate base and operating income. (3) Carefully consider all the constraints that must be taken into account regarding the nature and the operation of future plant additions. (4) Decide what is the most efficient method of obtaining the required revenues for each company in order to achieve not only the familiar rate design objectives, but also the goal of doing what will be most beneficial to the rate payer in the long run.

These four steps reflect issues which require not only measurement, but also a philosophical and logical approach.

Most important, we must squarely address the issue that the very process of regulation is probably the single most important element of the system and therefore must be part of the solution. We must keep in mind that the regulator needs comprehension of our methods. We must also remember that regulators, as well as consumers and utility management, need to know the facts. And those facts must be established in an orderly and clear manner. Just as regulators need facts so that they can arrive at sound decisions, utility managements obviously need the facts, because their decisions affect the future of stockholders and consumers. Finally, rate payers need the facts, not only because they are entitled to them, but also because they must come away with the absolute assurance that the entire process is fair: that there are no gimmicks, that there are no subsidies, and that no one is being ripped off.

In other words, whichever method is used, whichever pet theory is endorsed, it had better work! It had better work not only in theory and on paper, but also in reality. Because after that rate case is over, and we all go home or leave town, the rate payers and the stockholders must live with those results. And they had better be good!
The four papers make reference to ratemaking objectives and the best choice of time-differentiated ratemaking methodology given those objectives. Depending on which objective is stressed, one or another method will be preferable. I simply want to note that those who are prospective buyers of one of the four methodologies will have to be very careful to specify their objectives and analyze which of the methods best fits their needs — or which is the best point of departure.

One shared objective seems apparent in the papers of Mahoney, Leininger, and Crespo, namely, to comply with the National Energy Act. All say "let's go" on time-differentiated rates. The papers vary in how strongly they assert they have the answer to the complex question of designing time-differentiated rates. I should note that a lack of assertiveness with respect to the rightness of one's own answers is attractive in some respects. However, I must add that an unassertive posture, that is, that all of the methods provide "useful guidance" or are "helpful," is unattractive in another respect: Such a position does not solve the problem of exactly what should be done. Therefore, in what follows I want to discuss how well the various methodologies meet several ratemaking criteria I believe to be important.

Accounting or Embedded Cost Methods

The main virtue of any of the accounting cost methodologies, such as Crespo's, is their compatibility with the regulatory objective of fair return on historical or accounting value. A second virtue, I suppose, is their familiarity and understandability to accountants.

The main defect of accounting approaches is their failure to give companies, customers, and regulators a good sense of where costs are going and what types of consumption cause those cost increases. The methods are inherently backward looking. Even if one uses future test years, one is merely looking backward from a year or two in the future. Therefore, the method may be inconsistent with James Bonbright's efficient use objective. It remains to be seen whether it is inconsistent with the definition of the Public Utility Regulatory Policy Act (PURPA). It certainly seems inconsistent with the objective of giving customers good signals with respect to prospective changes in future costs.

A second flaw in accounting methodologies is their lack of clear, logical underpinnings. As an example, various ad hoc approaches to assigning costs to customers or to rating periods exist — Gorden Corey identified 29 competing methods years ago, and Ebasco recently invented yet another, but there is no clear, logical basis for selecting any one.
Marginal Cost

The marginal cost methodologies are almost uniformly strong against objectives in areas in which accounting methodologies are weak, and weak where accounting is strong. Briefly, marginal cost methodologies are of no help in determining revenue requirements. Marginal costing methodologies also tend to be deficient in their understandability, at least to noneconomists. Sometimes I doubt that understandability is even a criterion. The Tax Reform Act of 1976 has sometimes been labeled "the lawyers' and accountants' relief act"; the National Energy Act should go down in history as "the economists' relief act." However, the marginal costing methodologies are arguably superior in giving relevant price signals to consumers, and although the various marginal costing methodologies differ somewhat, they all have their basis in economic logic.

NERA

The NERA approach, outlined by Mahoney, is the first major commercial attempt to disseminate some powerful ideas. I would characterize the NERA approach as a conceptual framework for thinking about how to go from utility system plans to rates that reflect those plans. As a conceptual framework, it lacks the elaborate machinery of the Ernst & Ernst and Gordian approaches discussed by Leiningher and Schoor, but NERA's selling point may be that a knowledgeable and mature analyst may be better off in some cases without all the complex machinery. In the present stage of development of marginal costing methods, judgmental decisions are unavoidable in going from marginal costs to rates. And, in such circumstances, highly sophisticated marginal costs may be more of a distraction than an advantage. Their calculation is almost certainly more expensive.

Gordian and Ernst & Ernst

Both the Gordian (Schoor) and Ernst & Ernst (Leiningher) methodologies are evolving attempts to develop an elaborate computational methodology. Both are moving from theory to problems of implementation. Both are computing, or claiming to compute, short-run marginal costs, although I am not convinced Ernst & Ernst can do it.

The Ernst & Ernst methodology for computing marginal costs has several problems worth noting. One is its lack of documentation. In part for reasons of proprietary interests, in part for reasons of pace of development, it is hard to know exactly what is in the Ernst & Ernst methodology. It may, therefore, be difficult to use in an actual rate case.

As far as can be discerned, the Ernst & Ernst model focuses on long-range marginal costs. If so, this focus is potentially a problem for those companies which have current capacity configurations that they wish they did not have, that is, virtually all the companies with oil-fired baseload capacity. While Ernst & Ernst claims to be able to compute short-run marginal costs, this claim is difficult to accept because the model seems to have no dispatch or reliability modules.

Operationally, because the Ernst & Ernst model is not an optimizing one, it has to be used by someone with system planning knowledge. Moreover, the model makes some simplifying assumptions about costs that could adversely affect the quality of its results.

Nevertheless, the Ernst & Ernst approach is interesting in that it directly computes marginal costs by time period. This contrasts with NERA's method of allocating costs to time periods on the basis of loss-of-load probabilities.

The Gordian methodology can also be faulted on its documentation. This flaw may soon be corrected; Gordian evidently plans to divulge the contents of its model to clients in the relatively near future.

Because the starting point of the Gordian model is what appears to be a good engineering capacity expansion model, including an elaborate load reliability and dispatch module, Gordian probably arrives at system plans much like those of company planners. As would also be true of the Ernst & Ernst model, it would be interesting to know how well the model results correspond with those of company planners.

The criterion of understandability is surely bent, if not broken, by the Gordian model. It is complex. Nonetheless, it does have to rely on a number of simplified assumptions and on inputs from a knowledgeable user. Although some at Gordian would disagree, I also believe that the model employs a rather ad hoc allocation of capacity charges to rating periods.

Integration

In closing, I have five summary comments concerning the four papers. First, the regulatory objectives related to total revenues will necessitate at least some use of accounting cost methodologies. Perhaps I should say that I hope such is the case; I am not prepared for the Turvey prescription of prices equal to marginal costs and nationalization.

Second, the objective of informing customers on the future direction of costs and rates will be best served, I believe, by paying atten-
tion to marginal costs. The relative emphasis on accounting and marginal costs is a matter of choice — an important choice, to be sure.

Third, the rate debate has dragged on so long that we have all become much smarter about both theory and practice, and we have some reasonable — if imperfect — tools for tackling time-of-day costing and marginal costing.

Fourth, perhaps implicit in the foregoing, ratemaking remains a matter of art, not science.

Fifth, important questions remain in several areas: (1) What are the benefits versus the costs and risks of fundamentally different rate structures? (2) What should the pace of innovation be? (3) Who should prescribe that pace?

Comments

William H. Melody

Marginal cost is a concept for all seasons and all approaches. Since, in theory, there is an infinite number of marginal costs, there must be one for every cost calculation possible. But the purpose of the exercise is not to do a cost calculation and then find a marginal cost definition that will support it. The purpose is to define the appropriate marginal cost and then measure it correctly. Ralph Turvey has observed: "It cannot be emphasized too strongly that any estimate of long-term marginal cost has no significance in abstracto but only in relation to a specified load increment. There are as many marginal costs as there are conceivable load increments." Presented in the papers by Les Mahoney, William Leininger, Alan Schoor, and J. R. Crespo are four methods, from approximately ten or twelve that have been developed to date, of implementing marginal cost in the electric utility industry. A comparative overview of the four approaches is instructive. Although the implementers have varying degrees of allegiance to the neoclassical welfare theory from which the marginal cost concept is derived, each employs marginal cost for the stated purpose of improving the efficiency of resource allocation. The intent is to give consumers the correct price signals so that their consumption decisions will lead to the most efficient allocation of resources.
The theoretical deficiencies of marginal cost are by now well known and need not be repeated here. We can concentrate on implementation and look for indications as to whether resource allocation is likely to be improved. In terms of the basic general cost concept that is being implemented, NERA and Ernst & Ernst say they are implementing long-run marginal cost (LRMC), although they have adopted different approaches. Gordan says it is using short-run marginal cost (SRMC), but its approach is the same as NERA's. Elbasco is using a blend of SRMC, LRMC, and embedded cost. For each of the four approaches, the rating periods selected differ as each has made different judgments. The methods of calculating marginal energy costs, marginal demand costs, and other marginal costs differ in each study, and the data used to calculate each are drawn from different sources and incorporated in different ways into the calculations.

Even a casual review of the four methods provides ample support for the conclusion:

In the final analysis, it must be recognized that marginal costs...are entirely bound up in the personal judgment of the analyst, the "arbitrary" decisions that he makes in defining his problem, and his changing expectations. In fact, we should expect that the optimists in a firm could have quite different marginal costs than the pessimists. Marginal costs cannot be viewed as facts for which objectively valid values can be known. The marginal cost values will depend upon who applies the theory and the incentives under which he works in the environment in which it will be applied.

In addition, there are conceptual difficulties. As Turvey points out, any marginal cost calculation is only relevant to a specified output increment. For pricing purposes, that increment is determined by the alternative prices being considered. In Figure 1, if the utility is considering raising prices from $P_1$ to $P_2$, it is expected that output would decline from $Q_1$ to $Q_2$. The marginal or incremental cost of the $A$-$B$ segment of the marginal cost function is associated only with the output increment $Q_2 - Q_1$. This output increment is associated only with the price change from $P_1$ to $P_2$. A price change from $P_1$ to $P_4$ would yield a different output increment and a different marginal cost. Unfortunately, none of the four approaches attaches its marginal cost estimates to specific output increments. Rather, the marginal costs apply generally to all kinds of output increments that, for the most part, are not even precisely described.

Furthermore, none of the output increments is attached to any specific price changes. It appears that the marginal cost calculations are general purpose, applicable to almost any increments in output and any pricing alternatives that would be considered. This can only make sense if marginal cost is equal to average cost over the range of possible changes in output. Thus, although the cost exercises may have started out attempting to calculate marginal costs, what has been calculated is some hypothetical average cost that is then applied as one would apply an average cost calculation. Of course, this hypothetical average cost need have no necessary relationship to the utility's actual average cost.

There is also confusion between output increments and simple growth in output over time. The economic concept of marginal cost is the difference between the cost of supplying one output level and the cost of supplying a different output level at the same point in time. In Figure 1, it is measured by the $A$-$B$ portion of the marginal cost curve. In Figure 2, the marginal cost of $Q_2 - Q_1$ increment for 1978 is also the $A$-$B$ portion of the marginal cost curve. In each case, the utility cannot supply both $Q_2$ and $Q_1$. It can supply only one output level.

However, incorporated in the different approaches to measuring marginal costs are costs associated with the growth of output over time. A growth increment is illustrated in Figure 2 as the difference between $Q_2$ and $Q_1$, resulting in an increase in average cost from $C$ to $D$ between 1978 and 1980. The growth increment has no relationship whatsoever to the marginal costs of economic theory. This mixing of concepts explains how the supposedly similar marginal cost ap-
approaches can calculate marginal costs on some occasions that are substantially below average cost and, on other occasions, that are substantially above average cost.

This again illustrates the elusive character of marginal cost. With average cost or fully distributed cost studies based on actual incurred costs, there can be disagreement over the methods of allocation, but not over the basic cost values to be allocated. With marginal cost, there can be disagreement over the methods of allocation and over the cost valuations to be adopted for allocation. Under marginal cost, one begins and ends with hypothetical cost valuations.

Of equal significance to the costing methodology is how the marginal costs are used to set prices. It is the prices to which consumers must react, not the marginal costs. Accordingly to welfare theory, optimal resource allocation will occur only if all prices are set simultaneously equal to their respective marginal costs. None of the four approaches proposes to set prices equal to marginal cost. All approaches would have prices deviate from marginal costs so that the traditional revenue requirement, calculated on the basis of actual incurred costs, is realized.

Each approach adopts different rules for calculating the appropriate deviation of prices from marginal cost. In fact, within each approach, each firm has several options for determining the appropriate deviations. The constraint is simply that the revenues should equal the overall revenue requirement. None proposes to employ the inverse elasticity, or Ramsey price, rule that W. G. Baumol, Robert Willig, and others argue is necessary for optimal resource allocation under a revenue requirement constraint. The criterion of optimal resource allocation does not seem to be important as a basis for selecting the method for determining price deviations from marginal cost. Apparently, the criterion is some undefined basis of judgment.

Finally, let us examine how these rather imperfectly determined prices are expected to affect resource allocation. Is the general approach employed in these studies expected to change resource allocation between the electric utility industry and others in the economy? No! The traditional return on investment standard, calculated on the basis of actual costs as reported in financial statements, will continue to be the standard for attracting capital. If marginal cost were to be used for this purpose, it would have to be employed in calculating the revenue requirement. Such a proposal has not been made.

Is this general approach expected to change resource allocation among firms within the electric utility industry? No! Once again, the return on investment standard will continue to be employed. Thus, the adoption of marginal cost as a basis for pricing is not directed to have any effect on the fundamental allocation of economic resources.

Is this general approach expected to change resource allocation among classes of service within the utility? Under some approaches, it provides a basis for discriminating among service classes. It certainly will affect the allocation of revenue requirements and therefore the rate levels of different classes of service. But apparently, this is not its major purpose.

The major intent is to provide a basis for the employment of time-of-day, peak-off-peak rates within utility rate structures. But a change in time-of-day rates does not require marginal cost. Peak-off-peak pricing methods have been used since the earliest days of the industry. If the time-of-day refinement appears promising, it certainly should be evaluated. But it makes much more sense to consider time-of-day pricing by examining the cost variability of actual costs, and short-term forecast costs, than it does to go through these elaborate metaphysical exercises.

In the final analysis, then, as the electric utility industry attempts to chart its way through the uncertain waters of the future, what is the contribution of marginal cost? We have a theory of resource allocation
directed not to guiding the industry along an uncertain course, but to optimally rearranging the deck chairs on the Titanic. The implementation, as exhibited by the four approaches described here, can only be interpreted as unnecessarily elaborate and hypothetical calculation exercises. Each approach genuflects in the direction of marginal cost and then marches to its own hypothetical drummer.

PURPA has legislated economists into a major role in the electric utility costing exercise. This may be good for the welfare of economists, but these theories and methods would never survive a free market test. The allocation of resources here is certainly not socially efficient.

Notes
3. Ibid., p. 296.

Comments

Robert S. Gay

Costing methodologies for various rate forms is a very timely subject in view of the fact that the National Energy Act has recently been passed. Based on the information that I have gathered from these papers and based on my own interpretation of the National Energy Act, it appears that this timely discussion of cost methodologies is extremely important.

I believe the National Energy Act is a first step toward making puppets out of the state commissions. In other words, unless the state commissions acquiesce, future legislation will remove all of their responsibilities and concentrate them in federal agencies. I hope these events will not occur. I believe that the efforts of the National Association of Regulatory Utility Commissioners in the national rate study, which forms the basis for the papers under discussion, should show everyone that the state commissions are entirely competent to make individual decisions on rates as they affect their particular state.

Leo T. Mahoney's paper outlines his approach to determining marginal cost. His method has many advantages, but a main problem that I see is the use of loss of load probability (LOLP) for determining the on-peak seasons and base months seasons. Over the years I have
come to the conclusion that LOLP is not a very good way to determine seasonal on-peak and off-peak periods. Primarily because companies often underestimate their units in off-peak periods, it is possible that the highest LOLP occurs during a so-called low peak period month.

Maintenance scheduling also affects marginal running costs. Since maintenance is performed during the off-peak seasonal period, the replacement of base load generation by more expensive generation can result in marginal energy costs in the off-peak seasonal period which are greater than those of the on-peak period. While it may be true that from the viewpoint of the system planner, this maintenance schedule minimizes total costs, it should be noted that we are concerned with determining marginal costs and charging the correct price. The maintenance requirement is a function of generator use and is properly chargeable to the period in which it occurs. Therefore, to incur the cost in one period and defer it to another in the form of higher running costs may cause some significant problems.

To a substantial degree, part of the problem seems to arise from overemphasizing the viewpoint of the system planner, whose goals are felt to come close to satisfying the criteria of the economic theoreti-
cians. However, the planner is primarily concerned about supply considerations, and although he must meet the seasonal demand vari-
ations of the customers in a reliable, least-cost way, his job ends there. The planner will attempt to provide a constant system reliability throughout the year, although he may try to provide relatively higher reliability during peak months.

There is nothing basically incorrect about viewing things through the system planner's eyes. However, the marginal cost analyst is not relieved of his responsibility to evaluate critically the overall situation from both the supply and demand side.

Therefore, it is left to costing and rate design specialists to devise prices which properly reflect costs imposed on the system. It is the area of economics that customer response (which should not be possible) shortage costs. Since these are central to the economic theory of marginal costing, there is a need to incorporate this concept more completely into cost analysis. The shortage of data in this area indicates a need for additional research.

William J. Leininger claims that his is a pure economic approach to marginal costing. This claim notwithstanding, the approach evokes thoughts that range from puzzling to confusing. In responding to the criticism that the approach ignores reality, Leininger states: "If you define reality as the existing system, the statement is true." Earlier in the paper, a standard definition in economics is used: "The long run is that period of time within which all inputs can be changed." Very basically, all inputs could require from 30 to 40 years (or more) to change in the electric utility industry. If so, the 20-year period to which Leininger refers would be intermediate, rather than long run.

To ignore the reality of the existing system is one thing, but to look even 20 years into an uncertain future using today's technology seems highly questionable. It is understood that the long run is not defined by the passage of time, but, in the economic sense, according to input variability. Nevertheless, 20 years is a long time and may be too long to be considered relevant to setting prices in today's world, except in a purely conjectural sense.

Leininger's production function approach utilizing different fuel conversion technologies, based on data from generating units across the country, appears overly simplistic, and one can only wonder about its appropriateness.

All things considered, the methodology is difficult to follow. Although primarily covering long-run marginal costs, the following statement is confusing: "If the economists' marginal approach is selected as the basis for time-of-use pricing, then we favor the short-
run marginal cost basis over the long run because we do not have the world of perfect competition . . . . This preference may be tempered by the capacity position of the utility." The "capacity position" which led to the selection of long-run marginal costs must obviously be the controlling case since short-run marginal costs are so seldom mentioned.

Peter F. Drucker has stated in his book "In Search of Excellence" that it pays so much attention to the details of long-run costs without comment-
ing on the possible problems involved in pricing based on long-run versus short-run costs. In general, the mixture of theory and the "magic wand" universal generation system places the methodology on a somewhat contradictory basis.

Alan Schoor discussed his methodology, which differs from others by its use of a generation optimization model, although the use of the system planner's perspective and loss of load probability is very similar to that of the National Economic Research Associates. Concerning these two points and the need to specify shortage costs, the comments which were made about Mahoney's paper apply equally to the Gordian methodology.

J. R. Crespo's paper contains the premise that including expensive power plants in the rate base and resolving political forces in rate design "are two of the main issues that are troubling those who must work in the area of regulation today." Whether these two issues are relevant to everyone is a matter of personal opinion. However, the
remark illustrates the point that the primary theme of this presentation revolved around regulation and was not concerned very much with time-of-day rates. This, more than anything else, is probably most indicative of EBASCO's philosophy, which appears concerned with merely reacting to regulation rather than taking an active role in changing it. Observe the number of times the words regulation and regulator were used in Crespo's paper. The impression that emerges is that EBASCO's entire philosophical attitude lies within the limited bounds of conforming to rules laid down by regulators.

The foregoing statements are not to be construed to mean that embedded costs are wrong. They have been used in the past and will undoubtedly be used in the future, regardless of the final disposition of the marginal cost issue.

Generally speaking, one of the conclusions which can be reached is that although much has been accomplished so far in the debate on different costing methodologies for time-of-use rates, there are many questions to be answered. The debate on these unsettled issues may be clouded by the sometimes unclear provisions of the Public Utility Regulatory Policy Act (PURPA), which may raise new controversies and possibly obscure some of the progress which has been achieved to date. I hope that some of the more optimistic views of PURPA are well founded.
Diffusion of Central Station Solar Power

David A. Huettner

Any economic assessment of central station solar electric power is complicated by the fact that solar plants will not be operated in isolation but must be integrated into a utility's system of plants. The economic decision to adopt such power, therefore, will depend on the costs and operating characteristics of solar plants and of the system into which they are integrated.

While plant and system level factors are not easily separated, sufficient analysis has been published by this author and others to allow assessment of the economic potential of central station solar power at both the plant and system level.1 The plant level analysis will be presented in the next section. The following section will present the system level analysis. The final section will summarize and present conclusions regarding the potential diffusion of central station solar power.

Plant Level Assessment

The broadest economic framework for assessing solar plants is a comparison of the costs and benefits of solar and nonsolar alternatives. Since a study of this type has recently been completed, a summary of
that analysis will be presented. The major economic characteristics of solar plants will also be discussed before turning to a system level analysis.

Comparison of Solar and Nonsolar Options

The study summarized herein compared several solar alternatives available to utilities in the southwestern United States. Wind and geothermal energy were investigated but were not commercially viable at the plant site (Hobb's, New Mexico) selected for analysis. The 32

Table 1. Alternatives for Generating Power for Hobb's, New Mexico

<table>
<thead>
<tr>
<th>Options</th>
<th>LCOE factor (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central Receiver</td>
<td>20</td>
</tr>
<tr>
<td>Central Receiver</td>
<td>15</td>
</tr>
<tr>
<td>Central Receiver</td>
<td>10</td>
</tr>
<tr>
<td>Cylindrical Trough</td>
<td>10</td>
</tr>
<tr>
<td>Cylindrical Trough</td>
<td>25</td>
</tr>
<tr>
<td>Cylindrical Trough</td>
<td>30</td>
</tr>
<tr>
<td>Parabolic Dish Brayton System A</td>
<td>20</td>
</tr>
<tr>
<td>Parabolic Dish Brayton System A</td>
<td>25</td>
</tr>
<tr>
<td>Parabolic Dish Brayton System A</td>
<td>35</td>
</tr>
<tr>
<td>Fixed Mirror Distributive Focus</td>
<td>50</td>
</tr>
<tr>
<td>Fixed Mirror Distributive Focus</td>
<td>50</td>
</tr>
<tr>
<td>Dispersed Photovoltaic-Present Cost</td>
<td>10</td>
</tr>
<tr>
<td>Dispersed Photovoltaic-Present Cost</td>
<td>10</td>
</tr>
<tr>
<td>Dispersed Photovoltaic-Present Cost</td>
<td>10</td>
</tr>
<tr>
<td>Dispersed Photovoltaic-ERDA Cost</td>
<td>25</td>
</tr>
<tr>
<td>Dispersed Photovoltaic-ERDA Cost</td>
<td>30</td>
</tr>
<tr>
<td>Parabolic Dish Non-Brayton</td>
<td>20</td>
</tr>
<tr>
<td>Parabolic Dish Non-Brayton</td>
<td>25</td>
</tr>
<tr>
<td>Parabolic Dish Non-Brayton</td>
<td>30</td>
</tr>
<tr>
<td>Imported Oil</td>
<td>75</td>
</tr>
<tr>
<td>Gasified Coal</td>
<td>75</td>
</tr>
<tr>
<td>Low Sulfur Coal</td>
<td>75</td>
</tr>
<tr>
<td>Nuclear A — Burn-up Rate 53,000; Pu Recycle Cr., 0</td>
<td>60</td>
</tr>
<tr>
<td>Nuclear A — Burn-up Rate 53,000; Pu Recycle Cr., 96 mili/kwh</td>
<td>60</td>
</tr>
<tr>
<td>Nuclear A — Burn-up Rate 53,000; Pu Recycle Cr., 96 mili/kwh</td>
<td>60</td>
</tr>
<tr>
<td>Nuclear B — Burn-up Rate 53,000; Pu Recycle Cr., 0</td>
<td>75</td>
</tr>
<tr>
<td>Nuclear B — Burn-up Rate 53,000; Pu Recycle Cr., 96 mili/kwh</td>
<td>75</td>
</tr>
<tr>
<td>Nuclear B — Burn-up Rate 53,000; Pu Recycle Cr., 96 mili/kwh</td>
<td>75</td>
</tr>
<tr>
<td>Nuclear C — Burn-up Rate 24,000; Pu Recycle Cr., 0</td>
<td>60</td>
</tr>
<tr>
<td>Nuclear C — Burn-up Rate 24,000; Pu Recycle Cr., 96 mili/kwh</td>
<td>60</td>
</tr>
<tr>
<td>Nuclear C — Burn-up Rate 24,000; Pu Recycle Cr., 96 mili/kwh</td>
<td>60</td>
</tr>
<tr>
<td>Nuclear D — Burn-up Rate 24,000; Pu Recycle Cr., 0</td>
<td>70</td>
</tr>
<tr>
<td>Nuclear D — Burn-up Rate 24,000; Pu Recycle Cr., 96 mili/kwh</td>
<td>70</td>
</tr>
<tr>
<td>Nuclear D — Burn-up Rate 24,000; Pu Recycle Cr., 96 mili/kwh</td>
<td>70</td>
</tr>
</tbody>
</table>

NOTE: Pu = plutonium; Cr. = credit.

Table 2. Solar Options

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central receiver (Power Tower)</td>
<td>Consists of a field of heliostats all tracking the sun and focusing on a central collector at the top of a tower. A 10 Mw plant is under construction at Barlow, California. Original cost estimates of $50 million have increased to $120 million.</td>
</tr>
<tr>
<td>Parabolic dish (Brayton and non-Brayton)</td>
<td>Consists of a field of parabolic dishes all tracking the sun. Under the Brayton A option, electricity is generated by a small, Brayton cycle turbine generator associated with each parabolic dish. Under the non-Brayton option, the heat collected from each parabolic dish is pumped to a central location to generate electricity using one large rankine cycle turbine generator. JPL has just awarded contracts to develop a 1 Mw pilot plant.</td>
</tr>
<tr>
<td>Photovoltaic ($1750/m² and $500/m²)</td>
<td>This option consists of a field of flat plate photovoltaic arrays. No optical or reflecting concentrators are used, and no tracking of the sun is considered. The 3 Mw pilot plant at Mississippi College in Arkansas uses parabolic trough concentrators, and some tracking of the sun is used; hence this is not the same as this option. The current cost of the photovoltaic arrays is $1,750 per square meter, and the ERDA (now DOE) goal for this cost is $50 per square meter.</td>
</tr>
<tr>
<td>Fixed mirror distributive focus</td>
<td>This option consists of a field of hemispherical concentrators set in the ground, hence they do not track the sun. Hemispherical concentrators focus on a line, and the linear collector located at the focus tracks the sun by season. The heat collected is pumped to a central location where electricity is generated using one large rankine cycle turbine generator. A 5 Mw pilot plant is now under construction at Crosbyton, Texas.</td>
</tr>
<tr>
<td>Cylindrical trough</td>
<td>This option consists of a field of parabolic trough concentrators oriented east and west to minimize the amount of tracking of the sun required. The heat collected is pumped to a central rankine cycle turbine for electrical generation. No large-scale pilot plants using this option are under construction currently in the United States.</td>
</tr>
</tbody>
</table>

nonsolar alternatives, listed in Table 1, consisted of low sulfur coal, imported oil, gasified coal, and a pressurized water nuclear plant. These were selected because they can all be done with certainty, are technologically feasible, and are as comparable as possible to solar in terms of environmental impact. Table 2 provides a brief description of the five solar options.

Annualized, discounted, busbar costs in mills per kwh were com-
puted for each option. Mass production, \(^1\) optimal scale of production, \(^2\) and compliance with environmental laws were assumed for costing each option. Costs of energy storage and electrical distribution were not considered, but transmission costs to Hobbs were calculated if power transmission was required. Busbar energy costs were used to prevent the results from becoming overly site or utility specific. Note that busbar costs internalize the costs of plant unreliability only to the extent that plant factor is affected. Nonquantifiable costs (pollution, mortality, and so on) were not estimated due to the well-known problems involved, and the analysis, therefore, was of the cost-effectiveness rather than the cost-benefit type.

Based on the above assumptions, baseline 1980 costs were calculated for each option. Baseline construction costs were computed considering inflation and interest during construction. The total consumption cost estimate for each option was then divided by the kw of generating capacity to yield fixed costs per kw. Thirty-year plant lives and a 15 percent fixed charge rate were assumed for all options except the nuclear, for which a 15.55 percent rate was used. Up to three levels of load factor were specified for each option to allow for sensitivity analysis concerning this parameter. When combined, all these assumptions allow calculation of baseline 1980 fixed costs in mills per kwh.

Operating costs and fuel costs were also calculated for each option under baseline conditions as of 1980. The major sensitivity analysis allowed for in operating costs occurred in the nuclear option by use of two levels of load factor, two levels of burn-up rate, and two levels of plutonium recycle credit. Transmission costs were also calculated for the nonsolar options since the generating plants were assumed to be located in Amarillo, Texas.

Modification of baseline data to reflect future costs was accomplished by specifying various cost components and parameters (such as the rate of inflation) for sensitivity analysis. In particular, four scenarios were developed for parameters subject to inflation through time. Scenario I assumed that gradual technological change would offset input price increases; a zero annual inflation rate was used for all costs. Scenario II, a middle-of-the-road model, was based mainly on expert opinion about future rates of increase for a particular parameter. Scenario III assumed that recent (1970–1976) rates of inflation would continue beyond 1980. Scenario IV was actually a hybrid formed by replacing the fuel costs in Scenario I with those in Scenario III. In essence, this hybrid reflects what would happen if fuel costs inflated much more rapidly than nonfuel costs. The inflation rates of these scenarios were applied to salient parametric cost values of each of the alternatives considered. Some parameter values, such as plant factor, were, by their nature, not subject to inflation yet were allowed to take alternative values over the 1980–1995 period. Table 3 presents a summary of the parameter values subjected to sensitivity analysis, and Table 4 shows the inflation rates assumed for each scenario.

Solar plant cost estimates were obtained from two sources. One was the studies published by Monument Energy Co. (Moneco) between 1974 and 1976. The estimates typically were based on engineering knowledge at the time. The other source was recent detailed engineering analysis done for the Jet Propulsion Lab (JPL) in Pasadena, California, which indicated significantly higher costs. The higher estimates arose mainly because of a growing knowledge of the engineering complexities of solar power. Only a minor portion was attributed by JPL to inflation since the publication of the Moneco studies.

Two procedures were adopted to allow for the effects of increased engineering complexities on the costs reported by Moneco. First, the JPL cost estimates were used for comparison. Second, the raw Moneco estimates were multiplied by a factor of 1.25 and 1.50 to create two more sets of Moneco estimates for comparison with the nonsolar options. These multipliers were established by comparing the Moneco and JPL estimates when they overlapped.

Table 3. Summary of Parameter Values Subjected to Sensitivity Analysis

<table>
<thead>
<tr>
<th>Options</th>
<th>Parameter</th>
<th>Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solar</td>
<td>Discount rate (8%, 10%, 12%)</td>
<td>Load factor (29%, 25%, 30%)</td>
</tr>
<tr>
<td></td>
<td>Inflation (Scenarios I, II, III, and Hybrid)</td>
<td>Inflation (Scenarios I, II, III, and Hybrid)</td>
</tr>
<tr>
<td></td>
<td>Total costs multiplied by 1.25 and 1.50</td>
<td>Total costs multiplied by 1.25 and 1.50</td>
</tr>
<tr>
<td></td>
<td>Plant site (Hobbs, New Mexico)</td>
<td>Plant site (Hobbs, New Mexico)</td>
</tr>
<tr>
<td>Fossil</td>
<td>Discount rate (8%, 10%, 12%)</td>
<td>Load factor (75%)</td>
</tr>
<tr>
<td></td>
<td>Inflation (Scenarios I, II, III, and Hybrid)</td>
<td>Inflation (Scenarios I, II, III, and Hybrid)</td>
</tr>
<tr>
<td></td>
<td>Load factor (75%)</td>
<td>Load factor (75%)</td>
</tr>
<tr>
<td></td>
<td>Plant site (Amarillo, Texas)</td>
<td>Plant site (Amarillo, Texas)</td>
</tr>
<tr>
<td>Nuclear</td>
<td>Discount rate (8%, 10%, 12%)</td>
<td>Load factor (65%, 70%)</td>
</tr>
<tr>
<td></td>
<td>Inflation (Scenarios I, II, III, and Hybrid)</td>
<td>Inflation (Scenarios I, II, III, and Hybrid)</td>
</tr>
<tr>
<td></td>
<td>Burn-up rate (24,000 and 33,000 thermal kw</td>
<td>Burn-up rate (24,000 and 33,000 thermal kw)</td>
</tr>
<tr>
<td></td>
<td>days/year) (voided uranium)</td>
<td>(voided uranium)</td>
</tr>
<tr>
<td></td>
<td>Plutonium recycle credit (0.6 and 0.96 mills/Akw)</td>
<td>Plutonium recycle credit (0.6 and 0.96 mills/Akw)</td>
</tr>
<tr>
<td></td>
<td>Plant site (Amarillo, Texas)</td>
<td>Plant site (Amarillo, Texas)</td>
</tr>
</tbody>
</table>
Table 4. Annual Inflation Rates Beyond 1980, by Scenario

<table>
<thead>
<tr>
<th>Options and cost category</th>
<th>Scenario I (percent)</th>
<th>Scenario II (percent)</th>
<th>Scenario III (percent)</th>
<th>Scenario IV (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solar</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Construction cost</td>
<td>0</td>
<td>5</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Operation and maintenance cost</td>
<td>0</td>
<td>4</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Fossil</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Construction cost</td>
<td>0</td>
<td>5</td>
<td>10*</td>
<td>0</td>
</tr>
<tr>
<td>Production cost</td>
<td>0</td>
<td>4</td>
<td>1.3</td>
<td>0</td>
</tr>
<tr>
<td>Fuel Cycle cost</td>
<td>0</td>
<td>4</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Nuclear</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Construction cost</td>
<td>0</td>
<td>5</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Production cost</td>
<td>0</td>
<td>4</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>Fuel Cycle cost</td>
<td>0</td>
<td>5</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Transmission</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Construction cost</td>
<td>0</td>
<td>5</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Operation and maintenance cost</td>
<td>0</td>
<td>4</td>
<td>6</td>
<td>0</td>
</tr>
</tbody>
</table>

*For the imported oil option, an 11 percent rate of increase of construction cost was used since this was the average for 1970–1976.

The sensitivity analysis detailed above was applied to the engineering cost data developed by Moneco and JPL. The result was a series of matrices of discounted busbar cost values. In each matrix, rows were defined by various electrical generating options, and columns were defined by time, discount rate, and scenario combinations. A rank matrix corresponding to each cost matrix also was generated, the entries designating the rank of all options on a column-by-column basis. Summaries were developed and are presented in Tables 5 through 8.

Table 5. Summary of Ranking Patterns for Raw Moneco Data

<table>
<thead>
<tr>
<th>Scenario:</th>
<th>I'</th>
<th>I'</th>
<th>I</th>
<th>I</th>
<th>I and II</th>
<th>III</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construction Date:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Option</th>
<th>1'</th>
<th>1'</th>
<th>I'</th>
<th>I'</th>
<th>I and II</th>
<th>III</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low sulfur coal</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Imported oil</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>8</td>
<td>8</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Nuclear D</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Nuclear E</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>7</td>
<td>4</td>
<td>8</td>
<td>9</td>
</tr>
</tbody>
</table>

Table 5—Continued

<table>
<thead>
<tr>
<th>Scenario:</th>
<th>I'</th>
<th>I'</th>
<th>I</th>
<th>I</th>
<th>I and II</th>
<th>III</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construction Date:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dispersed photovoltaic option</th>
<th>ERDA goal</th>
<th>Central receiver</th>
<th>Fixed mirror</th>
<th>Parabolic dish</th>
<th>Brayton A</th>
<th>Gasfied coal</th>
<th>Parabolic dish</th>
<th>Cylindrical trough</th>
<th>Dispersed photovoltaic present cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERDA goal</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>7-8</td>
<td>6-7</td>
</tr>
<tr>
<td>Central receiver</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>7-8</td>
<td>6-7</td>
</tr>
<tr>
<td>Fixed mirror</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>7-8</td>
<td>6-7</td>
<td>5-6</td>
<td>7-8</td>
<td>6-7</td>
</tr>
<tr>
<td>Parabolic dish</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>8-9</td>
<td>7-8</td>
<td>6-7</td>
<td>7-8</td>
<td>6-7</td>
</tr>
<tr>
<td>Brayton A</td>
<td>9</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>7-9</td>
<td>9-5-7</td>
<td>5-7</td>
<td>7-9</td>
<td>9-5-7</td>
</tr>
<tr>
<td>Gasfied coal</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>10</td>
<td>10</td>
<td>9-5-7</td>
<td>7-9</td>
<td>9-5-7</td>
</tr>
<tr>
<td>Cylindrical trough</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>10</td>
<td>11-11</td>
</tr>
<tr>
<td>Dispersed photovoltaic present cost</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12-12</td>
</tr>
</tbody>
</table>

Note: I' is the hybrid scenario obtained by entering the fuel costs from Scenario III into Scenario I. This table is based on a 25 percent plant factor for the solar options.

Table 6. Summary of Ranking Patterns for Moneco Data Increased by 25 Percent

<table>
<thead>
<tr>
<th>Scenario:</th>
<th>I'</th>
<th>I'</th>
<th>I'</th>
<th>I'</th>
<th>I and II</th>
<th>III</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construction Date:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Option</th>
<th>1'</th>
<th>1'</th>
<th>I'</th>
<th>I'</th>
<th>I and II</th>
<th>III</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low sulfur coal</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Imported oil</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>8</td>
<td>8</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Nuclear D</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Nuclear E</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>7</td>
<td>4</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Gasfied coal</td>
<td>9</td>
<td>9</td>
<td>11</td>
<td>11</td>
<td>5-7</td>
<td>9</td>
<td>8-5</td>
</tr>
<tr>
<td>Parabolic dish</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Cylindrical trough</td>
<td>11</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>Dispersed photovoltaic present cost</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
</tbody>
</table>

Note: I' is the hybrid scenario obtained by entering the fuel costs from Scenario III into Scenario I. This table is based on a 25 percent plant factor for the solar options.
Table 7. Summary of Ranking Patterns for Monoeo Data Increased by 50 Percent

<table>
<thead>
<tr>
<th>Scenario:</th>
<th>I'</th>
<th>I'</th>
<th>I'</th>
<th>I'</th>
<th>I and II</th>
<th>III</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discount Rate:</td>
<td>1980</td>
<td>1985</td>
<td>1990</td>
<td>1995</td>
<td>All</td>
<td>All</td>
<td>All</td>
</tr>
<tr>
<td>Construction Date:</td>
<td>1980</td>
<td>1985</td>
<td>1990</td>
<td>1995</td>
<td>All</td>
<td>All</td>
<td>All</td>
</tr>
<tr>
<td><strong>Option</strong></td>
<td><strong>Low sulfur coal</strong></td>
<td><strong>Imported oil</strong></td>
<td><strong>Nuclear D</strong></td>
<td><strong>Nuclear E</strong></td>
<td><strong>Dispersed photovoltaic</strong></td>
<td><strong>ERDA goal</strong></td>
<td><strong>Central receiver</strong></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>10</td>
<td>11</td>
<td>11</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
</tbody>
</table>

**NOTE:** I' is the hybrid scenario obtained by entering the fuel costs from Scenario II into Scenario I. Also note that this table is based on a 25 percent plant factor for the solar options.

Table 8. Summary of Ranking Patterns for JPL Data

<table>
<thead>
<tr>
<th>Scenario:</th>
<th>I'</th>
<th>I'</th>
<th>I'</th>
<th>I'</th>
<th>I and II</th>
<th>III</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discount Rate:</td>
<td>1980</td>
<td>1985</td>
<td>1990</td>
<td>1995</td>
<td>All</td>
<td>All</td>
<td>All</td>
</tr>
<tr>
<td>Construction Date:</td>
<td>1980</td>
<td>1985</td>
<td>1990</td>
<td>1995</td>
<td>All</td>
<td>All</td>
<td>All</td>
</tr>
<tr>
<td><strong>Option</strong></td>
<td><strong>Low sulfur coal</strong></td>
<td><strong>Nuclear D</strong></td>
<td><strong>Imported oil</strong></td>
<td><strong>Nuclear E</strong></td>
<td><strong>Gasified coal</strong></td>
<td><strong>Parabolic dish</strong></td>
<td><strong>Brayton A</strong></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>6</td>
<td>9</td>
<td>9</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>8</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
</tbody>
</table>

**NOTE:** I' is the hybrid scenario obtained by entering the fuel costs from Scenario II into Scenario I. Also note that this table is based on a 25 percent plant factor for the solar options.

In developing the rank summaries, several options were eliminated for simplification. Various criteria were used to delete them. The 21 solar options based on the Monoeo data were reduced to 7 by deleting two levels of utilization. Storage was not included in the analysis, and a 25 percent rate of utilization of solar capacity is about the maximum attainable in the absence of storage. Since solar costs are higher at a 20 percent level of utilization, the 25 percent cut-off retains the most favorable solar results without storage. Using this same procedure, the 12 solar options based on the JPL data were reduced to 4 (parabolic dish Brayton A, central receiver, parabolic dish non-Brayton, and cylindrical trough), all at a 25 percent level of utilization.

The method for deleting nonsolar options was similar for both the Monoeo and JPL cost comparisons. Eight different nuclear options were created by considering two levels of utilization of capacity, two levels of burn-up rates, and two levels of plutonium recycle credit. The levels of these three factors were most favorable to nuclear generation under option nuclear D and least favorable under option nuclear E. The other six nuclear options always fell between these two extremes, hence they could be deleted. All three of the fossil fuel options were retained for the cost comparisons.

Tables 5, 6, and 7 summarize the rankings for the Monoeo data. The four options of low sulfur coal, parabolic dish non-Brayton, cylindrical trough, and dispersed photovoltaic present cost were consistently ranked first, tenth, eleventh, and twelfth, respectively. The only important exception is shown in Table 5 for the hybrid scenario and the 1995 construction date, in which case two solar options rank ahead of the low sulfur coal option. The best solar option was always dispersed photovoltaic ERDA goal, followed by central receiver, fixed mirror distributive focus, parabolic dish Brayton A, parabolic dish non-Brayton, and dispersed photovoltaic present cost.

As noted earlier, the raw Monoeo data presented in Table 5 are based on studies published prior to 1975 and therefore are likely to underestimate solar station costs. Tables 6 and 7 are thus better guides. In these two tables, the best solar options never surpass the low sulfur coal option but do manage to rank second, third, or fourth for the hybrid scenario and the 1995 construction date. These results indicate that the best solar options can become competitive with the best non-solar option by the end of the century, but only under conditions most favorable to solar (that is, the hybrid scenario). Under less favorable assumptions, Scenarios I, II, and III in Tables 6 and 7, the best solar options do no better than approach competitiveness with the imported oil and nuclear D options.

Similar analyses were applied to JPL data, excepting the escalation.
of costs by 25 and 50 percent (Table 8). The low sulfur coal option ranked first in all cases, while the imported oil and nuclear options ranked second and third with one exception: under the hybrid scenario and the 1955 construction date, in which case the parabolic dish Brayton A option ranked third. For the JPL data, the best solar options never surpass the traditional options of low sulfur coal or nuclear and, surprisingly, rank worse than the imported oil option by the end of the century in Scenarios I, II, and III. The best solar option was always parabolic dish Brayton A, followed by central receiver, parabolic dish non-Brayton, and cylindrical trough.

These results indicate that the best solar options are not competitive with the best nonsolar options on a pure cost basis. The best solar options could become competitive with the imported oil option by the end of the century, but only under conditions most favorable to solar (that is, the hybrid scenario).

Table 9 summarizes the most important factors included or excluded from this analysis. One important exclusion was wet scrubbers, since it was assumed that these would not be required on a low sulfur coal plant in the Southwest. Should they be required, the costs of the low sulfur coal option would increase by 20 percent, but low-sulfur coal would remain the lowest cost option except under assumptions most favorable to solar, in which case central station solar would be preferred by 1990.

A second point to note is that the study summarized here compared solar plants to baseload nuclear and fossil plants. Intermediate daytime loads are served by more expensive fossil plants, however. A reexamination (not shown) indicates that, under assumptions most favorable to solar, intermediate daytime loads could be served by solar as early as 1985. Under less generous assumptions, 1990 or later would be a more likely date (note that 1990 is the U.S. Department of Energy goal).

This analysis indicates that daytime intermediate loads are the best target for solar electric plants on a baseload cost basis. The solar options ranked from best to worst are shown in Table 10. The central receiver, FMDF, and parabolic dish options all performed reasonably well using either the Monco or JPL data, and large-scale pilot plants for all three are currently under construction (or in the advanced planning stage in the case of parabolic dish).

**Economic Characteristics**

A solar electric plant (except for photovoltaic options) is essentially a solar boiler connected to a turbogenerator. The solar boiler or collector system for a central station solar plant has very high fixed costs (see Table 11), extremely low operating costs, and generally produces steam at temperatures ranging from 600°F to 1000°F depending on the option considered. In addition, the collector system accounts for about 80–90 percent of total solar plant costs (see Table 11), hence any major reduction in solar plant baseload costs can only be achieved by a reduction in collector costs. The major characteristic of a solar collector, however, is that it only operates while the sun shines; therefore, with or without storage, it can never be utilized more than about 30 percent of the time. Furthermore, since the high temperatures are obtained by concentrating sunlight, the output of a solar electric plant is seriously impaired on cloudy days.
Table 11. Collector Costs for Central Station Solar Electric Options

<table>
<thead>
<tr>
<th>Collector</th>
<th>1980 collector cost ($/kW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central receiver</td>
<td>579</td>
</tr>
<tr>
<td>Cylindrical trough</td>
<td>1,692</td>
</tr>
<tr>
<td>Parabolic dish non-Brayton</td>
<td>1,278</td>
</tr>
<tr>
<td>Parabolic dish Brayton A</td>
<td>816</td>
</tr>
<tr>
<td>Fixed mirror-distributive focus</td>
<td>732</td>
</tr>
</tbody>
</table>

Source: Center for Economic & Management Research, A Study of the Feasibility of Utilizing Solar, Wind and Geothermal Energy in Hobbs, New Mexico, Final Report, prepared for the U.S. Department of Energy under contract No. E-491-1-5231, June 1978, vol. 2. Note that the $570/kW figure for the central receiver option is probably 25 percent to 50 percent too low based on more recent studies. Similarly, more recent studies indicate that the $690/kW figures shown for the other options are probably 25 percent too low.

The turbogenerator portion of the solar plant accounts for 10-20 percent of the total costs, and its utilization is also limited to a maximum of about 30 percent unless energy storage or a nonsolar boiler is used. While increased utilization of the turbogenerator is desirable, a doubling would reduce total costs no more than 3-10 percent, so this is not a critical factor.

Some insight into the collector cost problem can be gained by assuming that a turbogenerator unit is in place, but a new boiler needs to be constructed. Assume also that a solar boiler has zero operating costs. One would be indifferent between a solar and fossil boiler: (assuming equal reliability and durability) if the fixed costs of the solar boiler, $F_s$, equals the fixed plus operating costs, $O_f$, of the fossil boiler:

$$ F_s = F_r + O_f. $$

Then:

$$ L_sR_s = L_fR_f + O_f. $$

Further assume that $R_s = R_f$ and $C_s = C_f$; that is, the fixed charge rates are equal, and the boilers produce equal output during the day. Since the fossil boiler can also produce power at night, $U_s > U_f$, $O_f > Q_f$. Given that there are 8,760 hours in a year, and noting that $Q = 8760$CU; we have:

$$ \frac{Q_s}{U_s} = \frac{Q_f}{U_f} \quad \text{since} \quad C_s = C_f. $$

Substituting in equation (2):

$$ I_sR_s = I_fR_f + O_f(8760). $$

Rearranging:

$$ I_s = I_f \left( \frac{U_f}{U_s} \right) + \frac{O_f(8760)}{R_s}. $$

$U_s$ is known to equal about 30 percent at maximum, and a fixed charge rate of 18 percent is applicable to both solar and fossil boilers. For a base load fossil boiler, $I_f = 430/kW$, $U_f = 75$ percent, and $O_f = 0.020$ mills/kWh. Substituting these values yields a value for $I_s$ in $$/kWh at which solar boilers would be competitive in 1980 for baseload operation:

$$ I_s = 430 \left( \frac{0.30}{0.45} \right) + 0.020 \left( \frac{0.30}{0.18} \right) (8760) = 172 \quad + 292 = 464/kW. $$

For an intermediate load fossil boiler in 1980, $U_f = 45$ percent, $O_f = 0.023$, and $I_f = 460/kW$; hence:

$$ I_s = 460 \left( \frac{0.30}{0.45} \right) + 0.023 \left( \frac{0.30}{0.18} \right) (8760) = 267 + 365 = 632/kW. $$

These calculations indicate that solar boilers would be nearly competitive in 1980 with baseload and intermediate load fossil boilers if solar collector costs were $464/kW and $632/kW, respectively. Collector cost estimates in Table 11 indicate that the underestimated Monenco fig-
ures are generally higher than even the $632/kw. Equation (5) does suggest the importance of fossil fuel costs since this is the major component of O
costs.

Continued increases in fossil fuel prices relative to other prices would substantially improve the relative position of solar plants. Resolving equations (6) and (7) for the level of $O$ that would equalize solar and fossil plant costs indicates that fossil fuel prices would have to be twice their current level for this to occur.

Summary

A review of the plant level assessment indicates that solar plant busbar cost must be reduced relative to nonsolar alternatives if solar plants are to compete for even intermediate daytime loads by the year 1990. Solar plant cost reductions of any consequence can only be achieved by lowering collector costs. This cannot be achieved by increased utilization, even if a totally efficient, zero cost storage medium were available. A perfect storage medium of this type would improve solar plant reliability but would not reduce busbar costs. Technological change is the way to reduce solar plant collector costs. The busbar costs of nonsolar alternatives will undoubtedly increase due to tighter environmental controls and/or increased fuel prices, but it is doubtful that these nonsolar cost increases could compensate for a lack of technological improvement in solar plant collectors.

System Level Assessment

At the system level of analysis, the overall objective is to match the output of the generating plants to system load at minimum cost while meeting a given standard of system reliability. The latter is generally measured in terms of loss of load probability (LOLP). Reliability is only achieved at some cost, however, and hence a plant's busbar costs are only a partial measure of the desirability of integrating a given generating plant into a system.

One way of measuring a given generating plant's impact on system reliability and costs is through capacity credit. This shows the added peak system load that could be carried with no change in system LOLP by adding the plant in question. In effect, capacity credit measures a plant's ability to match a system's load pattern. Since solar plants are useless at night or on cloudy days, capacity credit will be a direct measure of their economic desirability.

The objective of this section is to assess, at the system level, the economic desirability of solar plants. This could be done in many ways, but we will examine the potential diffusion of solar plants. Utilities not only are generating systems but also are the only major potential adopters of central station solar generating plants. Solar plants must be integrated into utility systems, and they will match the needs of some better than others. Certain utility characteristics will make solar attractive to some utilities and not others, and an understanding of these will allow better assessment of the economic potential of central station solar power.

Several hypotheses concerning the diffusion of solar power will be developed. Published system level studies of central station solar plants will then be reviewed to assess these hypotheses.

The overriding characteristic of solar plants is their dependence on sunshine and energy storage to meet loads. Two hypotheses are suggested.

H1: Utilities with summer peaks and daily peak loads occurring before 4 P.M. in the summer would find solar easier to adopt since solar output and daily peak demand coincide.

H2: Utilities with higher amounts of pumped storage in their system would find solar easier to adopt.

Simulations of a limited number of utility systems indicated support for the first hypothesis, but the effect of storage was found to be negative.

Additional hypotheses concerning the diffusion of solar plants are suggested by the high fixed costs of solar plants, their operating characteristics, and the fact that they are generally smaller in scale than traditional facilities.

H3: Municipal utilities will adopt solar plants sooner than privately owned utilities due to their lower cost of capital.

H4: Utilities in regions with high average levels of insolation will adopt solar plants sooner since busbar costs vary inversely with the average level of insolation (all else being equal).

H5: Small, isolated utilities will adopt solar plants sooner due to the high costs of traditional options at small scales and their limited access to joint ventures.

Available studies offer strong support for H3 and H4, and modest support for H5.

The remainder of this section will review the available studies bearing on each of these five hypotheses.

Hypothesis One

The strongest support for the first hypothesis comes from a simu-
loration performed in a recent study. Only solar photovoltaic options were considered (the fixed-tilt flat plate collector performed the best), but most of the conclusions are applicable to solar thermal options. The utilities were selected for study: Arizona Public Service (APS), Florida Power and Light (FPL), and New England Electric Service (NEES). All three are summer peaking systems but differ in terms of: timing of peak demand and peak insolation; average insolation and projected mix of 1995 generating capacity (see Table 12). The most important difference is that FPL’s daily peak in August does not coincide with the insolation peak, while the reverse is true for both APS and NEES. Average insolation is lowest for NEES and highest for APS, indicating that the solar plant should be utilized most for APS and least for NEES, all else being equal. As for generating mix, all three have about 45 percent in baseload nuclear or coal units; about 40 percent in intermediate coal, oil, or combined cycle units; and about 10–20 percent in combustion turbine, pumped storage, or older fossil units for peaking. The major difference is that APS and NEES have considerable pumped storage-hydro, while FPL has none.

Table 13 shows the most important results of this simulation. A solar plant can conserve scarce fossil fuels when operated in a fuel saving mode, but if it can also displace existing generating capacity, the economics become much more attractive. The first column in Table 13 shows the utilization of solar capacity achieved in these simulations. Since a solar plant cannot achieve rates of utilization beyond about 50 percent, these results indicate that the solar plants would indeed conserve traditional fuels at either 5 percent or 20 percent penetration of system capacity. The utilization is highest for APS and lowest for NEES, mainly because of the low average level of insolation in Boston and high average level in Phoenix.

The second column of Table 13 shows capacity credit, that is the extent to which solar capacity could substitute for traditional capacity with no change in system loss of load probability. Two important conclusions can be drawn from column 2. First, capacity credit can be achieved, but they are much higher when peak season solar output coincides with peak demand, for example, the APS and NEES cases. Solar plants would be adopted sooner by utilities where insolation and daily demands coincide, since both capital costs and fuel costs are saved. Utilities with daily load patterns resembling FPL’s could have trouble adopting solar plants since their high fixed costs cannot be justified by fuel saving alone or by fuel savings plus minimal capacity credit.

The second major conclusion to be drawn from column 2 of Table 13 is that the capacity credit falls rapidly with the level of penetration

<table>
<thead>
<tr>
<th></th>
<th>Arizona Public Service</th>
<th>Florida Power &amp; Light</th>
<th>New England Electric Service</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily peak insolation (August)</td>
<td>1:45 P.M.</td>
<td>12:45 P.M.</td>
<td>1:00 P.M.</td>
</tr>
<tr>
<td>Daily peak demand (August)</td>
<td>3:00 P.M.</td>
<td>5:45 P.M.</td>
<td>11:15 P.M.</td>
</tr>
<tr>
<td>Average insolation (kWh/m²/day)</td>
<td>5.81</td>
<td>5.23</td>
<td>3.49</td>
</tr>
<tr>
<td>1995 generation mix (percentage of total generation)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nuclear steam</td>
<td>26.5</td>
<td>46.8</td>
<td>43.2</td>
</tr>
<tr>
<td>Coal-fired steam</td>
<td>42.1</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Oil-fired steam</td>
<td>6.8</td>
<td>39.6</td>
<td>44.7</td>
</tr>
<tr>
<td>Combined cycle</td>
<td>3.2</td>
<td>4.1</td>
<td>0.0</td>
</tr>
<tr>
<td>Combustion turbine</td>
<td>12.9</td>
<td>9.5</td>
<td>1.2</td>
</tr>
<tr>
<td>Pumped storage hydro</td>
<td>8.1</td>
<td>0.0</td>
<td>5.5</td>
</tr>
<tr>
<td>Hydro</td>
<td>0.6</td>
<td>0.0</td>
<td>5.4</td>
</tr>
</tbody>
</table>


Table 15. Flat Plate Photovoltaic Plant Simulation Results at 5 Percent and 20 Percent Penetration of Total System Capacity, Assuming No Dedicated Solar Storage, in Percent

| Capacity factor achieved as a % of maximum possible output at rated capacity | Capacity credit
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>5 percent penetration</td>
<td>20 percent penetration</td>
</tr>
<tr>
<td>5 percent penetration</td>
<td>20 percent penetration</td>
</tr>
<tr>
<td>APS</td>
<td>25.8</td>
</tr>
<tr>
<td>FPL</td>
<td>22.4</td>
</tr>
<tr>
<td>NEES</td>
<td>17.7</td>
</tr>
</tbody>
</table>


**NOTE:** Penetration is defined as rated photovoltaic capacity as a percentage of total system rated capacity. APS = Arizona Public Service; FPL = Florida Power and Light; NEES = New England Electric Service.

**Effective capability as a percentage of rated capacity:** Capability in Mw is measured by standard loss of load probability (LOLP) calculation and shows the additional peak load that could be carried, by adding this capability, with no change in system LOLP.
(fuel savings do not, however, as shown in column 1). Solar plants, therefore, would become uneconomic fairly rapidly as the degree of penetration of a given system increased.

This conclusion is further supported by two studies utilizing Southern California Edison load information and central receiver solar plants located at various sites (see Table 14). Capacity credit falls rapidly as penetration increases in all cases. Solar plant capacity credit is higher and falls less rapidly with increased penetration if dedicated energy storage is available, but, as the discussion of H2 indicates, dedicated energy storage is not cost justified.

In any event, column 2 of Table 13 indicates that the capacity credit declines less rapidly as solar penetration increases for APS and NEEs. This suggests that utilities in areas where insolation and demand coincide could adopt solar power to a greater extent, that is, degree of penetration, and that the diffusion of central station solar power is limited to a minor amount of total generating capacity. Furthermore, this minor amount is much less than that defined by daily sunshine. (More than one-half of all electrical demand is during daylight hours, but solar power would have to be substantially cheaper than traditional sources before 100 percent displacement could be achieved even during daylight hours).

### Table 14. Central Receiver Simulation Results for Southern California Edison (SCE)

<table>
<thead>
<tr>
<th>Plant/site</th>
<th>Storage</th>
<th>Capacity credit at four levels of penetration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>5 percent</td>
</tr>
<tr>
<td>SCE study*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inyokern, California</td>
<td>None</td>
<td>32%</td>
</tr>
<tr>
<td>Inyokern, California</td>
<td>6 hour</td>
<td>80%</td>
</tr>
<tr>
<td>Aerospace Corporation study*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Santa Maria, California</td>
<td>6 hour</td>
<td>64%</td>
</tr>
<tr>
<td>Inyokern, California</td>
<td>6 hour</td>
<td>74%</td>
</tr>
<tr>
<td>Yuma, Arizona</td>
<td>6 hour</td>
<td>n.a.</td>
</tr>
<tr>
<td>Dispersed</td>
<td>6 hour</td>
<td>n.a.</td>
</tr>
</tbody>
</table>

**Note:** Southern California Edison is a summer peaking system with peak summer demands occurring at about 2 p.m.


*Aerospace Corporation, Penetration Analysis and Margin Requirements Associated with Large-Scale Utilization of Solar Power Plants, prepared for EPRI, EPRI ER-198, August 1976, Figures 4-6 and 4-9. Note that capacity credit is computed at 100 percent minus percentage backup capacity and that the dispersed plant site option had 60 percent of the solar capacity at Yuma and 20 percent each at Santa Maria and Inyokern.

---

**Hypothesis Two**

The second hypothesis is not supported by available studies; in fact, simulations suggest that one should expect a negative relationship between pumped storage and solar power diffusion. The simulations show that solar plant capacity credit is higher and does not decline as rapidly when solar penetration increases if dedicated storage is available (see Table 14). Unfortunately, the results of other studies show that dedicated storage is undesirable: "Unlike solar thermal generating plants where some amount of thermal buffer storage is an operational requirement, there appears to be no advantage to dedicated storage in photovoltaic power plants."  

Nondedicated system storage is found to be desirable in these simulations (assuming the cost is low enough), but added system storage would not solve the solar plant capacity credit problem, since solar plants and storage are substitutes: "Results show that the prior existence of photovoltaic plants reduces the value of added system storage. Photovoltaic plants and system storage are thus shown to be in at least partial economic opposition; that is, the value of each is strongly influenced by the existence of daily peak loads and each, when applied to a system, tends to flatten the daily load curve leaving a less advantageous situation for the other."  

These conclusions are readily transferable to solar thermal plants. While it is technically feasible to improve the capacity credit of solar plants through storage, the cost of energy storage is currently too high to make this an economically attractive method of improving solar capacity credit. If the cost and availability of energy storage methods are improved, solar plants will be forced to compete with off-peak power from conventional plants. The outcome of this competition is widely known and would only postpone further the arrival of solar electric power.

---

**Hypothesis Three**

The high fixed costs of solar plants and the lower weighted average cost of capital of municipal utilities suggest that municipalities will adopt solar sooner than investor-owned utilities. A recent study supports this hypothesis and finds that levelized busbar costs of a given solar plant are about 20 percent lower for municipalities than for investor-owned utilities. Since municipalities account for less than 10 percent of U.S. generating capacity, the size of this market is limited.

---

**Hypothesis Four**

Insolation varies significantly among regions in the United States.
Both of the reports cited above have utilized insolation data to develop their results. In both, the busbar costs of solar plant output vary inversely and approximately in proportion to the average level of annual insolation in each of the cities studied. All other things being equal, utilities in regions with high average levels of insolation should adopt solar plants sooner than utilities in regions with low average levels.

_Hypothesis Five_

Small utilities (less than 10 Mw) were the focus of the Aerospace report cited above and constitute less than 3 percent of U.S. generating capacity. The calculations of this report indicate that the costs of one Mw solar plants are about 2.2 times those of 10 Mw solar plants. While this report was limited to central receiver solar plants, similar scale relationships have been reported for other solar technologies. At this small scale, solar plants are currently roughly competitive with coal plants but cost more than diesel generators. Diesel is about one-fourth the cost of solar at one Mw but only about one-half at 10 Mw. While the Aerospace report did not investigate beyond the 10 Mw figure, other studies indicate that solar costs are cut in half as scale increases from 10 Mw to 100 Mw, at which point they begin to level off.

Since small utilities frequently have problems obtaining access to joint ventures, it is possible that utilities of 50 Mw to 100 Mw could find solar plants attractive sooner than other utilities if joint ventures are unavailable to them. The size of this market is rather small, however, as a percentage of total U.S. generating capacity.

_Consclusions_

One of the most important conclusions of this paper is that overall solar penetration is limited since solar capacity credit declines as penetration of a given utility increases. Solar plant busbar costs will have to fall well below those of traditional plants if solar penetration is to go beyond 5 or 10 percent of a given system. The overall market for solar plants is probably much more limited than is generally recognized.

The most important utility characteristic is definitely the degree to which daily insolation and daily peak demand coincide during the peak season. When these coincide, the capacity credit for solar is higher (but still below that of traditional generating plants), and the capacity credit declines less rapidly as solar penetration increases toward 20 percent.

The average level of insolation in a utility's service area is relatively important. The utility's size and access to pumped storage are relatively minor factors, but they are characteristics worth considering. Finally,

the ownership status of the utility is another characteristic of limited importance due to its effect on the cost of capital.

The analysis of this paper suggests that some utility characteristics will affect the rate of diffusion of central station solar electric power. Further work needs to be done to quantify and clarify these effects, however. Industrial and other direct uses of solar energy were not considered here, but that area could be more promising since the high costs of electricity distribution and high inefficiencies of conversion to electricity could be avoided.

The ultimate penetration of solar plants into utility systems will depend on developments in energy storage. If energy storage costs remain high, the ultimate penetration of solar electric power should be limited to a small portion of daytime loads. Diffusion would be slow and highly variable from one geographic region and utility to another. If energy storage costs can be reduced substantially, the diffusion of solar electric power could be delayed due to forced competition with off-peak power from conventional plants, but the ultimate penetration would be greatly increased.

_Notes_


2. Although plant costs were adjusted, where necessary, to reflect cost reductions achieved by mass production. No adjustment was required for the non-solar options since these plants were commercially available, and their costs reflect mass production economics.

3. The optimal scale of production assumed was: 100 Megawatts (Mw) for the solar options; 1150 Mw for the nuclear options; and 350 Mw for the fossil options.

4. The use of fossil and solar boilers connected to one common turbine-generators unit is currently under investigation.

5. The fixed charge rate covers the cost of capital, depreciation, insurance, and taxes.

6. General Electric Co., _Requirements Assessment of Photovoltaic Power Plants in Electric Utility Systems_, prepared for Electric Power Research Institute, EPR1 ER-685, Project 051-1, Technical Report, June 1978, vol. 2. It should be noted that the relationship between solar penetration and capacity credit shown in Table 13 of this paper was developed by the author from data summaries presented in the GE report. The authors of the GE report made no mention of this relationship and are not responsible for its development or any resulting error.

Diffusion of Technological Innovations among Privately Owned Electric Utilities: 1950–1975

Robert D. Stoner

This paper will examine the spread of technological innovations in the electric utilities industry and evaluate the factors which cause some utilities to adopt innovations earlier than others [see Smith 1977]. It was necessary to gather data on innovation adoption dates that had never been compiled before. The principal source was a questionnaire sent to a sample of 117 Class A and B utilities. It requested the adoption date of sixty innovations. These were primarily cost savings measures and covered every major area of utility operation: distribution, transmission, generation, accounting and billing, metering, and engineering. In general, innovations were chosen whose first adoption was in the 1950s or early 1960s and whose use spread throughout the 1960s and early 1970s. In addition, an effort was made to include only innovations which were obviously cost saving to a wide variety of utilities, regardless of geographic location, fuel, or importance of hydroelectric power. To ensure that these criteria were met, I solicited advice from over 20 U.S. electric utilities. Completed questionnaires were received from 66 companies. Only data for the 20 innovations for which the responses were judged most reliable, and whose adoption was most widespread, were used.
The principal goal was to determine the extent to which structural factors, both economic and regulatory, influence the technological innovativeness of electric utilities. More specifically, four questions were considered paramount: (1) Is there any evidence that utility size is significantly related to innovativeness? In particular, are larger utilities more innovative, relative to their share of industry output, than small utilities, as the Schumpeterian hypothesis would suggest [see Schumpeter 1950]? Conversely, do moderately sized (or even small) utilities do more than their share of innovation, as F. M. Scherer [1974; 1965a; 1965b] postulated? (2) Can we discern any noticeable differences in the regulatory climate surrounding utilities which were early adopters as compared to those that were later adopters? In particular, are utilities that operate in jurisdictions in which there is considerable regulatory lag, and consequently an opportunity to earn excess profits, more innovative than utilities in states in which profits are monitored closely? (3) How important is the influence of the Averch-Johnson [1962] effect on utilities' innovative behavior? That model emphasizes the potential distortion in input choices caused by rate-of-return regulation. Do we find evidence for this effect in the adoption patterns of utilities in the sample? (4) What is the effect of competition on technological innovativeness? Do utilities which face greater competition in their residential, commercial, and industrial markets tend, on average, to be earlier adopters than those which face little competition in these markets?

In addition to these four concerns, the study had another important goal. There have been several recent efforts to assess static or managerial efficiency among electric utilities [see Julo 1961; Pace 1970]. H. C. Petersen [1973] and Bernard Tenenbaum [in progress] have developed separate but similar econometric models to measure relative static efficiency and develop rankings of managerial performance. My study of technological innovativeness among utilities offers an excellent opportunity to develop rankings based on technological innovativeness and to compare this assessment with the rankings of managerial performance. Thus, I will compare my results and the Tenenbaum findings on static efficiency. Since my sample and the period studied were chosen to coincide with Tenenbaum's, this comparison provides a unique opportunity to compare static and dynamic assessments of utility performance. In addition, I will explore the question of the most likely relationship between static (or managerial) performance and dynamic (or technological) performance in a regulated industry. In this regard, it seems very doubtful that one would expect to find a one-to-one correspondence between dynamically and statistically efficient utilities.
dummy variable categorizing utilities as being from either fair value or original cost jurisdiction states.

The second regulatory tightness variable attempts to quantify the regulatory lag a utility faces. This lag refers to the period between rate hearings during which utilities are not subject to formal controls on profits. During this time there is an opportunity to earn higher than normal profits. The regulatory lag exists because of the inability or unwillingness of most state commissions to engage in constant and close supervision of utility profits. Several economists have suggested that this lag, rather than being a flaw, is a saving grace of the regulatory process. They suggest that it is the very absence of regulation, which in turn leads to the possibility of earning above-normal profits, that provides utilities with the incentive to cut costs and innovate [Trebting 1963].

An important variant of this regulatory lag theory has been suggested by Paul Joskow [1974]. He observes that regulatory commissions are political entities and therefore subject to political pressures from their constituencies — the legislature, environmental groups, and, more recently, consumers. These groups are much more sensitive to increases in utility rates than to increases in utility profits. Commissions, in turn, reflect the concerns of their constituents. They are largely indifferent to rising profits as long as rate levels remain constant or are falling. Joskow notes that from 1962 to 1968, a period of high profits, there were very few rate hearings. This lack of regulatory action was quite the opposite of what would have been expected if commissions did, in fact, regulate profits. Both the standard regulatory lag model and the Joskow variant reach the same conclusion: The opportunity exists for utilities to earn above the allowed rate of return for long periods, and this opportunity may spur utilities' cost-conscious and innovative behavior. I tested this hypothesis by forming a variable measuring supranormal profit expectations. For my purposes, these were assumed to be an average of the difference, from 1961 to 1966, between the actual rate of return on equity earned by the utility and the adjusted rate of return on equity that the state commission allowed in the last previous electric rate case in the utility's state. This is basically a proxy for the degree of regulatory lag a utility faces. I would expect, a prior, a negative sign on this variable in the delay regressions. Negative or small positive values of the supranormal profits variable would tend to occur in firms facing internal or external difficulties, or in firms operating in tight regulatory climates. These conditions would tend to promote longer delays. On the other hand, large positive values of the regulatory lag variable would tend to indicate that the firm is operating in a regulatory setting in which there are opportunities to earn above-normal profits. A utility management's perception of this fact should induce it to innovate more quickly. A finding of this sort would provide empirical support for those who believe that regulatory lag should be used systematically as a policy tool to promote utility efficiency and innovativeness.

The third aspect of regulatory "tightness" that was investigated is the influence of the Averch-Johnson effect on a utility's innovative behavior. That model emphasizes the potential distortion in input choices caused by rate-of-return regulation. Averch and Johnson show that a profit-maximizing utility, subject to a rate-of-return constraint which is greater than the cost of capital, will tend to produce a given output at a capital-labor ratio greater than that which minimizes costs. In several later analyses, W. J. Baumol and Alvin Kleverick [1970] and Scherer [1970] have suggested that this A-J input distortion is smaller when the utility is "less tightly constrained" and becomes more pronounced the more "tightly regulated" is the utility. Tightness of regulation in the A-J model is measured as the difference between a utility's actual rate of return on capital and its cost of capital. Using the Scherer and Baumol and Kleverick result, it is hypothesized that the A-J bias toward the adoption of capital-intensive technology will be greater for those utilities in which the difference between the actual rate of return and the cost of capital is small ("tight" regulation) than for those utilities for which the difference is large ("loosely" regulated utilities). Therefore, if the A-J variable is included in the delay model, we would expect its sign to be positive in the case of capital-intensive (labor and fuel saving) innovations, and negative (or nonsignificant) in the case of capital-saving innovations. In other words, tightly regulated utilities are hypothesized to be, on average, faster adopters of capital-intensive innovations and slower adopters of capital-saving ones.

**Competitive Climate**

The third explanatory variable in the delay model consisted of two measures of the competitive climate faced by a utility.

The first attempts to measure the extent of competition an electric utility faces from the sale of natural gas in its service area. It is defined precisely as the percentage of natural gas customers in its electric service territory for which a utility, its affiliates, or its subsidiaries are responsible. This measure was developed as a more meaningful indicator of the competition an electric utility faces from gas sales than the usual "combination" versus "separate" dichotomy. A utility could be a "combination" company (selling both electricity and gas) and still face considerable pressure from gas companies in its service territory. Likewise, a utility could be a "separate" (selling only electricity) and not
face much competition from gas if little is sold in the utility's service territory. For this reason, a utility's competitive position with respect to gas has been defined by relating its own number of gas customers to the total number of gas customers in its electric service territory. The expected sign on this variable was positive; the greater the degree of monopoly status of a utility in electricity and gas, the longer were its expected delay times.

The second competitive climate variable attempts to measure the extent of competition a utility faces because of the size and nature of its industrial loads. It is postulated that the larger the percentage of a utility's sales to industrial customers (especially if these are energy-intensive customers such as primary metals, cement, and chemicals), the greater the pressure on a utility to be innovative. This effect is expected for two reasons. First, unlike the case for residential customers, electric energy costs can be a rather large percentage of total costs for energy-intensive industrial customers. These firms do take energy prices into account in their locational decisions. Therefore, within regions of the United States where energy-intensive industries locate (in other words, where favorable fuel costs or good hydroelectric sites permit low cost electricity production), a utility would in all probability actively compete to attract such industrial loads from other areas of that region, from other regions of the country, and even from other countries.

The second source of competition that electric utilities face in their industrial markets stems from the wide variety of substitutes for electricity that large industrial concerns (especially energy-intensive ones) can use. Among the possible substitutes are cold, geothermal, solar, and self-generated electricity. These interfuel substitution possibilities are greater for large users than in residential and commercial markets primarily because the bulk of large industrial customer energy demand is for applications other than lighting (for which electricity is virtually the sole possibility). To weigh both of these effects, a variable was formed which took into account both the percentage of a utility's sales to industrial customers and the degree of energy intensity of the industries in the utility's state. A negative sign would be expected on this composite competition variable, indicating that the greater the degree of competition in industrial markets perceived by a utility, the shorter its delay in adopting a given innovation.

Several other variables were included in the delay regressions, but these will only be listed: a utility's growth rate, the fuel and labor prices a utility faces, and regional dummy variables to account for geographic differences among utilities. These variables were all expected to have some effect on a utility's delay in adopting a given innovation.

Robert D. Stoner

Statistical Results

The model was run by statistically relating (through regression) the delay of each utility in adopting a given innovation to the explanatory variables described above.

The model was run separately for each of the 20 innovations, and the results were mixed. First, the R² for the 20 regressions varied between -09 and .45, indicating that the model had relatively low explanatory power. Second, the coefficient of the firm size variable was negative in 18 of the 20 cases, and statistically significant in half, indicating that large firms tended to be earlier adopters of innovations than did smaller firms, other things being equal. The Schumpeterian hypothesis seems corroborated by these results. Third, the dummy variable which categorized utilities as original or fair value regulated was of the expected sign in 15 of the 20 cases, but was significant (at the 5 percent level) in only two. Fair value regulation seems to be associated with some increase in utility innovativeness.

Fourth, the results of the A-J test were inconclusive. For capital-intensive innovations, the sign of the A-J variable was positive (as expected) in 8 of the 11 cases, but only significant in two. Among the capital-saving innovations, the sign of the A-J variable was positive in 5 cases (which was not expected) and negative in 4. One of the negative coefficients was significant. These findings offer weak support for my dynamic interpretation of the A-J hypothesis. Fifth, the variable measuring regulatory lag was significantly negative for 7 of the 20 innovations, and was negative, although not significant, in 9 of the remaining cases. These results suggest fairly strongly that the degree of regulatory lag has an important effect on a utility's innovativeness. Sixth, neither of the "competitive climate" variables was significant in more than one of the regressions. Furthermore, there was no discernible pattern of the signs of these coefficients.

In assessing the overall results, an interesting explanation would account for the less than perfect success of the model in explaining a firm's delay in innovation. It is conceivable that a firm's static and dynamic performance is not primarily determined by the economic and structural environment it encounters. An emerging body of literature is beginning to emphasize the firm's internal organization as an important determinant of its performance [Coleman et al. 1957; Nason 1971; and Rogers 1962]. According to these theories, such factors as the firm's bureaucratic structure and resultant attitudes toward risk, and the age and educational status of company executives, are more likely to be closely associated with measures of performance than are such "external" factors as the competitive and regulatory climate within which the firm operates. These external factors are important
only insofar as they affect the internal bureaucratic structure and decision-making process. If this theory is correct, one would not expect a model, such as the delay model outlined above, incorporating only these external factors to have a great deal of explanatory power.

Unfortunately, I do not have data for utilities in the sample on organizational variables. However, an indirect test was devised of what I will call the "internal structure" or "firm effect" hypothesis, by examining the residuals from the 20 delay regressions. The residual for a particular firm represents the difference between its delay time and the delay predicted for the firm based on the model. When a residual is positive, the firm is taking longer to adopt the innovation than the model predicts; when it is negative, the opposite condition prevails. I assume that the existence of consistently positive (or negative) residuals for a utility across all 20 regressions is strong evidence that factors associated with the firm itself ("firm effects") account for a large part of the unexplained variation in its innovative behavior. In fact, in investigating the pattern of residuals for each utility, I found considerable evidence of just such a pattern. This result gives fairly strong support for the "firm effect" hypothesis, and, as well, sheds light on the mixed performance of the delay model.

Comparison of Dynamic and Static Efficiency Rankings

I compared my results on technological innovativeness with a recent study by Tenenbaum, following in the path of Iulo and Pace, which attempts to measure static (or managerial) efficiency among electric utilities in the 1960s. Tenenbaum developed an econometric model to predict average cost per kilowatt-hour and tested it on a sample of 56 privately owned utilities. By specifying the model so that all independent variables are largely free of managerial discretion, he was able to interpret the residuals of the regression as measures of a utility's managerial efficiency. Highly negative residuals represented utilities whose costs are lower than the model predicts, and they are therefore designated "efficient." Highly positive residuals represent relatively "inefficient" utilities. Residuals whose value is near zero represent utilities of average efficiency.

To compare Tenenbaum's results with mine, I ranked on the basis of their technological innovativeness each of the utilities in my sample which overlapped with the Tenenbaum sample (there were 50 such firms). To form these rankings, I weighted the delay time of a company independently: the degree of cost savings represented by that innovation and the initial investment expenditures entailed in adopting a particular innovation. These estimates of cost savings and initial investment were established with the help of a number of utilities that I visited in the course of the project. This weighting procedure was performed because it was felt that early adopters of substantial cost-saving and high initial investment innovations should be given more credit in the rankings than were early adopters of less cost-saving and lower initial investment innovations. The weighted delay times for each company were then added across all 20 innovations to form an index of innovativeness for that company. This was done for each of the 50 companies, and comparison of the index of innovativeness across companies established a utility's relative rank. In order to compare the findings on static and dynamic performance, I computed a correlation coefficient between my rankings and Tenenbaum's. The result was a negative, although not statistically significant, relationship, that is, there is some indication that utilities measured as management efficient tend to be technologically less progressive, and vice versa.

The comparison of my results with Tenenbaum's lends some credence to the theory that there is not necessarily a one-to-one correspondence between statically and technologically efficient utilities. Such a theory becomes more clear if we explicitly view the innovative process as taking place in a dynamic setting. In particular, if the availability of a new technique is seen not as a once-and-for-all development, but as a series of improvements over some period, early adoption is not necessarily the most desirable course. For example, the digital computer for accounting and billing purposes, one innovation I sampled, has experienced considerable development after its initial introduction and is still being improved upon today. If we perceive the innovation process as taking place within this constantly changing framework, a firm's optimal strategy for embodying new technology in capital will depend on three major factors.

The first two are the pattern of past purchases of related equipment and the rate at which such investment is depreciating (in the example given, the timing of purchases of older accounting and billing machines and their rate of depreciation). The third is the rate at which the new technology is being improved, both in reliability and sophistication, in its successive vintages (in the example, the rate of improvement in succeeding generations of digital computers). In this more realistic setting, it is perfectly conceivable that a firm's optimal strategy will be adoption later than by so-called industry technological leaders. I have referred to this phenomenon as the 'optimality of being second.' By waiting, a firm is able to 'leapfrog' to a more sophisticated and reliable version of the new technology (or, in the extreme, to another altogether new technology) and more fully depreciate its past capital stock. Slower adoption also allows other firms to incur the development or break-in costs associated with the initial introduction of an innova-
tion. This effect is likely to be particularly important in an industry such as electric utilities. Since there is very little direct competition among private systems, the industry is characterized by a general willingness to share information about technological development. In many utility industry engineering journals, for example, one can read very detailed descriptions and analyses of the operating experiences of companies that have tried new techniques or new machines. There is strong reason to believe, therefore, that firms which lag behind in introducing innovations are less likely to experience high costs during the break-in period.

The “optimality of being second” effect is likely to be much more prevalent in a regulated industry, such as electric utilities, than in nonregulated sectors of the economy. This result is expected because a well-managed, stably efficient electric utility will see mainly the above benefits of “waiting” to adopt innovations, not the risks or costs of later adoption, as would an unregulated firm. Among these risks and costs are the likelihood that significant business would be lost to more innovative rivals and the possibility that the firm would fall irrevocably behind its rivals on the technological learning curve and not be able to catch up. Neither of these risks is as great for an electric utility company as for an unregulated firm; utilities have a largely captive customer base, and no utility is likely to fall far behind its rivals technologically, since technological information in the industry is so readily shared.

To conclude, then, if these various hypotheses are accepted, the expected relationship between statically and dynamically efficient utility companies would be at least random, and probably inverse. A finding of randomness is suggested if one weights heavily the consideration that each utility, whether or not it is managerially efficient has made a unique set of past investment decisions. These are likely to dictate to some degree, its present innovation patterns and produce some distribution of delay times. According to this view, therefore, a firm’s managerial efficiency or inefficiency would not be a major determinant of its expected delays. There would be a more or less random relationship between dynamic and static efficiency.

However, if, in addition, one stresses the benefits of “waiting” (that is, "leapfrogging" and the avoidance of break-in costs), and one further postulates that stably efficient utilities are more likely to be cognizant of these benefits than stably inefficient utilities, an inverse relationship would then be expected between static and dynamic efficiency. Stably efficient utilities would be later adopters of many innovations, while those utilities always among the first to adopt would incur the costs of excessive "progressiveness" and show up as statically inefficient. Such a theory is a possible explanation for the results presented in this section.

**Future Work**

At least four suggestions may be made concerning the direction of future research.

First, although I have made a number of improvements on previous efforts to measure the competitive pressures a utility experiences, these efforts fall short on several counts. I have attempted to measure only the competition a utility faces from natural gas sales and in attracting industrial customers; there are several other sources of competition a utility is likely to face which I was not able to quantify. Competition for fringe loads in commercial and residential areas markets and competition for wholesale customers are two important examples [see Pace 1971, pp. 753–84 and 753–59].

These areas of competition lie mostly in the realm of what has been called “potential” competition. Schumpeter points out that even if a firm’s “actual” market position is stable and well defined, it is likely to experience to some degree a fear of erosion of its position. In Schumpeter’s words [1950, pp. 84–85], “the businessman feels himself to be in a competitive situation even when he is alone in the field.” The degree of this perceived threat is termed “potential competition,” and Schumpeter maintains that the forces of potential rather than actual competition are largely responsible for a firm’s innovativeness. If we grant this argument, then it becomes very important to attempt to measure various sources of potential competition, such as that for fringe loads and wholesale customers. However, it is precisely the “perceived” nature of these competitive threats that makes them hard to pin down empirically. A great deal of work remains to be done in developing quantifiable measures of potential competition.

Second, a comparison of the measure of dynamic efficiency used in this study— the technological innovativeness of utilities— should be made with other measures of dynamic efficiency. In the applied economics literature, dynamic efficiency is often measured using an index of total factor productivity (TFP). [See Kendrick 1973.] Rodney Stevenson [in progress] is currently in the process of refining just such a TFP measure for a large number of electric utilities. Important work remains to be done in contrasting the rankings of utilities based on the total factor productivity results with the rankings I developed based on innovativeness.

A third promising possibility for further work would be the extension of my analysis of privately owned electric utilities to publicly owned utilities. The innovations that have been chosen apply equally
well to many federally, municipally, and cooperatively owned U.S. utility systems. Given the growing interest in the subject of public ownership of electric utilities in this country, such a comparison of private and public utility innovative performance would certainly have important policy implications.

A fourth avenue of exploration would be to compare the innovative performance of both private and public electric utilities with that of relatively unregulated U.S. industries [see Neubergr 1976]. Certain of the innovations studied here, such as computerized accounting and billing procedures, apply equally well to various industries in which rate-of-return regulation is not used, such as banks, department stores, or even large national magazines. An examination of the diffusion of such an innovation across the three primary categories of industries in this country (private, public, and regulated) would certainly be an important contribution to our understanding of the relative merits of these different forms of ownership.

The tentative nature of many of the conclusions of this study and the numerous suggestions for further research indicate that more questions were raised by the inquiry than were answered. However, that is a possible outcome for any truly worthwhile research effort. It can only be hoped that future inquiries into the process of technological innovation in the electric utilities industry will be able to build on the groundwork, however tentative, of the present study.

References


Obsolescence, Depreciation, and New Technology

F. W. Sinden

Setting depreciation rates is inherently difficult because the economic value of a machine depends not only on its physical state of repair but also on its ability to compete in the future with new machines, even with machines that may not yet exist. Thus, depreciation depends inescapably on an assessment of an unknown future. It is therefore not possible to set output prices (since depreciation is a component of these prices) without making judgments, implicitly or explicitly, about future technology.

The aim of this paper is to show, in the context of an idealized model, how information about the future goes into the determination of an ideal depreciation schedule, and how to show how obsolescence, either sudden or gradual, affects ideal output prices. It also says a few words about how depreciation in the real world may deviate from the ideal, and what some of the consequences of that deviation may be.

To explain ideal depreciation, it seems best to use a simple, concrete example. The conditions 1 will assume for this example are ideal in two main ways: (1) the setting is an ideal competitive market, and (2) the future is perfectly known. Even with these idealizations, the example is not quite so simple as one might suppose.

Island Air Service

Transportation between the mainland and a remote island is provided by a single airplane which shuttles back and forth. Demand is constant and inelastic. The airplane is replaced from time to time by a new and better one. Complete information about all future airplane models, including dates of introduction, maintenance schedules, operating costs, and so forth, is given in a large catalogue published by the Perfect Preciscience Society. Cost information from this catalogue can be summarized by the function of two variables $z(t, r)$:

$$z(t, r) = \text{cost (\$ per day) at time } t \text{ of owning and operating an airplane of vintage } r.$$  

All payments are assumed made when due without borrowing or credit. At the initial moment (when $t = r$), $z(t, r)$ is a spike representing the capital cost of the airplane. Thereafter, $z$ may fluctuate with $t$, displaying large bumps at times of major overhaul and gradually rising as the airplane wears out.

From the function $z(t, r)$ may be derived an optimal sequence of replacement times, $r_1$, $r_2$, $r_3$, ..., for every starting time $r$. More specifically, $r_1$, $r_2$, $r_3$, ..., is the sequence which minimizes the present worth of cost looking forward from time $r$. (The discount rate in the present worth is taken as given.) The cost at each moment under optimal replacement will be denoted by $c(t, r)$:

$$c(t, r) = z(t, r), \quad t \leq r \leq r + 1.$$  

(1)

The situation is illustrated in Figure 1.

Given the cost function $c(t, r)$, what fare $f(t)$ should be charged at each moment of time? To answer this question, let us set up a competition among all possible fare functions $f(t)$. Let us suppose that the air shuttle is set up as a franchise which is continuously up for bid. Competitors bid by stating the fare they will charge at the current moment if given the franchise. The incumbent can stay only by meeting the lowest bid fare at every moment.

It might seem that this system would be unstable, that an outsider could bid uneconomically low in order to gain access, with the intention of raising the fare later to make up the loss. But that strategy would not work, because the later attempt to raise the fare above the competitive level would result in instant loss of the franchise and a net loss for the period of tenure.

Anyone taking over the franchise is free to start fresh with the latest technology: He or she is in no way committed to any existing equipment or past decisions made by the predecessor. Thus the in-
cumbent must constantly compete with the latest technology. This does not mean, however, that he must instantly scrap the current airplane and buy a new one every time an improvement occurs, but it does mean that he must replace prudently from time to time and that the replacement schedule is influenced by the continuing competition among new models. It is in this way that the fare determined by the competitive franchise bidding incorporates obsolescence.

\[ C(t, T) : \begin{cases} Z(t, T_0) & T = T_0 \\ Z(t, T_1) & T_1 < T < T_2 \\ Z(t, T_2) & T = T_2 \end{cases} \]

Figure 1. \( C(t, T) \) is the Expenditure Rate of Time \( t \) for an Optimal Replacement Sequence Starting at Time \( T \)

The franchise bidding leads to a unique competitive fare \( f(t) \) that is smoother over time than \( c(t, \tau) \) and that yields a revenue \( R(t) \) with the following property: The present worth of \( R(t) \) looking forward from any time \( \tau \) is equal to the present worth of cost for the starting time, \( \tau_0 = \tau \). In symbols,

\[ \int_\tau^\infty e^{-r(t-\tau)} R(t) \, dt = \int_\tau^\infty e^{-r(t-\tau)} c(t, \tau) \, dt, \text{ for every } \tau. \tag{2} \]

The unique function \( R(t) \) is such that the air service exactly breaks even in the long-run present worth sense, no matter when the franchise starts. If we denote the right-hand side of Equation (2), that is, the present worth of cost looking ahead from \( \tau \) by \( C_{pw}(\tau) \), that is,

\[ C_{pw}(\tau) = \int_\tau^\infty e^{-r(t-\tau)} c(t, \tau) \, dt, \tag{3} \]

then the solution of Equation (2) yields for the competitively determined revenue:

\[ R(t) = rC_{pw}(\tau) - C_{pw}(t). \tag{4} \]

Despite its simplicity, this equation is interesting and is capable of producing results that may seem surprising. These arise from the second term, \( C_{pw}'(t) \) (which stands for the derivative of \( C_{pw}(t) \)). Let us pause for a moment to consider the effect of this term. Note that it occurs with a minus sign. When \( C_{pw}(t) \) declines, the revenue \( R(t) \) is larger (see Figure 2). What would make \( C_{pw}(t) \) decline? Technological improvement is the most likely cause: When technology improves, costs for potential new entrants decline. Improving technology, then, causes the fare to be higher than it otherwise would be, at least momentarily.

![Figure 2: Example Illustrating the Effect of the Second Term](image-url)
This increment represents the cost of obsolescence. I will call it, therefore, the "obsolescence premium." In the long run, of course, it is more than made up for by improved costs as reflected in the first term (see Figure 2 again). But technological improvement does have the effect of making prices higher now because it forces earlier replacement and faster depreciation.

If technological advance is steady, let us say exponential, then the obsolescence premium is imbedded in the smoothly declining price and is not very visible. There is another case, however, in which the obsolescence premium shows up in a very visible and even startling way. This is the case in which a single technological improvement occurs at one point in time.

To lead into this example, consider first a simple case in which no technological improvement occurs, as shown in Figure 3. As can be seen from the figure, technology is completely stagnant: the future looks exactly the same from any starting point. Mathematically, this stagnation is expressed by the fact that \( c(t, \tau) \) depends only on the difference \( t - \tau \). From this it follows that the present worth of cost \( C_{aw}(t) \) is constant, hence that the revenue \( R(t) \) is constant.

Let us consider, however, what this constant revenue means to the franchise holder. Just after \( \tau \), the time of takeover, the franchise holder faces a bulge in cost that temporarily exceeds the constant revenue. During the period of the bulge, therefore, he must incur debt. Later, the cost is just enough lower than the revenue so that the difference can pay the interest on the debt.

Figure 4 shows a slightly more elaborate example in which an improved airplane becomes available at time \( \tau_0 \). The new airplane is so much better than the old that even its initial cost bulge is less than the old operating cost. Thus there is no doubt that the incumbent must switch to the new plane right at \( \tau_0 \) no matter what the vintage of the present plane.

What does the competitively determined revenue stream \( R(t) \) look like in this case? We know what it must look like for \( t > \tau_0 \) because in that range, looking forward, the situation is the same as in the previous example: The cost stream is invariant with respect to starting time, hence \( R(t) \) is constant. But what does \( R(t) \) look like before \( \tau_0 \)? One plausible possibility is that \( R(t) \) is constant at the level appropriate to the old technology right up to \( \tau_0 \), whereupon it drops abruptly to the constant level appropriate to the new technology. Another plausible possibility is that \( R(t) \) declines gradually before \( \tau_0 \) as competitive bidders for the franchise anticipate the low cost airplane that will soon be available.

In fact, neither of these possibilities is the right one. What actually happens is that \( R(t) \) rises in anticipation of the new airplane, as shown at the bottom of Figure 4. (The dashed line in this figure shows the constant value \( R(t) \) would have assumed if no improvement had occurred.) In the period before \( \tau_0 \) the revenue (hence fare) is actually higher than it would have been if the new airplane had not been developed.

How is this possible? How can the franchise holder get away with raising the fare? Why is it not bid down by the competitors? The reason is that the competitors are in a difficult position in the period just before \( \tau_0 \). If they offer to charge only the dashed line fare they will lose money. Recall that in the previous example an entrant had to incur debt to pay for the initial cost bulge, and that the interest on the debt was paid indefinitely thereafter by the excess of revenue over cost. But
in the present example this comfortable excess is wiped out after \( t_0 \); competition forces the price down to a level just sufficient to cover the costs associated with the new airplane, with nothing left over to pay interest on old debts. In the period before \( t_0 \) then, when entrants have to start with the old type of airplane, they are faced with an initial debt on which they will not be able to pay interest after \( t_0 \). Therefore, they must plan to recover the cost of the initial bulge quickly in order to be free of debt when the new airplane becomes available at \( t_0 \). Mathematically, the revenue rise preceding \( t_0 \) is generated by the previously mentioned second term in the equation for revenue — Equation (4). From the viewpoint of the incumbent, the revenue rise just pays off the old debt in time for the replacement at \( t_0 \). Without the technological advance at \( t_0 \), only one plane is ever purchased, and it is never depreciated. With the advance, it must be depreciated fully by \( t_0 \).

Now suppose that the technological advance at \( t_0 \) is unexpected, that is, no one knows it will occur until shortly before \( t_0 \). Then up to the time when the news breaks, the fare fails to include the obsolescence premium. When the news does break, the fare jumps up to take account of the obsolescence premium, but by then it is too late to recover the past premium that should have been paid. The incumbent operator suffers an irretrievable loss equal to the depreciation that he should have taken in the past.

In an uncertain world there is always the possibility of events such as the one just described which will result in loss. Under these conditions, therefore, all competitors for the air service franchise will include in their calculations an allowance for insurance against risk. One may expect the competitively determined fare to be higher and depreciation faster than it would be if the future were known.

One characteristic of an uncertain future is that there is bound to be disagreement about it. Different competitors will have different opinions about the probabilities. The most optimistic competitor will bid lowest and therefore win the bid. But the most optimistic view is not necessarily the most realistic, and the optimist may well go out of business. In the airplane example, then, under uncertainty one might expect the franchise to be held by a series of Pollyannas who charge too little and go bankrupt one after another!

Although the revenue \( R(t) \) in this example (and in general the \( R(t) \) given by Equation (4)), is determined by a competitive process, can we be sure that it is socially optimal? Under our assumption of fixed, inelastic demand, there is nothing, for example, to prevent the revenue (hence fare) from fluctuating widely, as in the case just considered. Would some smoother \( R(t) \) be socially better? Or would \( R(t) = c(t, \tau) \), which in general is less smooth, be better? (This is the case in which all costs are recovered instantaneously.)

If we take the present worth of consumers’ surplus as the criterion of social welfare, then the case of the island air service is indeterminate: Since demand is fixed, the utility derived from the air service by its customers is fixed and independent of the fare. So long as cost is minimized, any revenue function \( R(t) \) whose present worth equals the present worth of cost will maximize the present worth of consumers’ surplus.

The case we have been considering, however, is a limiting special case. In the more general case in which demand varies with price and in which cost is proportional to output, maximization of the present worth of consumers’ surplus does yield a unique \( R(t) \), and in fact this \( R(t) \) is given by Equation (4), provided this equation is reinterpreted to
apply to costs and revenue per unit of output. In the air service example, then, the revenue stream determined by competitive franchise bidding does maximize the present worth of consumers’ surplus in the limit as demand becomes inelastic.

Depreciation

The only given data in the air service example are the cost data from the Perfect Prescience Society catalogue (summarized by \( c(t, \tau) \)) and the discount rate, \( r \). We return now to the cost function \( c(t, \tau) \) and derive from it the competitively determined depreciation schedule.

Consider a franchise holder who takes over at time \( \tau_0 \). Looking forward from a later time, \( t^* \), the present worth of revenue is generally greater than the present worth of the cost stream in which the franchise holder finds himself. The difference between these present worths is the economic value of the current airplane at the moment \( t^* \), because that is what someone else, taking over the franchise at time \( t^* \), could pay for the airplane and just break even in the long run. If we let \( V(t, \tau_0) \) equal the value of equipment at time \( t \) in an optimal replacement sequence that began at time \( \tau_0 \), then:

\[
V(t, \tau_0) = \int_{\tau_0}^{t} e^{-(t' - \tau_0) r} \left[ R(t') - c(t', \tau_0) \right] dt'.
\]

(5)

The revenue \( R(t) \) can be eliminated from Equation (5) by means of Equation (2), so that the equipment value is given directly in terms of the given cost data. Generalizing the present worth notation introduced in Equation (5), we let

\[
C_{PW}(t, \tau_0) = \int_{\tau_0}^{t} e^{-(t' - \tau_0) r} c(t', \tau_0) dt'.
\]

(6)

\( C_{PW}(t, \tau_0) \) is the present worth of cost looking forward from time \( t \), if the replacement sequence started at the earlier time \( \tau_0 \). We can rewrite Equation (5) as:

\[
V(t, \tau_0) = C_{PW}(t) - C_{PW}(t, \tau_0).
\]

(7)

Unless \( t \) happens to coincide exactly with one of the replacement times \( \tau_0 \), the franchise holder is in possession of a used airplane at time \( t \). According to Equation (7), the value of this airplane just equals the cost advantage it affords over starting fresh.

Graphs of the two cost terms in Equation (7) are shown in Figure 5 for a typical case. Note that \( C_{PW}(t, \tau_0) \) coincides with \( C_{PW}(t) \) at replacement times. This is true for every starting time \( \tau_0 \). In fact, \( C_{PW}(t) \) is the upper envelope of the ensemble of functions \( C_{PW}(t, \tau_0) \) in which \( \tau_0 \) ranges over all values. Technological change is reflected in the shape of \( C_{PW}(t) \). If technology improves, then \( C_{PW}(t) \) declines as shown in Figure 4.

---

F. W. Sindlen

---

Figure 5. \( V(t, \tau_0) \) Equals the Value of Equipment at Time \( t \) in an Optimal Replacement Sequence that Began at Time \( \tau_0 \).

Equation (7) shows how costs, present and future, determine the economic value of a machine. Two kinds of cost enter into this determination: those associated with the machine itself and those associated with other competing machines, including machines of the future. In principle, even technologies of the distant future affect the current value of the current machine. Fortunately, the magnitude of this effect diminishes exponentially with distance into the future.

The ideal economic depreciation rate is the negative rate of change of \( V(t, \tau_0) \) with respect to \( t \). If this depreciation schedule is followed, the initial capital will always be exactly recovered at the end of each machine’s life. This is true under very general conditions, including inflationary conditions. The idea that during inflation depreciation should accumulate the cost of the next machine rather than pay back the cost of the current machine is not reflected in this theory. If cost \( c(t, \tau) \) inflates without a corresponding increase in the interest rate \( r \),
then lifetimes may be affected, but capital still will be exactly recovered at the end of each machine's life.

Let us consider now what the market-determined ideal depreciation looks like in the simple special case in which technology improves like $e^{-\lambda t}$ as shown in Figure 6. In this case the optimal replacement times are uniformly spaced. The faster the technological improvement, the shorter the interval between replacements. Within one lifetime the optimal depreciation schedule is not straight line, but is definitely accelerated, as shown in Figure 7.

If for some reason the air service used straight-line depreciation rather than the economically optimal accelerated depreciation, then the fares charged to customers would be too low at the beginning of each airplane's life and too high at the end. This could occur, for example, if the air service were running under a regulatory arrangement whose rules only approximated the effects of market forces. If this arrangement were replaced by competitive bidding in the middle of an airplane's life, then the depreciation schedule would jump from straight line to the accelerated economic schedule. That would mean that the capital that should have been recovered but was not would be irretrievably lost.

Figure 7. Deviation of Straight-Line Depreciation from the Ideal in the Case Illustrated in Figure 6

*For exponentially improving technology

Figure 6. Example of a Simple Cost Function $Z(t, \tau)$ with Steady Improvement

In the air service example, different technologies are characterized entirely by different fixed cost streams over time. If demand and output are variable, and economies of scale are present, then the cost function $z(t, \tau)$ must be generalized at least to the form $z(t, \tau, q)$.
where $q$ is the level of output. In addition to economies of scale, there may be constraints on the rate of expansion of a new technology. In that case, the cost function might have the form $z(t, q, q)$. Except for the variable demand case with no economies of scale, in which

$$z(t, q, q) = z(t, \gamma t),$$

the mathematics is much more complicated in these cases, and it is no longer clear that competition leads to socially optimal results.

If a technology has a memory, then the integral of $q$ over time may be involved. This is true, for example, if a learning curve describes the development of a new technology. It is also true if the technology uses a finite resource, so that costs rise as the resource is used up. The competition in this case can become a rather complex game in which society may or may not be a winner.

Notes

4. A remark by Paul Samuelson may be interpreted as saying that depreciation should cover the replacement of a machine rather than the machine itself: “Suppose prices are rising sharply. If I sell my goods for enough to cover labor and other costs and also to cover depreciation, you might think I am breaking even. What would an accountant say who figures depreciation on the basis of the past low prices originally paid for my machines and building? He, too, would say I am breaking even. But in fact I can be said to have been selling my goods at a real loss; for when my machines and buildings have worn out, I shall not have enough money to reproduce them at the new higher price level.” P. A. Samuelson, *Economics*, 7th Ed. (New York: McGraw-Hill, 1967), p. 103. See also A. E. Kahn, *The Economics of Regulation* (New York: Wiley, 1970), vol. 1, p. 54.
6. For an example of such a game see David Pullin, “Approaches to the Structuring of Energy Systems,” Ph.D. diss., Department of Physics, Cambridge University, 1978.

Comments

John E. Cummings

My purpose is to relate the papers of Robert Stoner and David Huet-
ner to the development of new energy resources as future electric power options. More specifically, I would like to offer several com-
ments in extension of these papers directed toward the acceptance of sol-
lar central power applications into the marketplace.

The work described by Stoner deals with the process of privately
owned electric utility acceptance and integration of technological inno-

vations. As the basis for his statistical analysis, Stoner surveyed electric utilities regarding the time required for acceptance of specific technical innovations. Results indicated spans of several months to several years. The innovations studied were already commercially available and also were more or less independent system elements, generally a replace-
ment product or a better process. Examples included pad-mounted trans-
formers, impact wrenches, chain saws, buckets on hydraulic derricks, and automatically switched capacitors for voltage regulation.

Let us consider the acceptance of new electric power generation
options. The solar power tower is one example. The point to be made
is that its introduction would involve not a single technical innovation,
but many. What is more, the numerous collection, conversion, and
utility integration innovations are dynamically interrelated. This is a

317
very difficult systems problem insofar as technical innovation is concerned, and the result would be protracted delay before significant advance.

A second aspect, quite often ignored, relates to the research and development phases necessary to make technological innovation commercially available. Before a utility integrates a new option, technical and economic feasibility must be demonstrated. In addition, the all-important requirements of reliability and availability must be satisfied before a capacity credit can be assigned to a new plant option. Often forgotten is the considerable time necessary to complete this overall process. The research and development stages for several new electric power generation options are shown in Figure 1. The solid lines represent an optimistic expectation, while the dotted lines reflect uncertainties associated with the present early stages of development.

The message is, regardless of what the new power options might be, a minimum of twenty years should be anticipated before any significant use can be achieved. There is little doubt in my mind that the conventional power generation options presently available will be our mainstay well into the next century. The real danger lies in deferring construction of conventional power plants while waiting for some breakthrough in the advanced technologies.

I would like to offer one further comment regarding solar energy options and the research and development process just discussed. It is a common mistake to view all solar options as having reached the same phase of technological development. In fact, they have not. While solar heating and domestic hot water applications have demonstrated performance capabilities and are on the threshold of commercial dissemination, solar thermal conversion is still in the engineering feasibility stage, as is wind energy conversion. Photovoltaic conversion remains in the scientific feasibility phase. Given this diversity, we should anticipate significant differences in the research and development processes for each of the solar options, and we must be prepared to accept the cost and time needed to bring various technologies through the intermediate phases.

I would also like to offer a few comments to supplement Huettners paper. In my view, much more of this kind of study must be undertaken. Quite often, supporters and promoters of the new solar applications, while eager to push into hardware stages, ignore the marketplace into which the applications are to be integrated. It is essential that solar central power plants be viewed from the standpoint of the plant's value when integrated into utility systems.

Huettners study by General Electric sponsored by the Electric Power Research Institute (EPRI) aimed to determine the requirements and impact associated with the integration of photovoltaic electric power generation. Computer codes and methodologies familiar to the electric utility industry were used to
model reliability, economic dispatching, and generation expansion. In addition, photovoltaic power plants were modeled and dynamically simulated for specific utilities using actual weather and utility system data. Under the EPRI Solar Program, similar analyses have been carried out for solar thermal-electric power generation (power tower) and wind electric conversion. At present, studies are under way to extend wind and photovoltaic analyses to dispersed applications, including electric power production in homes and by industry.

Some general results from these studies reinforce Huetter's conclusions. They are summarized below.

(1) The value of solar thermal, photovoltaic, and wind electric power plants varies among utility systems and according to weather conditions. Correlations between solar plant availability and utility demand functions displaced are of key importance.

(2) Storage, except buffer storage for stability, should always be considered for its value to the system as a whole. If inexpensive storage were available, it could be used in conjunction with conventional base load capacity.

(3) One should be suspicious about accepting and using cost estimates because advocates tend to favor their own systems. Very few studies feature a common cost basis, and early in development stages there is bound to be considerable uncertainty in estimating cost.

In conclusion, I am happy to see a growing awareness of the importance of market forces when analyzing future power options such as those utilizing solar energy. If new applications are to displace conventional power generation options, they must be developable and demonstrated in such a way that utility planners are convinced they can be reliably integrated into utility systems and operational practices.

References


Comments

William G. Shepherd

Although F. W. Sinden's discussion is quite interesting, I would like to direct my remarks to the papers by Robert Stoner and David Huetter.

Stoner's paper offers much of substance, but I found most telling his insight that regulated utilities will rationally want to innovate. As parallel monopolists, they each gain by letting others be the first to try new techniques and work out the bugs. Of course, this is a prescription for technological stagnation, punctuated by abrupt limitations once an innovation has been tried by a plucky innovator. The sudden growth of nuclear power after 1963 fits that pattern, for example. Similar conditions hold in the telephone sector, where the use of new inventions has often been voluntary—and slow. It is only direct competition that stirs innovation, in utilities and elsewhere.

Stoner's "competitive climate" variables did not turn out to correlate with the speed of innovation. Yet, these variables do not really reflect competition, because there is not, in fact, any direct competition. Gas and other alternative fuels only exert pressure some users to hold down electricity prices. Such soft "competitive pressure" probably would not affect innovation in any case.
On the whole, Stoner found what one expects: Electric utilities have wide latitude in making innovations and little pressure to make them quickly. Nor do the regulators — given such talents, resources, and powers as they may have — have any clear effect. That, too, is what one would expect from observing the common run of state regulation. Stoner could go farther to draw this lesson: If the public wants better innovation in this sector, drastic policy changes will be needed. The most obvious would be to open direct competition in bulk power, as Leonard W. Weiss, James Meeks, and others have proposed. At any rate, Stoner gives us a highly competent analysis of a disturbing reality.

Huetter's analysis stirs the inevitable misgivings that one has toward a study from an oil state — Oklahoma — about solar power. It duly finds that solar energy has a cloudy, or should we say dim, future, at least for the next few decades. My best efforts have failed to locate any basic flaw in his work, so I accept the main shape of his findings. Solar power will indeed be economic first, and perhaps also last, in the Sun Belt. It may help heat buildings during the mild winters there, but it will do its heaviest duty during the scorching summer, when sunlight abounds. But for what? Perhaps mainly for air conditioning.

The lesson is disheartening. Solar power's main foreseeable use will apparently be to cool buildings, using a lot of capital for collecting sunlight and then manufacturing cold air. Perhaps this main growth of solar power can be obviated by simply designing buildings with better natural ventilation. Solar power may be an expensive way to accomplish what is partly unnecessary.

Yet, we can expect to be bombarded with claims that massive solar investments are a must for America. That may be a "soft" energy path down which we should tread softly, very softly, and perhaps not very far. Only if other, less capital-intensive versions of solar energy can be found are we likely to use the renewable energy of the sun as a major source of heat.
Risk Measurement and Rate of Return under Regulation

Ronald W. Melicher

An essential element of public utility regulation is the determination of a “fair” rate of return. Such an allowed return, in theory, is to be substituted for what otherwise might be monopoly profits. It is inadequate, however, to refer to minimum acceptable rate of return levels unless they are first established in a risk-return framework. Theory and evidence suggest that investors are risk averse and thus expect to receive returns that are commensurate with risks. This paper explores concepts of risk measurement as they relate to establishing a “fair” rate of return under conditions of regulation.

We begin our discussion by reviewing the public utility or regulatory concept in equation form. The equation often is stated as follows:

\[ R = E + (V - d) r, \]  

(1)

where:

- \( R \) = total operating revenues;
- \( E \) = total operating expenses (operating expenses plus depreciation and taxes);
- \( V \) = value of plant, equipment, and working capital;
- \( d \) = accumulated depreciation; and
- \( r \) = allowed rate of return.
For our purposes, we focus only on the r, or rate of return variable. The importance of establishing a "fair" rate of return in a risk-return framework can be traced to the Bluefield and Hope cases. In brief, a fair return should be adequate to maintain existing capital and to attract new capital; thus, it must be commensurate with the returns earned by other firms with corresponding risks.1

The ability to maintain and attract capital funds implies that risk must be viewed in a "market" context. That is, regulated and unregulated firms must compete on a risk-adjusted basis in the bond and stock markets. Modern investment theory maintains that the U.S. capital market is efficient in two basic ways. First, current bond and stock prices reflect all available information about the underlying form. Furthermore, as new information is provided, security prices are characterized by an instantaneous or rapid adjustment.2 This implies that, at any given moment, security prices reflect "real" or intrinsic values.

Second, U.S. security markets are efficient in a risk versus return framework. This is to say that investors expect higher returns on investments that are characterized by higher risks. While realized returns may be lower for riskier investments in any given period, they should be higher, on average, over the long run. This implies that investors should be rewarded with higher average returns for their willingness to accept greater risks.

Thus, the expected rate of return, \( E(R_t) \), for any risky security (denoted as i) is comprised of a risk-free rate of interest \( (R_f) \) plus a risk premium \( (R_p) \) and can be expressed in equation form as:

\[
E(R_i) = R_f + R_p.
\] (2)

The risk premium, \( R_p = E(R_i) - R_f \), reflects the anticipated incremental reward an investor expects to receive for investing in risky securities.

The risk-free return can be specified for a particular investment holding period as the yield on a U.S. government debt security of comparable maturity.

Risk premiums, in contrast, have been viewed in several different ways. For example, since U.S. government debt securities are regarded as being free of default risk, it can be argued that risk premiums reflect the probability of default. Investment theory contends that the return versus risk relationship can be expressed in a mean-variance or mean-standard deviation framework.3 Risk viewed in this context is measured as the variability in rates of return. Greater variability is associated with higher risk premiums. More recent developments, referred to as capital asset pricing theory, suggest that the return versus risk relationship should actually be expressed in a mean–systematic risk framework.4 Systematic risk is associated with movements in the macroeconomy and thus cannot be eliminated through diversification by holding a portfolio of securities. This theory contends that premiums will be paid in the securities markets only for that portion of risk which cannot be diversified away.

To summarize, risk in the form of risk premiums can be viewed as (1) risk of default, (2) total variability of rates of return, and/or (3) systematic variability of rates of return. Earlier it was suggested that a fair return should be commensurate with the returns earned by other firms with corresponding risks. This leads us to ask which is the proper measure(s) of risk to be used in regulatory proceedings in determining a fair rate of return for public utilities. As a first step, it seems worthwhile to examine historical relationships between realized returns and measures of risk.

Risk-Return Relationships:
Some Empirical Evidence

Investors traditionally have been concerned about the quality of or risk of default on debt obligations. Bondholders, of course, fear default by issuing firms on interest payments and/or bond principal. Stockholders also are concerned about debt default since such action often results in corporate bankruptcy. Presumably, then, investors expect to be compensated with higher returns as the probability of default increases. Empirical evidence by Lawrence Fisher [1959] and others indicates that risk premiums on corporate bonds can be largely explained by measures of default risk. It also follows that, for a given firm, risk premiums should be higher for the common stockholders relative to the bondholders, since the latter have a prior claim against the firm’s earnings and assets. Thus, for a given firm, common stockholders should expect a relatively higher rate of return.

Bond ratings should be closely tied to risk premiums since they also reflect risk of default. The most comprehensive study of the relationship between bond ratings and default rates was conducted by W. Braddock Hickman [1958]. He examined risk-return relationships by bond quality over the period 1900–1943. Some of his findings are summarized in Table 1. Bonds were rated I through IX, with I–IV representing investment quality (comparable to AAA through BBB today) and V–IX reflecting increasingly lower quality. The ratings in Table 1 were made at time of issue and indicate that default rates (percentage of issued bonds that defaulted) increased as bond ratings decreased. However, realized returns or yields also increased, on average, as bond ratings declined. Returns on investment quality bonds averaged slightly more than 5 percent, in contrast with 8.6 percent for lower quality bonds. This evidence supports the contention of a
trade-off between bond ratings or risk premiums which reflect default risk and rates of return — both expected and realized.

Also of interest is whether bonds as a whole provide relatively lower risk-adjusted returns when compared to common stocks. This brings us to our second measure of risk. Investors are said to be risk averse and thus expect to be compensated for uncertainty of returns. Greater variability in returns, and thus higher uncertainty, implies greater risk and the expectation of higher average returns.

A recent study by Roger G. Ibbotson and Rex A. Sinquefield [1978] investigated average annual returns on high quality (AAA and AA) corporate bonds and New York Stock Exchange common stocks for the period 1926–1975. Table 2 summarizes these findings in a risk-return framework. It can be seen that, on the basis of arithmetic average returns, corporate bonds failed to outperform common stocks in any of the five decades. (On the basis of ten-year holding periods, however, corporate bonds had compound annual rates of return that exceeded those of common stocks during 1926–1935 and 1966–1975.) It also can be seen that the variability in annual returns was substantially less for corporate bonds relative to common stocks.

The risk-return relationship can be normalized by calculating the coefficient of variation. In brief, the higher that coefficient, the greater the risk. Table 2 supports the contention that, on average, common stocks are riskier than corporate bonds in terms of variability of returns; thus, common stockholders should expect higher returns. Of course, in any given period, returns on bonds may exceed returns on common stocks, and vice versa.

Now we turn to the risk-return relationships for common stocks as a group. Shannon Pratt [1965] examined New York Stock Exchange common stocks over the period 1929–1959. He grouped stocks into quintiles on the basis of their variability (standard deviation) of monthly rates of return during three- and five-year periods. This historical variability then was compared with future average rates of return. Average returns were found to increase with risk (measured by historical variability) for all groups except for the most risky. It should be remembered, of course, that Pratt was trying to use past risk measures to explain future or ex ante rates of return.

Empirical evidence supports the contention that risk-averse investors expect, and usually realize, higher returns for making riskier investments. This relationship seems to hold across bond quality differences, between corporate bonds and common stocks, and across stock quality differences (reflected in variability of returns), with the possible exception of very risky common stock investments.

### Table 2. Risk-Return Relationships for Corporate Bonds and Common Stocks

<table>
<thead>
<tr>
<th>Period</th>
<th>Average annual returns (percent)</th>
<th>Standard deviation of annual returns (percent)</th>
<th>Coefficient of variation (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1926–1935</td>
<td>7.17</td>
<td>4.35</td>
<td>1.15</td>
</tr>
<tr>
<td>1936–1945</td>
<td>3.98</td>
<td>1.29</td>
<td>1.35</td>
</tr>
<tr>
<td>1946–1955</td>
<td>1.90</td>
<td>2.54</td>
<td>1.19</td>
</tr>
<tr>
<td>1956–1965</td>
<td>6.29</td>
<td>5.00</td>
<td>1.26</td>
</tr>
<tr>
<td>1966–1975</td>
<td>3.92</td>
<td>8.26</td>
<td>2.11</td>
</tr>
</tbody>
</table>

*These data represent arithmetic average annual rates of return, with ten-year compounded annual rates of return shown in parentheses.*

*The coefficient of variation is calculated as the standard deviation divided by the arithmetic average return and reflects a standardized risk-return relationship.*

---

Systematic or Market Risk

Systematic or market risk needs to be addressed separately from default risk or total variability of return risk measures. The measurement of systematic risk is founded in modern investment theory. Rates of return on all securities are influenced by the "market" and thus are correlated with "market" returns. The market, in turn, is influenced by macroeconomic conditions. Thus, this market risk is common to all firms and has a systematic impact on the returns of individual common stocks.

It is common practice today to measure the market in terms of the Standard & Poor's 500 Stock Price Index or the New York Stock Exchange Index. Systematic risk indicates the extent to which common stock returns move with returns on one of these market measures. Figure 1, using a simple linear regression format, illustrates how systematic risk is measured. Monthly percentage rates of return on a common stock or portfolio of stocks are plotted against the monthly percentage rates of return on the S&P 500 Index for the same intervals.

The calculation of systematic or market risk can be expressed in simple regression form as:

$$ Y_i = a_i + b_i X + e_i, \tag{3} $$

where:

- $Y_i$ = monthly return on stock $i$;
- $a_i$ = intercept term or alpha coefficient;
- $b_i$ = slope of the regression line or beta coefficient;
- $X$ = monthly return on the market index; and
- $e_i$ = error term in the regression equation.

The alpha coefficient indicates how well a stock performed after the systematic effect was eliminated. Systematic risk, as measured by the slope of the regression line, may be less than, equal to, or more than the market's risk of 1.00. Stocks or portfolios with beta values of less than 1.00 are said to be less risky than the market. The opposite is true for beta values greater than 1.00.

The error term in equation (3) also is important in that it captures that portion of the total variability in a stock's return which is not explained by systematic risk. For example, in Figure 1, if all the $x$'s had fallen on the regression line, then the error term would be zero. The "goodness-of-fit" of the regression equation also can be viewed in terms of the coefficient of determination ($R^2$), which measures the extent to which all of the variation in the common stock's return is explained by variation in the market index. $R^2$'s can range from 0.00 (no correlation) to 1.00 (perfect correlation). Thus, the total variability in a stock's returns can be separated into systematic risk and unsystematic (unique to the firm) risk components. A coefficient of determination value of 1.00 implies that systematic risk explains all of the variability in a stock's returns, and all $x$'s in Figure 1 would fall on the regression line.

Empirical evidence suggests that $R^2$'s for individual stocks probably fall, on average, in the .20 to .30 range. Thus, for the "average" firm, approximately 20–30 percent of the variability in its stock returns is explained by systematic risk, with the remaining 70–80 percent of the variability being due to unsystematic or unique firm risk. At the same time, empirical evidence also indicates that $R^2$ values rapidly approach 1.00 for reasonably diversified portfolios. This means that unsystematic or unique risk can be eliminated through portfolio diversification.

William Sharpe [1978a] and others have examined the relationship between systematic risk and rates of return for diversified portfolios.
Table 3 shows systematic risk-return results for New York Stock Exchange common stocks during 1951–1967. One can see consistent and nearly linear relationships between average annual returns and ex post beta values. However, for less diversified portfolios (that is, when all risk is not systematic risk), Sharpe indicates that the return-beta relationship is reasonably consistent but not as good as the relationship shown in Table 3. In summary, systematic risk reflects nearly all risk in diversified portfolios and is linearly related to realized returns.1

Table 3. Systematic Risk-Return Relationships for Common Stocks, 1951–1967

<table>
<thead>
<tr>
<th>Risk-return group</th>
<th>Average annual returns (percent)</th>
<th>Ex post beta values</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 (highest)</td>
<td>22.67</td>
<td>1.42</td>
</tr>
<tr>
<td>9</td>
<td>21.12</td>
<td>1.24</td>
</tr>
<tr>
<td>8</td>
<td>20.73</td>
<td>1.21</td>
</tr>
<tr>
<td>7</td>
<td>20.24</td>
<td>1.14</td>
</tr>
<tr>
<td>6</td>
<td>19.83</td>
<td>1.05</td>
</tr>
<tr>
<td>5</td>
<td>18.68</td>
<td>.95</td>
</tr>
<tr>
<td>4</td>
<td>17.67</td>
<td>.90</td>
</tr>
<tr>
<td>3</td>
<td>15.84</td>
<td>.79</td>
</tr>
<tr>
<td>2</td>
<td>13.73</td>
<td>.66</td>
</tr>
<tr>
<td>1 (lowest)</td>
<td>11.58</td>
<td>.58</td>
</tr>
</tbody>
</table>


Capital Asset Pricing Theory and Regulation

Modern investment theory revolves around the development of what is referred to as capital asset pricing theory, known in model form as the capital asset pricing model (CAPM). The model is expressed as:

\[ E(R_i) = R_f + (E(R_m) - R_f) B_i, \]

where:

- \( E(R_i) \) = expected rate of return on risky security \( i \);
- \( R_f \) = risk-free or riskless interest rate;
- \( E(R_m) \) = expected rate of return on the market; and
- \( B_i \) = the systematic risk or beta for security \( i \).

The reader should note that the CAPM model is very similar to equation (2), the only difference being that the risk premium \( (R_m) \) is expressed as consisting only of systematic risk in the form \( E(R_m) - R_f \) \( B_i \).

Proponents of the CAPM contend that it can be applied to inefficient stocks or portfolios when systematic risk does not explain total variability of returns, as well as for efficient portfolios. Even when the returns on an individual firm’s common stock are affected substantially by unsystematic risk or risk unique to that firm, CAPM proponents argue that the proper measure of risk is that risk which cannot be eliminated through diversification. This is because, under capital asset pricing theory, security markets will pay premiums only for systematic or nondiversifiable risk. It is further argued that it is up to the risk-averse investor to eliminate unsystematic risk through the construction of efficient portfolios.

However, the usefulness of CAPM in regulatory proceedings should not be accepted without question. We need to ask whether systematic risk is the only risk concept important in identifying firms with corresponding risks for purposes of determining fair rates of return. Furthermore, is CAPM an adequate method for determining the cost of equity capital for regulated firms? Before addressing these questions, it seems useful to examine historical systematic risk characteristics for regulated firms.

Systematic Risk Characteristics for Electric Utilities

Table 4 shows alpha, beta, and coefficient of determination characteristics for 77 electric utility common stocks. The calculations are based on five years of monthly stock price data and the use of Standard & Poor’s 500 Stock Price Index to represent the market. Average beta values are substantially below 1.00, as would be expected. However,

Table 4. Systematic Risk Characteristics for Electric Utility Common Stocks

<table>
<thead>
<tr>
<th>Period</th>
<th>Alpha</th>
<th>Beta</th>
<th>Coefficient of determination (R²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960–1966</td>
<td>-.133</td>
<td>.885</td>
<td>.272</td>
</tr>
<tr>
<td></td>
<td>(.351)</td>
<td>(.208)</td>
<td>(.999)</td>
</tr>
<tr>
<td>1964–1968</td>
<td>-.288</td>
<td>.642</td>
<td>.164</td>
</tr>
<tr>
<td></td>
<td>(.236)</td>
<td>(.727)</td>
<td>(.666)</td>
</tr>
<tr>
<td>1966–1970</td>
<td>-.286</td>
<td>.767</td>
<td>.294</td>
</tr>
<tr>
<td></td>
<td>(.397)</td>
<td>(.183)</td>
<td>(.880)</td>
</tr>
<tr>
<td>1968–1972</td>
<td>-.256</td>
<td>.708</td>
<td>.244</td>
</tr>
<tr>
<td></td>
<td>(.319)</td>
<td>(.196)</td>
<td>(.883)</td>
</tr>
<tr>
<td>1970–1974</td>
<td>-.515</td>
<td>.654</td>
<td>.259</td>
</tr>
<tr>
<td></td>
<td>(.447)</td>
<td>(.198)</td>
<td>(.100)</td>
</tr>
</tbody>
</table>

SOURCE: Based on systematic risk calculations made by Merrill Lynch, Pierce, Fenner & Smith, Inc.
NOTE: Data calculations are based on 77 electric utilities. Mean values with standard deviations are given in parentheses.
they also exhibit volatility over the 1962–1966 through 1970–1974 period. The average $R^2$ values fell in the .24 through .29 range (with the exception of the years 1964–1968), indicating that, on average, over 70 percent of the variability in returns was due to unsystematic risk unique to the industry and/or individual utilities.

The existence of an industry effect is further supported by the negative alpha values throughout the several periods under study. Capital asset pricing theory holds that alpha values usually should be zero. The possibility of an industry effect is also supported by the fact that electric utility stock prices were declining during 1965–1975, while industrial stock prices experienced a slight upward trend over the same period.

In a recent article, Eugene Brigham and Roy Crum [1977] contend that fundamental risks increased for the electric utility industry between the mid-1960s and mid-1970s. They cite environmental problems, fuel shortages, future demand uncertainties, inflation-related problems (cost of capital investments and cost of financing), and deterioration in "quality" of earnings as reasons for increased fundamental risk. However, as can be seen from Table 4, average beta values were not increasing to reflect this perceived additional risk.

### Table 5. Beta Calculations for the Electric Utility, Natural Gas, and Telecommunications Industries

<table>
<thead>
<tr>
<th>Regulated industry</th>
<th>July 1976</th>
<th>October 1978</th>
<th>t-test of means</th>
<th>Kendall-rank correlation coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electric utility industry</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eastern United States (55 firms)</td>
<td>.750</td>
<td>.740</td>
<td>.598</td>
<td>.922</td>
</tr>
<tr>
<td>(102)</td>
<td>(101)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Central United States (48 firms)</td>
<td>.718</td>
<td>.705</td>
<td>.467</td>
<td>.752</td>
</tr>
<tr>
<td>(102)</td>
<td>(105)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Western United States (18 firms)</td>
<td>.731</td>
<td>.714</td>
<td>.669</td>
<td>.072</td>
</tr>
<tr>
<td>(107)</td>
<td>(108)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural gas industry (51 firms)</td>
<td>.787</td>
<td>.776</td>
<td>.283</td>
<td>.841</td>
</tr>
<tr>
<td>(196)</td>
<td>(185)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Telecommunications industry (19 firms)</td>
<td>.760</td>
<td>.727</td>
<td>.469</td>
<td>.897</td>
</tr>
<tr>
<td>(177)</td>
<td>(179)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall average</td>
<td>.750</td>
<td>.736</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Sources:** Based on beta estimates calculated by the Value Line Investment Survey.

**Notes:** Mean values with standard deviations are given in parentheses.

### Recent Beta Values for Regulated Firms

Table 5 provides systematic risk estimates for the electric utility, natural gas, and telecommunications industries as of July 1976 and October 1978. These estimates reflect calculations made for the Value Line Investment Survey and are based on weekly stock return data for individual firms relative to the New York Stock Exchange Index. Betas are estimated on the basis of five years of stock price data. Thus, while the beta estimating process differs between Merrill Lynch and Value Line, the systematic risk estimates in Tables 4 and 5 are reasonably comparable. Value Line, however, does not provide estimates of alpha and $R^2$ values.

Table 5 also indicates that beta estimates made in 1976 and 1978 seem to be highly stable. This is shown by the fact that the $t$-tests for differences in mean values were not significant, while the Kendall rank correlation coefficients indicate that estimated betas remained relatively constant for individual firms between the two periods.

### Table 6. Bond Ratings and Systematic Risk Relationships

<table>
<thead>
<tr>
<th>Regulated industry</th>
<th>Standard &amp; Poor's bond ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AA or higher</td>
</tr>
<tr>
<td>Electric utility industry (84 firms)</td>
<td>.725</td>
</tr>
<tr>
<td>Natural gas industry (29 firms)</td>
<td>.700</td>
</tr>
<tr>
<td>Telecommunications industry (13 firms)</td>
<td>.600</td>
</tr>
<tr>
<td>All three regulated industries (126 firms)</td>
<td>.700</td>
</tr>
</tbody>
</table>

**Sources:** Standard & Poor's bond ratings and the Value Line's betas are as of October 1978.

**Notes:** Average beta values with number of firms are given in parentheses.

### Systematic Risk and Bond Ratings

This paper began with a concept of default risk reflected in risk premiums and bond ratings. Next, total variability of returns was recognized as a measure of risk. Systematic or market risk then was introduced as a measure of risk. All three measures should be interrelated. The total variability measure captures both systematic and unsystematic risk for individual common stocks. Bond ratings also probably...
reflect the total riskiness of the firm, whereas beta captures only the systematic portion of a firm's total risk.

Table 6 explores the relationship between bond ratings and systematic risk. Relatively few regulated firms have AAA ratings on their debt today, and thus they are grouped together with the AA rated bonds. Similarly, only a few regulated firms have debt rated below BBB, and these are included with the BBB rated bonds. Theory would suggest that low beta firms would have relatively high bond ratings, and vice versa. This relationship seems to hold when all three regulated industries are grouped together. Such a relationship does not hold, however, for the electric utility industry. In fact, average betas are relatively greater for the AA bond rating group. Furthermore, statistical tests do not indicate significant differences in average beta values across bond rating categories. This suggests that bond ratings may reflect more than a regulated firm's systematic risk.*

Estimating a Fair Return

Utility rate of return regulation, based on the previously cited Blissfield (1928) and Hope (1944) decisions, evolved as a comparable risk—comparable earnings test. Under this approach, utilities are allowed to earn rates of return on their book equity comparable with those returns being earned by firms of similar risk.

All three of the risk measures discussed earlier can be used in conducting the "comparable earnings" analysis. Risk of default can be judged in terms of existing bond ratings or by examining those factors (such as interest coverage ratios, debt ratios, and profitability ratios) which explain bond ratings.18 Total variability of returns also can be used to assess comparable risk. While the variability in book equity returns should serve to confirm default risk assessments, total variability in stock market returns (dividend yields plus price changes) reflects the stock market's assessment of the firm's total riskiness. These measures of total riskiness should, of course, serve to corroborate one another. Finally, firms could be compared to establish similarity in the proportion of total stock return risk which is explained by systematic or market risk.

Since the early 1960s, however, the discounted cash flow (DCF) approach increasingly has been used to estimate the return required by investors that will enable the firm to maintain and attract capital. The DCF model is generally expressed as:

\[ k_e = \frac{D_t}{P_s} + g, \]  

where:

- \( k_e \) = cost of equity capital;
- \( D_t \) = expected cash dividends per share;
- \( P_s \) = current stock price; and
- \( g \) = expected long-run growth rate in cash dividends.

Note that this \( k_e \) is not the same as \( r \) in the regulatory equation (equation (1)) because \( r \) represents a weighted average cost of debt and equity capital.

In theory, the \( k_e \) for firm \( i \) should be equal to the expected return, \( E(R_i) \), such that

\[ E(R_i) = \frac{D_t}{P_s} + g = R_f + R_p. \]  

(6)

However, at issue is whether \( R_p \) represents total risk (total variability) or only systematic risk. It seems plausible that proponents of the DCF method, at least prior to the development of capital asset pricing theory, thought that \( k_e \) reflected a firm's total riskiness. But, for DCF and CAPM to produce comparable estimates of the cost of equity capital, CAPM proponents must assume that \( k_e \) in the DCF model reflects only systematic risk and thus ignores unsystematic risk (unique to the firm or industry).

Evidence shows that CAPM and beta applications increasingly have been used as part of testimony in rate of return hearings.11 The use of beta, along with other risk measures, in conducting the comparable risk—comparable earnings test seems reasonable. However, the use of CAPM to estimate a regulated firm's cost of equity capital is more questionable. Is it reasonable for investors to be attracted to utility common stocks if they are to be compensated only for systematic risk? Based on the evidence supplied earlier in this paper, investors who concentrate their equity investments in the common stocks of utilities while being guaranteed a relatively low portfolio beta would fail to diversify away all unsystematic risk. This is due to an investment or regulated firms effect. At the same time, it is not necessary to hold utility common stocks in order to construct efficient portfolios.

The process of regulation itself also seems to contribute to investor risk. Utility common stock investors are subject to "systematic" efforts on the part of regulators to prohibit excess profits. Willard Carleton [1974], as well as others, points out that rate of return—rate base regulation is designed to equate stock prices and book values per share — a phenomenon which does not exist for the nonregulated firm. Furthermore, the presence of regulatory "lag" adds another dimen-
sion of risk for the utility common stock investor. While regulation restricts the upward potential for returns, regulators may not act to the same degree on the downward side, according to Brigham and Crum [1975]. They suggest that the present political climate is more conducive to utilities earning less than, as opposed to more than, their costs of capital. They conclude that this may lead to nonsymmetrical returns. Note that utility stock prices have trended downward since the mid-1960s as compared with the stock prices for industrial firms. Market-to-book values dropped markedly for utilities during this period in contrast with a lesser drop for industrials. Without question, utility common stockholders suffered substantially from unsystematic or unique firm and/or industry risk during this period.

While it is difficult to forecast whether these risk-return relationships will continue in the future, from recent past experience it appears questionable whether utility common stock investors will accept expected compensation only for their systematic risk. The process of maintaining and attracting capital in the future may require, at least in part, compensation for unsystematic risk associated with regulation and regulatory lag. At this time, it might be shortsighted to use CAPM as the sole basis for estimating the cost of equity capital for utilities. CAPM applications to utilities thus should be viewed with caution. Similarly, if the DCF method reflects only systematic risk, it also should be administered with caution.

Summary

This paper has examined three measures of risk—default risk, total variability of stockholder returns, and systematic or market risk—as they relate to determining a fair rate of return under regulation. Expected compensation only for systematic risk may prove to be a shortsighted approach to effective regulation. While systematic risk measures may have changed little in recent years for utilities, many would argue that total riskiness has increased due to an increase in unsystematic risk. And, since all of the unsystematic risk associated with regulation may not be diversifiable, estimates of fair rates of return to utility common stockholders should include more than compensation for systematic risk.

Notes


2. Modern investment theory distinguishes three forms of the efficient markets hypothesis. These encompass examination of the value of (1) historical stock price patterns (weak form), (2) publicly available information (semistrong form), and (3) privileged information (strong form). Tests of the efficient markets hypothesis are summarized by James Lorie and Mary Hamilton [1973].

3. Attempts to quantify risk-return relationships in a mean variance framework generally are traced back to efforts by Harry Markowitz [1952] in terms of portfolio theory and selection.

4. Several authors were working on capital asset pricing theories during the early 1960s. The development of the capital asset pricing model (CAPM), however, is attributed to William Sharpe [1964].

5. For example, see Merrill Lynch, Pierce, Fenner & Smith, Inc., Investment Performance Analysis, July 1975, Appendix D.

6. In diversified portfolios of 30 or more common stocks, systematic risk accounts for approximately 85–95 percent of total risk. That is, $\beta^2$ are in the 85–95 range. For example, see ibid.

7. While beta coefficients are not very stable over time for individual stocks, Marshall Blume [1971] and others have shown that betas are relatively stable over time for larger portfolios. The relationship between betas and stock returns for individual firms also is questionable. For example, Robert Levy [1974] identified a rather weak return-beta relationship for individual stocks over the 1961–1970 period when he used beta coefficients calculated in one year to predict returns in the next year. Other researchers, such as James Farrell [1975], suggest that return-beta relationships may be distorted even for portfolios if the portfolios are concentrated in homogeneous groups (growth, cyclical, or stable stock classifications).

8. Some researchers also have been concerned about the relationship between systematic risk and business and financial characteristics. It seems reasonable to expect that differences in betas between firms should be explainable, at least in part, by differences in accounting measures of risk. For example, Ronald Melicher [1974] examined accounting and beta relationships in the electric utility industry during 1967–1971. He found that accounting measures such as dividend payout ratios, equity return ratios, capital structure ratios, asset size, and stock market trading activity were significant in explaining 33–41 percent of the variation in beta levels exhibited by electric utilities. In another study, Melicher and David Rush [1974] were able to explain only approximately 25 percent of the change in betas exhibited by electric utilities between the 1963–1966 and 1967–1971 periods on the basis of changes in financial characteristics.

9. In a study of the relationship between betas and bond ratings, Carl Schwediman and George Pinches [1973] found that, on the average, higher beta values were associated with lower bond ratings; however, statistical significance was found only between relatively extreme bond rating categories. Melicher and Rush [1974] examined the relationship between bond ratings and betas in the electric utility industry and found that, while average betas tended to increase as bond ratings declined, betas did not differ significantly across investment quality (AAA, AA, A, and BBB) bond ratings.

10. A recent attempt to explain electric utility bond ratings was conducted by George Pinches, Clay Singleton, and Ali Jahanian [1978]. Shyam Bhan-
dari, Warren Boe, and Robert Soldosky (1979) also have attempted to forecast changes in electric utility bond ratings.

11. Early academic discussion of systematic risk and CAPM applications in regulatory proceedings include the efforts by William Breen and Eugene Lerner (1972) and Stewart Myers (1972). Recently, Richard Pettway (1978) examined the impact of Consolidated Edison’s 1974 dividend omission on utility bets.


References


---. 1978. "Reply to Comments on "Use of the CAPM in Public Utility Rate Cases."


The Cost of Equity Capital: A Model for Regulatory Review

Basil L. Copeland

The current practice in public utility regulation is to permit regulated firms to earn a return on rate base equal to the cost of capital.1 As a practical matter, this cost of capital is computed as the weighted average of the costs of the various sources of capital — debt, preferred stock, and common stock — that make up the firm's capital structure. Since the costs of debt and preferred stock are matters of contractual agreement, rarely is there any disagreement among cost of capital witnesses as to the proper computation of such costs. The cost of equity, however, is almost always a subject of dispute in a rate case, and it is often the single most intensively litigated issue. To a certain extent, this is surprising. The "cost of equity" is well established conceptually as a part of the canon of modern corporate financial theory, and it is surprising that so little has been done to develop rigorous estimation procedures.2 To fill this void, the regulatory community has adopted a bewildering variety of ad hoc procedures to "estimate" the cost of equity. These often appear to be little more than thinly veiled searches for evidence to confirm rate of return recommendations developed prior to any actual capital market analysis. The problem this poses for regulatory commissions as they review the reams of testimony presented on this issue in a typical rate case was given formal recognition recently in a rulemaking proposed by the Federal Power Commission (now the Federal Energy Regulatory Commission). The intent of the rulemaking, as proposed, was to provide "reasonably specific guidance to the participants in rate proceedings as to the type of evidentiary showing that is required to justify a recommended rate of return."3 While it is debatable whether the proposal would have accomplished its objective, there is no question as to the genuine need for guidance from commissions as to the type of evidence that is necessary to justify a recommended rate of return. The presentation of "evidence" on this issue has become little more than a frenetic exercise in casual empiricism designed to give the appearance of evidence and expertise while avoiding the substance of them. The commission appeared to recognize this fact when it stated in a recent Nevada Power Company decision that, "to the extent that current market data is made a part of the evidentiary presentation, the Commission is oftentimes not given, nor can it determine, a logical basis on which to connect this record data with the recommended rate of return. This situation not only severely hampers the Commission's decision making process but, as noted above, reduces the overall value of that participant's rate of return recommendations."4 It is not at all an exaggeration to suggest that the "evidence" often presented in support of rate of return recommendations is of no scientific or probative value whatsoever. The redeeming virtue of such evidence appears to lie in the fact that it clutters up the hearing record and creates uncertainty in the minds of commissions and courts as to the actual cost of capital. In the face of uncertainty, commissions and courts tend to be more liberal in their rate of return allowance than they might otherwise be.

Myth and Method in Estimating Equity Capital Costs

The tendency toward arithmomania in rate of return testimony takes its worst form in the testimony of witnesses who present "evidence" on a variety of methods — comparable earnings, discounted cash flow, risk premium methods, earnings-price ratios, and so forth — which, remarkably enough, all result in the same estimate of the cost of equity. It does not take a particularly bright person to suspect that the witness's prior conception of what the return should be, rather than the logic of a particular methodology, is responsible for the findings, and we should not be surprised when the testimony of such a witness fails to provide a commission with "a logical basis on which to connect [the] record data with the recommended rate of return." The ability to make all methods produce the same result suggests a level of sophistication
achieved only by those capable of transcending logic and reason. This may suffice if the only purpose of expert testimony is to go through the motions, but it is certainly unsatisfactory to anyone seriously interested in estimating the cost of equity.

A present state of affairs, with respect to the rate of return testimony is appalling. Commissions are confronted by competing witnesses, each supposedly an expert. Never is there any consensus among these experts as to the proper methodology to be employed to estimate the cost of equity, or the results that follow from the application of different methodologies, or even the results that follow from the application of similar methodologies. It is now the convention for regulators and witnesses alike to assume that the problem of estimating the cost of equity is essentially intractable, that expert witnesses are constrained by the nature of the problem to rest their recommendations and conclusions in the final analysis upon their "judgment" as "experts." Indeed, the necessity of falling back on one's "informed" judgment to justify one's rate of return recommendation has become a badge, not of shame for failing to develop a suitably rigorous and objective methodology, but of "expertise."

Are we constrained by the nature of the problems of estimating the cost of equity in relying upon the intuition of self-proclaimed experts? Does the lack of agreement among expert witnesses reflect the fact that (1) the cost of equity cannot be objectively determined, or that (2) too many witnesses do not know how, or do not want to know how, to estimate the cost of equity objectively? At the risk of incurring the wrath of almost anyone who would make. The cost of capital is said to be a matter of public interest; in this subject, I will suggest that the latter is a more likely explanation for the present state of affairs. In attempting to demonstrate this, I hope that the point of view presented here is at least thought-provoking, and I even dare to believe that it is correct. I do not doubt that it will be controversial, for the myth of "informed judgment" has almost become dogma in regulation. It brings to mind a passage from John Kenneth Galbraith's introduction to his book "Money": "Those who talk of money and teach about it and make their living by it gain prestige, esteem, and pecuniary return, as does a doctor or witch doctor, from cultivating the belief that they are in privileged association with the occult— that they have insights that are nowise available to the ordinary person." Galbraith also might have been talking about cost of capital witnesses.

The regulatory community has exaggerated a difficulty into an impossibility. Indeed, the cost of equity is the subject of much controversy, but if obtaining objective estimates of the cost of equity is not easy, neither is it impossible, as we have been led to believe. We may honestly wonder how many expert witnesses are interested in knowing how to estimate the cost of equity objectively. As it now stands, it does not require much in the way of training or talent to become an "expert" cost of capital witness, and whenever there is free and unimpeachable evidence, in a market, buyers usually find little trouble in being furnished with exactly what they demand. If a utility intends to request a return on equity of 15 percent, it will experience no difficulty in finding a cost of capital expert willing to testify that the firm's cost of equity is (lo and behold!) 15 percent. We are not deluded into thinking that this is mere coincidence. As consultants want to earn the lucrative fees utilities are willing to pay for such services, it is at least tacitly understood that they will supply what is demanded; if they cannot, or will not, the utilities simply shop around until they find someone who offers a product more in line with what they desire. Those willing to supply utilities with what they want no doubt prefer that regulatory commissions remain under the misimpression that the cost of equity is largely a matter of judgment; the ability to rely on ad hoc and subjective estimation procedures provides sufficient opportunity to make sure that the outcome of their analysis of a firm's cost of equity is not inconsistent with the result they were paid to produce. This greatly enhances the marketability of their services.

Not everyone, of course, is convinced that the cost of equity is largely a matter of judgment. Charles Donahue, writing in the Michigan Law Review, has said: "Granted the importance of precision in this area, one can only be disturbed at the way the calculation of the cost of capital is currently made. The cost of capital is said to be a matter of public interest; in practice this means that its computation is a kind of guessing game in which a jumble of numbers, including the returns allowed by other commissions and those earned in vastly different industries, are thrown before the commission, which then pulls a compromise figure out of its hat." But only a few recognize the regulatory process for what it is — an organic and casual empiricism. The rest are content to believe that cost of capital witnesses cannot estimate the cost of equity without making judgments along the way that make impossible any claim to objectivity with respect to their findings. This opinion cannot be taken seriously by anyone with an understanding of the philosophy and methodology of science, for a fundamental characteristic of science is its refusal to accept the authority of "informed" judgment. In science, any appeal to authority that cannot be tested is inadmissible.

Scientists assure an acceptable level of objectivity in research by a rigid adherence to method. This is not to say that their judgment has nothing to do with their approach to problem solving, but by adhering
to a logical and scientific method they can bring their judgment under the control of scientific discipline. How this is done has been described in an excellent way by Loyd Fischer and Howard Osborne.

Hypotheses are to be formulated in anticipation of empirical (i.e. inductive) analysis. The researcher must identify and delineate problems and hypothesize probable relationships among variables and possible solutions to problems prior to empirical investigation else the result is inefficient research. Inadequate conceptualization of an inquiry prior to data collection increases the likelihood that personal bias and preconceptions will adversely affect the objectivity of the analysis. All researchers work within value structures which affect not only the problems selected for investigation, but also, the questions asked, the hypotheses posed, and the conclusions found. But the impact of the researcher's value structures and preconceptions can be nullified if empirical tests of the hypotheses are sufficiently objective. 7

Objectivity in research is thus predicated upon (1) conceptualization of the problem prior to data analysis with the object of deriving testable propositions and (2) actual tests of the propositions. Deductive elaboration of alternative hypotheses leads to predictions that are then tested against reality and experience. One of the principal contributions of Karl Popper to the philosophy of science has been to point out that we can only falsify hypotheses, we can never verify them. Testability thus implies falsifiability: Confirmations of a theory do not count unless they are the result of a serious effort to falsify the theory. Testing then consists of formulating a theory in such a way as to predict some novel fact which may not be true, and a theory is said to be corroborated to the extent that it withstands a serious effort to falsify it.

According to the philosophy of science, the influence of judgment on the objectivity of one's findings regarding the cost of equity can be brought under the control of scientific discipline, and the thought that it cannot is not to be taken seriously. 8 It requires an a priori concept of theory of how investors determine the prices they are willing to pay for shares of stock in a public utility. But a theory is not enough. It must be formulated in such a way as to produce testable propositions about the relationship between empirically measurable phenomena. These propositions must be genuinely testable, that is, they must predict novel phenomena that may not be true. In the words of Popper, this means that "criteria of refutation have to be laid down beforehand; it must be agreed what observable situations, if actually observed, mean that the theory is refuted." 9

The present state of affairs with respect to cost of capital testimony can therefore be attributed to the lack of scientific discipline on the part of those who present themselves as experts on the cost of capital. They have inadequately conceptualized the problem of estimating the cost of equity, and they have not been concerned with stating which observable situations, if actually observed, would mean that their theory as to the investor's required return is refuted. But the fact that they have not done these things must not be mistaken for proof that they cannot be done. In the remainder of this paper we explore what can and has been done to displace myth with method in the estimation of equity capital costs.

Econometric Models of the Cost of Equity

The Traditional Approach

The cost of equity is the return investors require before they will invest in the shares of a firm's equity. This return is not contractually arranged, as is the interest paid on debt or the dividend paid on common stock, and those who invest in the shares of a company's stock do so because of an expectation of current and future dividends. The equilibrium market price of a company's stock reflects this expectation, and the problem of estimating the cost of equity is thus one of assessing investors' expectations as reflected in equilibrium market prices. Recalling the discussion above about the role of method in research, the first step in developing an objective methodology for estimating the cost of equity is to have a conceptual model or theory of how investors establish the prices they are willing to pay for shares in the equity of a corporation. While investors' expectations are not directly observable, deductive elaboration of the theory should produce testable hypotheses, propositions that are testable by reference to data that are observable. If the hypotheses are genuinely testable, that is, falsifiable, then nonjudgmental inferences regarding investors' expectations and the cost of equity may be fairly inferred from the data as long as objective statistical procedures are used to test the hypotheses.

The pioneering attempt to develop an objective model of the cost of equity was made by Merron Miller and Franco Modigliani [1966]. From the very beginning they noted that the problem of estimating the cost of capital is necessarily one of inference: "Now, because it is based on anticipations, is the cost of capital any longer a directly observable magnitude. It must, rather, somehow be inferred from what is observable, namely, the market prices of the various kinds of claims represented by the different securities." 10 The fact that market prices are based on expectations did not lead Miller and Modigliani to assume, as do so many today, that the cost of equity cannot reasonably and objectively be inferred from data that are observable. Using rather traditional valuation theory, they derived a single-equation econometric model of the valuation process as a basis for inferring the cost of capital.
for electric utilities from observable data. Their model provoked extensive comment, much of it critical, but it was nevertheless pathbreaking and suggested new areas of research to others. Since it was a seminal effort to develop an objective methodology for estimating the cost of capital, it is worth at least a brief review here.

On a per share basis, the finite growth model used by Miller and Modigliani as the conceptual basis for their approach can be written as

\[ P = \frac{E}{k} + I \left( (\pi - k) \frac{b}{k} \right) T, \]

(1)

where \( P \) is the equilibrium price per share of common stock, \( E \) is expected earnings per share, \( I \) is expected new investment per share, \( \pi \) is the expected profit rate or return on equity, \( k \) is the required profit rate or cost of equity, and \( T \) is the number of years in which \( \pi \) is expected to be greater than \( k \). The theoretical interpretation of Equation (1) is fairly straightforward. The first term on the right-hand side, \( E/k \), represents the capitalized value of the firm's current expected earnings per share. The second term, \( I(\pi - k)b/kT \), represents the capitalized value of the firm's future earnings or growth prospects, on a per share basis. When the firm only earns a return on equity equal to the cost of equity (\( \pi = k \)), or when there is no net investment (\( I = 0 \)), the second term drops out, and the price per share is simply the capitalized value of the current earnings. If we assume that new investments are financed entirely from retained earnings, then Equation (1) can be rewritten as

\[ P = \frac{\pi k}{k} + \frac{\pi b T}{k} \left( (\pi - k) \frac{b}{k} \right) T, \]

(2)

where \( B \) is the book value of investment per share, and \( b \) is the earnings retention rate. Dividing Equation (2) through by \( B \), we get

\[ \frac{P}{B} = \frac{\pi}{k} + \frac{\pi b T}{k} \left( (\pi - k) \frac{b}{k} \right) T, \]

(3)

Equation (3) provides the basis for a single-equation regression model,

\[ \left( \frac{P}{B} \right)_{i} = \gamma_{0} + \gamma_{1} n_{i} + \gamma_{2} \pi_{i} b_{i} + \epsilon_{i}, \]

(4)

where the regression coefficients have the following expected values: \( \gamma_{0} = 0 \), \( \gamma_{1} = 1/k \), and \( \gamma_{2} = -(\pi - k) b/k T \). Theory would thus suggest that the cost of equity can be inferred from a regression equation in which the ratio of market value to book value is the dependent variable, and the earnings growth rate and the earnings-to-book ratio (return on equity) are the independent variables.

The estimates of the cost of equity reported by Miller and Modigliani using this approach were too low to be intuitively acceptable to most observers, and several comments published in response to their article suggested a number of possible explanations for this result. But one important flaw in the Miller-Modigliani approach seems to have been overlooked. One usually presumes in regression analysis that the slope parameters or regression coefficients are the same for all of the observations in the sample. When they are not, the model is misspecified. A. J. Boness and G. M. Frankfurter [1977] demonstrated that the single-equation model of Miller and Modigliani fails to conform to a fixed coefficient regression regime. In retrospect, this is not surprising, because the ratio \( \pi/k \), which is embedded in the \( \gamma_{2} \) coefficient, is presumed to be the same for all firms in the sample used to specify the model. But it is not, and variations in \( \pi/k \) are one of the principal sources of variation in \( P/B \) for firms of homogeneous risk.

What this means, then, is that there is some unique \( \gamma_{2} \) for each firm in the sample, and any attempt to fit a single-equation model to the data will result in obvious specification error. Subsequent versions of the single-equation regression approach were specified in such a way as to be immune to this particular criticism. Note that Equation (3) can be rewritten as

\[ \frac{P}{B} = \frac{\pi}{k} + \frac{\pi b T}{k} \frac{b}{k} \left( (\pi - k) \frac{b}{k} \right) T, \]

(5)

which can be taken as the basis of a regression model of the form

\[ \left( \frac{P}{B} \right)_{i} = \gamma_{0} + \gamma_{1} n_{i} + \gamma_{2} \pi_{i} b_{i} + \gamma_{3} \pi_{i} b_{i} + \epsilon_{i}, \]

(6)

where the regression coefficients have expected values of \( \gamma_{0} = 0 \); \( \gamma_{1} = 1/k \); \( \gamma_{2} = T/k \), and \( \gamma_{3} = -T = -\gamma_{2} \).

This avoids the specification error inherent in the Miller-Modigliani model because of cross-sectional variations in \( \pi/k \), but it is subject to its own specification error if \( T \) is not the same for all firms in the sample used to estimate the model. Since it seems unlikely that \( T \) is the same for all firms in any sample taken from the electric utility industry (some firms always seem to earn a return greater than the cost of capital, while others demonstrate a persistent difficulty in even earning the cost of capital), it appears that this approach is no less subject to inferences than the original Miller-Modigliani model.

The shortcomings of the traditional approach would seem to indicate that single-equation models of the valuation process require excessively restrictive assumptions to represent the valuation relationship
for all firms in a typical cross-section. A more robust approach, conceptually, would be to permit all of the relevant valuation parameters to vary from firm to firm. Note, for example, that Equation (3) can be rewritten as

$$P/B = \left(1 - kT\right)\pi + \left(kT\right)\pi^2$$.

(7)

Since $k$, $b$, and $T$ are all constants for an individual firm at any given time, it follows that the valuation relationship for an individual firm can be completely specified as a quadratic function of the expected profit rate with the roots 0, and $(k - 1/bT)$. The relationship between the market-to-book ratio and the expected profit rate may be termed a valuation line. From Equation (7) it follows that all valuation lines for firms of a given risk class pass through the points (0,0) and $(k,1)$ in $\pi$-$P/B$ space. Except for intersection at these points, the valuation lines for individual firms in a risk class will differ if their earnings retention rates differ.

In order to visualize the functional relationship between $P/B$ and the expected profit rate for firms with different retention rates, consider Figure 1. The figure shows hypothetical valuation lines for two firms, A and B. Both are in the same risk class, and their valuation lines pass through the points (0,0) and $(k,1)$. Except for intersecting at these points, the valuation lines differ. The valuation line for firm A, $f_A(\pi)$, is drawn on the assumption that its retention rate is greater than $B$’s. In the range of interest where $\pi > k$, the valuation line for firm A is steeper than the valuation line for firm B. Thus two firms of identical risk, expected earnings, and book value will have different market values if their earnings retention rates vary. This is shown in the figure for the case in which $\pi_A = \pi_B$. The firm with the higher retention rate, A, has a higher market value in relation to its book value than does B, the firm with the lower retention rate. The explanation is obvious. Since both firms have the same expected profit rate, the earnings growth rate of the firm with the higher retention rate exceeds the earnings growth rate of the firm with the lower retention rate. This higher growth rate is capitalized in a higher market price. The firm with the higher retention rate has a lower dividend yield and higher growth rate than does the firm with the lower retention rate. Their market values, even though their profit rates and book values are the same, differ.

This result — that the value of the firm is not constant across all firms of equivalent risk and expected profit rates — is not surprising, but the failure to keep this fact clearly in mind is probably the single most important explanation for the parameter instability in equity valuation models reported by Michael Keenan (1970). When valuation models are misspecified, regression coefficients can be unstable from one period to the next, even though statistically significant. For example, suppose that we are trying to estimate a single quadratic valuation model for two firms, again A and B, with different retention rates. Empirically, we only observe single points along the true lines, such as the points $(\pi_A, P/B_A)$ and $(\pi_B, P/B_B)$ in Figure 2. If we attempt to fit a single valuation line through these points as our empirical estimate of the valuation relationship, we end up with the line $CC'$. This empirically determined valuation line provides a perfect fit, insofar as the least-squares criterion is concerned, but it bears no relationship to the structural relationships that exist between $P/B$ and $\pi$ for either firm. In
While there are definite limitations to the single-equation econometric models that have been published so far, their contribution to the problem of developing objective models of the cost of equity should not be underestimated. They are in a class by themselves in terms of theoretical sophistication and empirical objectivity when compared to the methods usually employed by cost of capital witnesses. They permit inferences regarding the cost of equity from empirical data on the basis of hypothesized relationships rigorously specified in the form of econometric models, and the results that follow reflect the logic of the theory and method employed rather than the a posteriori musings of an analyst interpreting the data in a manner consistent with his "informed" judgment. They can be faulted in only two respects.

First, single-equation models make unnecessarily restrictive assumptions about the constancy of certain valuation parameters across firms of equivalent risk. This criticism is made knowing full well that the realism of assumptions is not a sufficient test of the credibility of a theory or model. Most economists agree that the credibility of a theory rests in its predictive or explanatory power, not in the realism of its assumptions. If a theory predicts some novel fact which is potentially untrue, then the theory is credible even though it rests on assumptions that are undeniable abstractions from reality. But as Popper pointed out, there are degrees of testability: "Every genuine test of a theory is an attempt to falsify it, or to refute it. Testability is falsifiability, but there are degrees of testability: some theories are more testable, more exposed to refutation, than others; they take, as it were, greater risks." This brings us to the second sense in which single-equation regression models of the cost of equity can be faulted. As it happens, single-equation models of the valuation process are only exposed to a modest possibility of refutation. Depending on the exact form of the model, they predict a relationship between the ratio of market value to book value and the ratio of earnings to book and the earnings growth rate. Were stock valuation predicated on something entirely different than present and expected future earnings, it is likely that the regression coefficients of such models would be statistically insignificant or of the wrong sign, and the theory would be rejected as untenable. But the regression coefficients are significant, and of the expected signs, and to this extent current valuation theory withstands the possibility of refutation. But this possibility of refutation is of only modest proportions because the theory, in its present form, only predicts the signs of the coefficients, and not their values or magnitudes. In the absence of more specific predictions, it is difficult to say that the assumptions embodied in the model are justified. As demonstrated by Figure 2, we cannot rule out the possibility of empirically significant results even
when the model is structurally misspecified because of improper assumptions about the constancy of certain valuation parameters across firms of equal risk.

In the next section we demonstrate an alternative approach designed to overcome these difficulties. We derive a multiple-equation model of the valuation process that permits certain conceptually relevant valuation parameters to vary from firm to firm. The misspecification that results from attempting to conform all firms to a single equation is avoided. Each firm has its own valuation equation. But more important from the standpoint of developing objective methods of estimating the cost of equity, the model developed in the next section leads to specific predictions about the values of the valuation parameters. The theory is thus exposed to a greater possibility of refutation, and this enhances our confidence in the findings that follow.

An Alternative Approach

Our goal is to derive an econometric model of the cost of equity not subject to the infirmities of models derived in the traditional mode. This is not difficult, given certain simplifying assumptions. The first assumption we make is that stock valuation is based on the infinite growth model of Myron Gordon (1962; 1974) rather than the finite growth model of the traditional approach.\(^{25}\) On a per-share basis, the Gordon model can be written as

\[
P = \frac{D}{k - \pi_B}, \tag{8}
\]

where \(D\) is the current dividend per share.\(^{26}\) Since \(D = (1 - b)\pi_B\), Equation (8) can be rewritten to show that

\[
P = \frac{(1 - b)\pi_B}{k - \pi_B},
\]

and therefore

\[
\frac{P}{\pi_B} = \left[\frac{-\pi_B}{k - \pi_B}\right] + \left[\frac{1}{k - \pi_B}\right] \pi. \tag{9}
\]

With certain additional assumptions, it can be argued that Equation (9) implies a valuation line that is a linear function of the expected profit rate, with a slope equal to the inverse of the dividend yield (since \(1 - \pi_B\) equals the dividend yield) and an intercept along with the \(\pi\)-axis equal to the earnings growth rate (because \(\pi_B\) equals the growth rate).\(^{17}\)

Figure 5 is a graphical representation of such a line, shown for comparison’s sake along with a quadratic valuation line from a finite growth model.\(^{28}\) This linear valuation line is an important conceptual tool in our quest for an objective method of estimating the cost of equity because it implies a testable relationship between market-to-book ratios and return on equity. Even casual empiricism is sufficient to establish that firms with high market-to-book ratios tend to have higher returns on equity than do firms with low market-to-book ratios. What is unique about Equation (9) is that it implies a structural relationship that is specific and testable. It suggests that in every homogeneous risk class of \(n\) firms, each firm will have its own valuation line with a slope equal to the inverse of the firm’s dividend yield, and an intercept along the \(\pi\)-axis equal to its expected earnings.
growth rate. The empirical task is to estimate valuation lines for each firm in order to determine whether the data support this hypothesis.

A single-equation model of Equation (9), such as would occur with a simple regression of market-to-book ratios on return on equity, is inappropriate because it presumes that every firm in the sample has the same dividend yield and expected growth rate. What we need is a set of structural equations that permit the slope and intercept of the individual valuation lines to vary from firm to firm. A model that accomplishes this objective may be derived in the following way. First, note that the valuation line of the ith firm in a sample of n firms may be written as

\[ m_i = (\pi_k + (k - \pi_b) \gamma) \left( \frac{P}{B} \right). \] (10)

Except for the subscripts, Equation (10) is nothing more than an algebraic transformation of Equation (9). It may seem awkward to derive for \( \pi \) as if it were a "function" of \( P/B \), but there are sound reasons for doing so. The conceptual relationship is deterministic: Solving for \( \pi \) in terms of \( P/B \) is nothing more than an exchange of axes, with \( \pi \) now being read along the vertical axis, and \( P/B \) being read on the horizontal axis. Stating the relationship in this way has certain empirical advantages. We can only estimate the expected profit rate; we do not know for sure what return on equity investors are expecting when they capitalize a firm's earnings. By specifying \( \pi \) as the dependent variable, we impound any measurement error in measuring \( \pi \) in the residual error term, and thus avoid an "errors-in-variables" problem.

Equation (10) is the first of three equations necessary to specify a structural model that allows for cross-sectional variations in dividend yields and expected growth rates. Next, we specify that

\[ (\pi_k) = \gamma_0^* + \gamma_1^* (D/P)_k, \] (11)

and

\[ (k - \pi_b) = \gamma_2^* (D/P)_k. \] (12)

These equations imply the following. If \( (\pi_b) \) is a good estimate of the expected growth rate, then \( (k - \pi_b) \) will equal \( (D/P)_k \), and \( \gamma_2^* \) will equal one. If \( (\pi_b) \) is a good estimate of the expected growth rate, and if the firms in the sample are of equivalent risk, then \( \gamma_2^* \) will equal \( k \), and \( \gamma_2^* \) will equal minus one. Thus, on theoretical grounds, we expect \( \gamma_0^* \), \( \gamma_1^* \), and \( \gamma_2^* \) to be \( k \), \( -1.0 \), and \( 1.0 \), respectively. Substitution of Equations (11) and (12) into (10) produces a reduced form model:

\[ m_i = \gamma_0^* + \gamma_1^* (D/P)_k + \gamma_2^* (D/B)_k + \epsilon_i. \] (13)

Equation (13) can either be estimated as an ordinary regression equation or as a simple linear model of the form

\[ \pi_i = \gamma_0^* + (D/P)_k + (D/B)_k + \epsilon_i. \] (14)

The latter method is appealing because it provides a direct test of the model and avoids any complications that might arise because of multicollinearity between \( (D/P)_k \) and \( (D/B)_k \). This approach is equivalent to the fitting of a plane in xyz space such that it has the slopes hypothesized a priori. The residual errors are then evaluated with reference to this plane rather than with reference to the obs plane, and the F-test from a simple analysis of variance provides the basis for accepting or rejecting the model. It is also appealing because the computations necessary to estimate \( \gamma^* \) using this approach are quite simple. It is only necessary to note that the linear model must go through the point of means; therefore,

\[ \gamma_2^* = \frac{\bar{\epsilon}}{\bar{(D/P)} - \bar{(D/B)}}. \] (15)

Given an estimate of the expected return on equity, estimates of the cost of equity follow simply and directly.

Estimation of the cost of equity using this approach depends on the development of a suitable estimate of \( \pi_i \), the expected profit rate or return on equity. The first estimator we might consider is the current actual return on equity. But stock prices reflect normal profitability. Poor performance in one quarter will not adversely affect the price of shares if the dividend is unaffected and if investors expect the firm's earnings to return to normal in the next period. This might suggest that an average of the firm's earnings or return on equity is an appropriate estimator. Indeed, empirical work has shown that average returns on equity are more highly correlated with market-to-book ratios than are current returns on equity.

But there is another way of estimating the return on equity investors are expecting that is superior not only empirically, but also conceptually. Note that earnings per share can be stated as dividends per share divided by the payout ratio:

\[ E = D/p, \]

where \( p \) is the payout ratio. We can impute the earnings per share being capitalized by investors by dividing the current dividend by a "normal" or expected payout ratio. While we still must assume something about investors' expectations, the possibility of error is less when estimating the expected payout ratio than when estimating the expected return on equity or the expected growth rate, for the simple reason that most firms follow a stable payout policy. And since estimates of the normal
return on equity determined in this way have a high degree of correlation with dividend-book and dividend-price ratios, the procedure is defendable on empirical grounds.

Using this approach, the expected return on equity for the $i$th firm is:

$$\bar{\hat{r}}_i = E(B_i) = \left(\frac{D_i}{\bar{P}_0}\right) / B_i .$$  \hspace{1cm} (16)

A firm's average payout ratio for the past several years is generally not the best estimate of its payout ratio in the future. Firms that have had unusually high payout ratios (such as Bostom Edison in recent years) will tend to have future payout ratios below the average of their recent past, and firms with unusually low payout ratios will tend to have future payout ratios that are above the average of their recent past. This tendency for individual averages to regress toward the grand mean is a common phenomenon, and there are statistical techniques for predicting future values based on this tendency of individual values to regress toward the grand mean over time. Briefly, superior predictions of future averages are made by weighting individual past averages with the grand mean, basing the weights on the variance on the individual averages and the variance across individual averages. In the study reported here, estimates of $\bar{P}_0$ were based on ten-year averages for individual firms adjusted toward the grand mean in accordance with this method. Estimates of the expected or normal return on equity were then computed using Equation (16).

The model was tested on a sample of 45 electric utilities for the years 1961–1976. The cost of equity was estimated from Equation (15), and $F$-tests for the simple linear model of Equation (14) were computed. The results are shown in Table 1, and they suggest a fairly strong degree of corroboration of the theory underlying the model. When the linear model represented by Equation (14) is fitted to the data, it generally provides a good fit, one that cannot be rejected on the basis of an $F$-test. Only in 1974 did the model fail to perform well, in the sense that for that year the error sum of squares of the model exceeded the total sum of squares, producing, in effect, a negative $R^2$ and a negative $F$. In retrospect, there appears to be a reasonable explanation for this result. The model assumes that investors are capitalizing normal earnings (as reflected in the current dividend, and the normal payout ratio) and in retrospect that does not appear to have been the case in late 1974. Earlier that year, Con Ed passed its dividend, and the market for electric utility stocks appears to have reacted as if it expected other utilities to follow suit, at least to the extent of reducing their dividends, if not to pay them entirely. Since our estimate of expected return on equity was based on actual dividends, the model would consistently overpredict the return expected by investors. This would account for the unusually high error sum of squares.

An effort to improve the model for 1974 was undertaken by excluding from the sample any firm with a dividend yield in excess of 13 percent, the supposition being that in the case of firms with extremely high dividend yields, the market was expecting negative growth, or a reduction in the dividend. This reduced the sample size to 31 and reduced the ratio of the error sum of squares to the total sum of squares considerably, although not enough to result in a positive $R^2$ or $F$. In any event, this improvement is sufficient to suggest that the cost of equity in late 1974 was more likely in the range of 13–14 percent than the range of 15–16 percent. Furthermore, the model's failure to work in at least one case should be considered a point in its favor. The fact that the model can be falsified gives us confidence in its general usefulness because it will be falsified, and the test statistics will lead us to reject the model, unless it is supported by empirical evidence. From the standpoint of objectivity, this is far superior to the methods of witnesses which always "work," which always confirm their prior judgment of what they think the return should be.

From the standpoint of the regulator, a very practical concern is the actual estimates of the cost of equity produced by the model. It is

Table 1. Summary of Empirical Findings

<table>
<thead>
<tr>
<th>Year</th>
<th>n</th>
<th>$\bar{P}_0^*$</th>
<th>Standard error</th>
<th>$R^2$</th>
<th>$F$-test</th>
<th>Preferred stock yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>1961</td>
<td>45</td>
<td>0.066213</td>
<td>(0.09615)</td>
<td>0.859</td>
<td>128.432</td>
<td>0.0468</td>
</tr>
<tr>
<td>1962</td>
<td>45</td>
<td>0.092599</td>
<td>(0.09005)</td>
<td>0.769</td>
<td>70.067</td>
<td>0.0444</td>
</tr>
<tr>
<td>1963</td>
<td>45</td>
<td>0.071798</td>
<td>(0.08686)</td>
<td>0.807</td>
<td>88.027</td>
<td>0.0447</td>
</tr>
<tr>
<td>1964</td>
<td>45</td>
<td>0.071024</td>
<td>(0.09160)</td>
<td>0.851</td>
<td>103.080</td>
<td>0.0444</td>
</tr>
<tr>
<td>1965</td>
<td>45</td>
<td>0.074863</td>
<td>(0.10747)</td>
<td>0.794</td>
<td>50.074</td>
<td>0.0472</td>
</tr>
<tr>
<td>1966</td>
<td>45</td>
<td>0.083656</td>
<td>(0.12466)</td>
<td>0.794</td>
<td>49.835</td>
<td>0.0546</td>
</tr>
<tr>
<td>1967</td>
<td>45</td>
<td>0.091867</td>
<td>(0.13058)</td>
<td>0.649</td>
<td>58.803</td>
<td>0.0625</td>
</tr>
<tr>
<td>1968</td>
<td>45</td>
<td>0.090955</td>
<td>(0.10728)</td>
<td>0.762</td>
<td>67.194</td>
<td>0.0632</td>
</tr>
<tr>
<td>1969</td>
<td>45</td>
<td>0.105103</td>
<td>(0.10728)</td>
<td>0.906</td>
<td>40.788</td>
<td>0.0734</td>
</tr>
<tr>
<td>1970</td>
<td>45</td>
<td>0.098562</td>
<td>(0.11592)</td>
<td>0.744</td>
<td>60.885</td>
<td>0.0736</td>
</tr>
<tr>
<td>1971</td>
<td>45</td>
<td>0.106064</td>
<td>(0.10756)</td>
<td>0.753</td>
<td>63.886</td>
<td>0.0715</td>
</tr>
<tr>
<td>1972</td>
<td>45</td>
<td>0.103062</td>
<td>(0.11959)</td>
<td>0.555</td>
<td>24.092</td>
<td>0.0732</td>
</tr>
<tr>
<td>1973</td>
<td>45</td>
<td>0.117028</td>
<td>(0.13428)</td>
<td>0.547</td>
<td>11.167</td>
<td>0.0804</td>
</tr>
<tr>
<td>1974</td>
<td>45</td>
<td>0.154352</td>
<td>(0.05006)</td>
<td>&lt;0</td>
<td>&lt;0</td>
<td>0.1017</td>
</tr>
<tr>
<td>1975</td>
<td>45</td>
<td>0.156077</td>
<td>(0.04950)</td>
<td>&lt;0</td>
<td>&lt;0</td>
<td>0.1012</td>
</tr>
<tr>
<td>1976</td>
<td>45</td>
<td>0.121949</td>
<td>(0.13765)</td>
<td>0.459</td>
<td>17.791</td>
<td>0.0941</td>
</tr>
<tr>
<td>1977</td>
<td>45</td>
<td>0.141903</td>
<td>(0.11108)</td>
<td>0.664</td>
<td>41.518</td>
<td>0.0953</td>
</tr>
</tbody>
</table>
interesting, therefore, to compare the estimates produced by the model with comparable yields on fixed income instruments such as bonds or preferred stocks. Such a comparison is not a rigorous test, in the statistical sense, but it can be suggestive. Table 1 lists the yield on high-grade preferred stocks for the years of the study, and Figure 4 plots both trends over the period examined. The very close correlation between the two certainly lends some credibility to the model. The mean difference between the estimated cost of equity and the preferred stock yields was about 2.6 percent.

![Figure 4. Estimates of the Cost of Utility Common Stock in Comparison to the Cost of Preferred Stock 1961-1976](image)

**Conclusion**

The purpose of this paper will be well served if it provokes some thought as to the role of myth and method in the estimation of equity capital costs. In any research, objectivity is predicated upon the development of hypotheses to guide the inductive or empirical process. Only if hypotheses are formally stated and tested can the influence of individual and subjective values be brought under the control of scientific discipline. All cost of capital witnesses, by virtue of their exposure to various kinds of data, formulate conjectures, theories if you will, as to the rate of return on equity required to attract capital. The mere concatenation of such data into tables, figures, exhibits, and so forth, does not transform conjectures into reliable inferences unless the conjectures are subject to a genuine possibility of refutation. This requires that witnesses state prior to data analysis which observable situations, if actually observed, would mean that their conjectures are refuted. While the logic of this approach to research is widely accepted among economists, it has not been widely employed by cost of capital witnesses in their development of methods to estimate the cost of equity. Indeed, their methods often seem to be designed with the opposite in mind; instead of seeking to minimize the influence of judgment and personal values on the outcome of their research, their methods often seem most remarkable for the numerous opportunities they afford the witnesses to exercise their "informed" judgment. In any event, it is precisely the lack of attention to the logic of the scientific method that makes it difficult for regulatory commissions to determine "a logical basis on which to connect [the] record data with the recommended rate of return."

An application of research methodology to the problem of estimating the cost of equity will generally suggest an econometric model of the valuation process. To be sure, an "econometric model" is not an automatic embodiment of truth. In a remark that deserves broad circulation, it has been said that if you torture the data long enough, it will confess, even to crimes it did not commit. The emphasis here on the use of econometric models is not made under the naive assumption that statistical methods are necessarily more "right" than other methods: Even econometric models can be made to "confess" to crimes they did not commit, if they are tortured long enough. The real advantage of an econometric approach is that it requires one to explore thoroughly the relevant market relationships conceptually, prior to data analysis, in order to develop specific and testable propositions which then, and only then, suggest the proper empirical approach. While no one can deny that decisions and judgments are made along the way, they can be made in such a manner as to bring
them under the control of scientific discipline. This is all we can ever expect of an econometric model, or of any method of estimating the cost of equity for public utilities; that it be a sufficiently precise statement of our theory to expose to the critical scrutiny of our judgments and assumptions. When this is the case, our conclusions are not subjective value statements, but are objective findings that reflect the logic of our method.

Notes

1. This paper weaves together several strands of thought that have already appeared, in some instances, elsewhere. See Basil Capeland (1977, 1978, 1979). Intellectual debts are not always easy to recognize, but Myron Gordon, Harry Trebing, John Peketic, Jim Herden, and Eugene Rasmussen were certainly sources of encouragement and insight. Others might be mentioned, but it is unlikely that anyone would want to share this writer's responsibility for what appears here, and even those who are mentioned might be a little embarrassed (although I hope not).

2. There is, discussed below, a literature relating to the development of econometric models of the stock valuation process as a basis for the estimation of equity capital costs, but this literature has had little impact on the actual practice of estimating equity capital costs before regulatory commissions.


5. J. K. Gwirrath (1975, p. 4.5).


8. The fact that it is taken seriously in regulation suggests something about the level of "expertise" of those who dominate the scene as "experts" on the cost of capital. It seems doubtful, in view of this widespread belief, that many of them have formal training in research methodology or the philosophy of science. If they do, they have done a remarkable job in preventing their knowledge of research methodology from having any influence on their estimation procedures. But rather than doubt their intellectual integrity by suggesting that they have ignored what they know in order to preserve the ability of predetermining the outcome of their research, I would suggest the possibility that they do not know, for the most part, what they are doing.


12. The Miller-Modigliani model was actually specified in terms of the firm's total value (stocks and bonds), but little is lost by specifying it on a per-share basis, and the subsequent studies have often done so.

13. The derivation of Equation (1) assumes that $T$ is small.

14. In either of these two situations, $k$ would equal the earnings-price ratio.

15. The assumption that new investments are financed entirely by retained earnings permits us to abstract from the effect of stock financing on the earnings growth rate. External financing is a significant factor in earnings growth only if prices are considerably above book value and the firm makes continuous use of external funds to finance new investment. But firms with high market prices relative to book value will have high profit rates and be able to fund most, if not all, new investment internally. Only firms with low market-to-book ratios and profit rates are likely to require extensive external financing, and in such cases external financing has little effect on earnings growth. The assumption is therefore pragmatically defensible. In any event, the assumptions results in a testable model that ties the empirical data quite well. It is therefore inappropriate to argue whether or not the assumption is "realistic."

16. Miller and Modigliani divided through by total assets on the basis of a need to avoid heteroskedasticity because of variations in firm size. But as we show here, one can justify a relationship between the ratio of market value to book value and the earnings-to-book ratio on purely theoretical grounds.

17. There are estimation problems with this model because of "errors-in-variables," Miller and Modigliani used two-stage least-squares to deal with this problem.

18. This explanation for the instability of the Miller-Modigliani model was not mentioned by Bonez and Frankfurter, who attributed the specification error to nonhomogeneity of risk, that is, variations in $k$. But while nonhomogeneity of risk would be sufficient to explain such instability, it is not the only explanation. Since there is substantial variation in $r$ among electric utilities, it seems likely that nonhomogeneity with respect to return, rather than nonhomogeneity with respect to risk, is the more plausible explanation.

19. The model derived by Higgins (1974) was of this form.

20. Note that $T$ is also assumed constant in the Miller-Modigliani model.

21. They will also differ if $T$ differs from firm to firm, but the discussion here focuses on the influence of retention rate and payout ratio on the valuation line.

22. Movements in $CC$ would produce results analogous to the "parameter instability" reported by Michael Keenan. As the example shows, the true parameters may actually be stable, and what we observe as parameter instability may be the result of specification error from attempting to use a single-equation regression model to model variation relationships that are too complex for a single-equation model.

23. When a theory is corroborated, it is possible to say, in retrospect, that the assumptions were not critical, that nothing was lost by abstracting from reality.


25. The models discussed above were based on finite growth models of the valuation process. The appeal of finite growth models may have some-
thing to do with "realism." But as demonstrated below, the substantive predictions of valuation theory with respect to the relationship between P/B and π are not materially different from the two versions of the theory over the relevant range of interest. In other words, the "realism" of the finite growth model does not actually offer any more insubstantial value process than what we achieve with the more "unrealistic" infinite growth model.

26. This version of the Gordon model, as does the version of the finite growth model discussed above, abstracts from the effect of external financing on the earnings growth rate. See note 15 for a defense of this assumption.

27. The assumptions under which it can be argued that P/B is a linear function of π are discussed by Eugene Brigham and T. A. Bankston (1973). But even when their assumptions do not exactly hold, the slope of the linear approximation is equal to the slope of the true nonlinear function when π = 4, and it deviates significantly from the true slope only when π is substantially greater or less than 4. Furthermore, for the general case when π ≠ 4, use of the linear approximation will overstate the cost of equity, but only by an immeasurable amount.

28. If the finite growth model is thought to be superior to the infinite growth model because of more "realistic" assumptions about the nature of future growth possibilities, Figure 3 shows that there is not any particular justification for this point of view. Over the relevant range of interest, the mathematical properties evidenced by the slopes of the two models are nearly the same, and it is doubtful that our skill in measuring the cost of equity is sufficient to notice any differences of statistical significance.

29. Several studies have been published in Public Utilities Fortnightly demonstrating empirical relationships between return on equity and market-to-book ratios that use these relationships, in some cases, to infer the return necessary to result in a market price equal to book value. See L. S. Hyman (1974), J. J. T. Baquet (1975), and T. F. Simson and G. Spears (1975). The regression models employed in these studies, and the results produced by them, are misleading because they presume a constant functional relationship (a single slope) between P/B and return to equity for all firms. But as should be clear from our discussion above, that is not the case. The functional relationship varies from firm to firm depending on the payout ratios and the dividend yields of the individual firms.

30. Charles Stein (1956) is credited with showing that the mean is not always the most efficient estimate of future values. Bradley Efron and Carl Morris [1974] provide a recent survey. Efron's method, with applications to time series, and Morris [1977] presents a less technical but sufficient (and excellent) discussion of the method.

31. More recently, the model has been applied to a larger sample, comprising most market traded electric, with even better success. (Because of the increase in the degrees of freedom, the estimates produced from the larger sample have lower variances. The larger samples thus provide more precise estimates than the estimates reported here.)

32. An important test of any methodology is its acceptance by regulators. It is significant, therefore, that the methodology described here has been "tested" in a rate case and found to be acceptable. In its opinion, dated February 1979, in a case involving Southwestern Bell Telephone Company, Docket U-2916, the Arkansas Public Service Commission adopted the testimony of this author on the issue of rate of return and the cost of equity. The testimony in that case was predicated upon an empirical analysis similar to that described here, but with a data base consisting of telecommunications companies.

33. This paraphrases a remark attributed by Dennis Logue [1977] to Bubert Gutkein.

References


thing to do with “realism.” But as demonstrated below, the substantive predictions of valuation theory with respect to the relationship between P/B and π are not materially different between the two versions of the theory over the relevant range of interest. In other words, the “realism” of the finite growth version weighs, does not actually offer any more insight into the valuation process than what we achieve with the more “unrealistic” infinite growth model.

26. This version of the Gordon model, as does the version of the finite growth model discussed above, abstracts from the effect of external financing on the earnings growth rate. See note 15 for a defense of this assumption.

27. The assumptions under which it can be argued that P/B is a linear function of π are discussed by Eugene Brigham and T. A. Bankston [1973]. But even when their assumptions do not exactly hold, the slope of the linear approximation is equal to the slope of the true nonlinear function when π → 0, and it deviates significantly from the true slope only when π is substantially greater or less than 0. Furthermore, for the general case where π > 0, use of the linear approximation will overstate the cost of equity, but only by an immeasurable amount.

28. If the finite growth model is thought to be superior to the infinite growth model because of more “realistic” assumptions about the nature of future growth possibilities, Figure 5 shows that there is not any particular justification for this point of view. Over the relevant range of interest, the mathematical properties evidenced by the slopes of the two models are very nearly the same, and it is doubtful that our skill in measuring the cost of equity is sufficient to notice any differences of statistical significance.

29. Several studies have been published in Public Utilities Fortnightly demonstrating empirical relationships between return on equity and market-to-book ratios that use these relationships, in some cases, to infer the return necessary to result in a market price equal to book value. See L. S. Hyman [1958], J. A. Jaquette [1976], T. G. Mars [1975], and G. Spears [1975]. The regression models employed in these studies, and the results produced by them, are misleading because they presume a constant functional relationship (a single slope) between P/B and return on equity for all firms. But as should be clear by now from the discussion above, that is not the case. The functional relationship varies from firm to firm depending on the payout ratios and the dividend yields of the individual firms.

30. Charles Stein [1956] is credited with showing that the mean is not always the most efficient estimate of future values. Bradley Eron and Carl Morris [1971] provide a recent review of Stein’s method, with applications.

31. More recently, the model has been applied to a larger sample, comprising most market traded electrics, with even better success. (Because of the increase in the degrees of freedom, the estimates produced from the larger sample have lower variances. The larger samples thus provide more precise estimates than the estimates reported here.)

32. An important test of any methodology is its acceptance by regulators. It is significant, therefore, that the methodology described here has been “tested” in a real case and found acceptable. In its opinion of 15 June 1979, in a case involving Southwestern Bell Telephone Company, Docket U-2916, the Arkansas Public Service Commission adopted the testimony of this author on the issue of rate of return and the cost of equity. The testimony in that case was predicated upon an empirical analysis similar to that described here, but with a data base consisting of telecommunications companies.

33. This paraphrases a remark attributed by Dennis Logue [1977] to Bulent Gulenke.

References


Economic and Financial Models as Regulatory Tools

Robert E. Wayland

The complexity and number of issues faced by regulators suggest that economic and financial models could make a significant contribution to the quality of regulatory decisions. Experience with models to date, however, indicates that their contribution, while positive, has been disappointing relative to their potential. Some of the reasons economic and financial models have not made a greater contribution to regulation include: models’ failure to address issues of concern to regulators, regulators’ skepticism of model results or assumptions, and utilities’ resistance to the implementation of models in the regulatory process.

Regulatory modeling (whether computer simulations or mathematical formulas) involves reducing complex phenomena to their essential and most significant elements in order to analyze the economic and regulatory environment and to make informed estimates of future conditions. This paper reviews the contribution and the problems of three types of models that have been applied to the regulatory process for the electric utility industry: demand forecasting, cost of capital, and corporate finance models. Since the recent passage of the Public Utility Regulatory Policies Act (PURPA) will require more extensive regu-
Demand Forecasting Models

Many demand forecasting models are not designed to address issues which concern regulators, who may find demand forecasts by themselves of very little use. Forecasts are important to regulators and utilities because of their relationship to the capacity planning decisions and the consequent financial requirements. Therefore, a demand forecast is an important part of a firm's rigorous assessment of the future market conditions it is likely to face. But a single forecast reflecting only one set of assumptions is generally inadequate to assess policy alternatives or economic uncertainty. Since utility management share authority with their regulators, the forecast should address areas of concern to regulators. In addition, the uses of a demand forecast extend beyond the power siting decision and include security issues and ratemaking applications. The design and scope of demand forecasts should recognize the strategic needs and concerns of all parties involved.

Regulators, like all decision makers, are interested in evaluating alternatives. For example, regulators are interested in alternative rate structures such as time-of-day or life-line rates, but utility forecasts are infrequently capable of assessing the impact of alternative rate structures on the rate of growth in peak and average load. Despite the obvious concern of regulators about the levels of future rates, utility forecasts often make unrealistic assumptions about them. Many utility forecasts do not ensure that the rates are consistent with the level necessary to support the capacity additions that are required to meet the forecast demand. Although the federal government is supporting intensive research into competing energy sources such as solar, and while many environmentalist intervenors are promoting the use of alternative energy sources, most utility forecasts fail to consider such alternatives. This failure to address issues of comparability and regulatory decision making usefulness of the utilities' forecasts and promotes the proliferation of competing forecasts by other parties.

Utility resistance to implementation of forecasting models ranges from a lack of cooperation in supplying data to a reluctance to perform the type of experiments and load research necessary to gain greater understanding of the electricity market. Probably no industry of comparable size knows as little about its customers as the electric utilities. In recent years, the government, not industry, has taken the lead in load research and rate structure experimentation. Failure to pursue rate and load research aggressively has contributed to the erosion of management independence.

Since utility models often do not address issues of concern to other parties, intervenors and agency staffs have developed models that serve their interests. The result is often an overlong debate focusing on econometric minutiae rather than a general assessment of the reasonableness of the utility's position. When competing forecasting models produce widely differing results, it is difficult for the regulator to assess their contribution to the regulatory process.

One of the regulatory responses to help eliminate the confusion resulting from many competing forecasts is to strengthen the role of the state in preparing estimates of future demand for electricity. Although California's Energy Office has probably gone further than most, several states have acquired relatively sophisticated forecasting capabilities. Some states, such as Ohio, have attempted to establish guidelines for the preparation of demand forecasts. The increased involvement of the states in forecasting, rather than in reviewing the work of the utility, is a potential threat to the freedom of utility management to function in their strategic planning role.

Cost of Capital Models

One of the most important regulatory applications of models is the estimation of the required rate of return or cost of common equity. In recent years, two empirical applications of modern capital theory, the discounted cash flow model (DCF) and capital asset pricing model (CAPM), have been applied to the problem of estimating the cost of common equity. These models are used to interpret investors' expected return and the perceived risk of securities based on market data. Regulators are concerned with the measurement of investor expectations and the perception of risk, in part because regulatory actions influence investor expectations and risk perceptions. Regulators are often justifiably skeptical of rate of return recommendations based on DCF and CAPM because the empirical methods used to apply the theoretical principles often appear overly sensitive to the selection of time periods or other measurement variables.

Most regulatory applications of DCF and CAPM rely almost exclusively on historical data and return relationships to measure the cost of common equity. For example, proponents of DCF usually determine the expected dividend growth by trending or averaging past rates of increase in earnings per share, dividends per share, and historical retention rates and earnings on book equity. Analysts employing CAPM usually determine the risk premium solely on the basis of
historical returns relative to the market and a "risk-free" security — usually government securities.

Historical data can be useful in establishing relationships among variables, but reliance solely on historical data is clearly limited. The use of such data inevitably forces the analyst to deal with the ex ante/ex post problem. Prices reflect investors' expectations, which may or may not be realized.

CAPM can be useful in establishing a firm's relative risk position in order to augment other forms of analysis such as DCF. One problem with CAPM, however, is the usually implicit assumption that relative risk, measured by beta, is stable and that historical betas can be projected into the future. Although most studies indicate that the assumption of relatively stable betas is reasonable for portfolios, this is sometimes questionable for the marginal firm, as the experience of numerous rating changes in the 1974–1975 period indicates. Furthermore, many of the policies considered by regulatory agencies can affect the perceived riskiness of firms. Therefore, more work on the determinants of risk perception as they relate to the traditional financial measures would be useful to regulators in assessing their policy options.

The failure to address more directly the formation of expectations and perceptions of risk sometimes leads regulators to question the appropriateness of these models in determining a fair return figure. Often the regulator is aware of current or anticipated conditions or events that influence the price of a company's stock but that do not seem to be reflected in the data used by DCF or CAPM proponents.

The skepticism of regulators toward economic models to estimate the cost of common equity is indicated by the fact that the DCF model is still only the second choice among regulators. Comparable earnings — nearly devoid of theoretical content — is the leading methodology (see Table 1).

The utility industry's feelings about DCF were illustrated by its reaction when the Office of Economic Analysis of the FPC (now FERC) recently proposed to determine "zones of reasonableness" on equity returns on the basis of a DCF analysis. When the initial recommendations were released for public comment, the utility industry voiced strong objections. The effect of the utility industry response was that the FPC put the matter aside for an indefinite period.

The most legitimate arguments usually advanced by the industry questioned the wisdom of endorsing any particular methodology. Given the state of the art in empirical applications of capital theory, this reservation has a great deal of merit.

Table 1. Method Favored in Rate of Return Determination

<table>
<thead>
<tr>
<th>Method</th>
<th>Number of states</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discounted cash flow</td>
<td>20</td>
</tr>
<tr>
<td>Comparable earnings test</td>
<td>24</td>
</tr>
<tr>
<td>Earnings-price ratio</td>
<td>7</td>
</tr>
<tr>
<td>Midpoint approach</td>
<td>3</td>
</tr>
<tr>
<td>Other</td>
<td>19</td>
</tr>
</tbody>
</table>


Corporate Finance Models

The recent development of generalized corporate finance models should enable regulators to improve the quality of policy analysis. Corporate finance models are useful in determining the revenue and profit implications of various ratemaking alternatives. For example, a finance model can be used to determine the revenue, cash flow, and external financing implications of various regulatory treatments of deferred income taxes and investment tax credits, including construction work in progress (CWIP) in the rate base or adopting different rates of return.

A corporate finance model is simply a mathematical description of the financial and accounting relationships underlying the basic financial statements. These models can generate pro forma income statements, balance sheets, and sources and uses of funds statements, given forecasts of construction and capital expenditures and historical data on the firm.

One serious problem in applying finance models to the regulatory process is that many have come to believe that they provide an easy route to answers to complex questions. Many potential users of these models appear to believe that they simply have to load the data, crank up the model, and wait for answers to difficult questions to pop out. At a recent conference, Harry Trebing expressed the fear that these models might become a placebo and divert attention from addressing many of the underlying analytical problems faced by commissions. Properly used, finance models can facilitate the analysis of important economic issues. Finance models cannot tell us if the income redistribution resulting from time-of-day rates is good or bad. Nor can they tell us if the cost of attaining environmental quality is worth the effort. But they can help measure the expected costs and financial implications of these policies and thus contribute to the quality of regulatory analysis.
The use of corporate finance models such as the Regulatory Analysis Model (RAM) and Empire State Power Resources, Inc. (ESPR) model to set rates is limited by several factors: statutory restrictions on future test years, the shortcomings of analytical models used to estimate important inputs such as future demand and cost of capital (discussed above), limited staff capability to handle large complex computer models, and industry resistance to exposing its internal forecasts.

The projections of a corporate model are no more accurate than the inputs provided by analytical models. For example, a poor demand forecast will result in inaccurate revenue forecasts and inappropriate estimates of financing requirements. Nevertheless, a corporate finance model can make a contribution to improving the quality of demand forecast and cost of capital models simply by demonstrating the sensitivity of a firm's financial health to such things as load growth.

Most states do not possess the necessary foundation to implement large integrated computer-based corporate models. For example, a recent NARUC survey indicated that only 12 states utilized computers in their data processing in 1976. In addition to the lack of computer facilities, the wage structure of many state commissions precludes the hiring of adequately trained staff to implement advanced automated analysis techniques.

Utilities have sometimes resisted the adoption of forecasts generated by corporate models in regulatory proceedings. One of the most common reasons is the reluctance to disclose their own forecasts lest they be liable for stockholders' suits. Many firms also appear reluctant to expose their management's planning to criticism, especially given the adversarial nature of regulatory hearings.

Note

1. FPC Docket No. RM77-1.

Comments

John L. O'Donnell

Ronald Melicher does a masterly job of summarizing the basic assumptions and key concepts of modern financial theory dealing with the measurement of risk and returns on assets traded in an efficient market. His discussion unfolds in a logical fashion by breaking down total investment risk into the riskless rate and a risk premium appropriate for each individual asset. He uses this basic framework to review some of the best scholarship covering risk and return relationships of debt and equity securities. Both theory and empirical evidence support the central notion that, over the long pull, risk and return are linearly related.

Of course, many difficulties remain when it comes to translating generalities into the specifics of estimating the cost of equity capital for an individual utility. For example, Melicher correctly observes that the riskless rate is usually taken to be the yield on government bills or bonds which have no default risk. This still leaves losses in purchasing power due to unexpected changes in the value of money unaccounted for. A 10 percent loss in real terms is still a loss. In practice, there may be no such thing as a risk-free rate.

There are also strong arguments against using variability of market
returns as a surrogate for measuring risk premiums. Returns may not be normally distributed, and the size of a standard deviation is closely related to the particular investment holding period selected. For example, low short-term variability may be associated with a long-term increasing or decreasing trend in the market value of an investment. Perhaps the best defense for using the standard deviation of market returns is that it lends itself to statistical manipulation, and we have nothing better to offer. Even so, convenience born of necessity should not be unnoticed. We must also note that regardless of how sophisticated our statistical manipulations of historic data, the past remains a dangerous platform from which to forecast future events.

Portfolio theory weaves together the concepts of an efficient market and variability of returns into a subtle explanation of how asset values are determined. Once again, Melcher does a brilliant job of presenting the elements of the theory, culminating with a concise explanation of the capital asset pricing model (CAPM). In this model the only intelligent investment strategy is to form efficiently diversified portfolios. As a result, the market only rewards non-diversifiable risk, and this is measured by a portfolio’s beta. It follows that the action of forming efficient portfolios determines the market yields on individual securities, and portfolios having the same beta carry identical risks and hence promise identical returns.

Portfolio theory is intellectually satisfying. It constitutes a powerful pedagogical tool for explaining the way capital assets are priced in a highly competitive market constrained by a number of demanding assumptions. Many important corollaries flow from the theory. By drawing attention to the relationship between risk and return, it generates a more meaningful evaluation of portfolio management. It reveals that many financial managers were not clever at all, but simply are “gunslingers” who turned in high returns by exposing their funds to inordinate risks. As already noted, portfolio theory stresses the vital role of intelligent diversification. Since only non-diversifiable risk is rewarded, it is quite pointless to try to beat the market by looking for undervalued securities. This conclusion has caused great consternation among security analysts, who recognize a direct assault on their livelihoods. In fact, the University of Chicago has abolished all its traditional courses in security analysis.

In recent years, empirical tests of CAPM as applied to portfolios have revealed some worrisome differences between what the model predicted and the numbers revealed. Nevertheless, dedicated believers in the new church have extended the theory from diversified portfolios to individual securities. It is at this point that betas, alphas, and epsilons of regression analysis become of keen interest to regulators and utilities. Obviously, if a security’s beta is a reliable index of relative risk, then we have found the key to unlock the rate of return riddle.

A new beta boom industry has arisen, but the flood of different calculations demonstrates that individual betas are unstable, typically have high standard errors and low $R^2$, and often yield different values depending upon the method of calculation. It is unfortunate that Melcher pursues the use of beta in regulatory proceedings by reference to different sources for his betas rather than using exclusively either Value Line or Merrill Lynch.

Many students feel that individual betas underestimate the current risks confronting utilities. The explanations advanced by this school involve fairly complicated statistical theory. Melcher also seems to think some understatement of risk is occurring and concludes that regulatory risk requires that a special premium be added to the relative risk revealed by the beta of an individual utility. This solution leaves two problems. First, it denies the theoretical foundation upon which the CAPM rests. No reason is advanced to explain why regulatory risk cannot be diversified away by combining utility stocks in efficient portfolios with stocks of nonregulated firms. If, indeed, regulatory risk cannot be diversified away, then portfolio theory decrees it will be impounded in beta with all other non-diversifiable risk. Second, if we admit some risk premium must be added to beta, then we are left wondering how this extra unit of risk can be determined. Melcher forces us to the conclusion that a great deal remains to be done before the CAPM can be used as a reliable tool to determine the fair rate of return.

In many respects the central theme of Basil Copeland’s paper resembles the familiar dispute about normative versus positive economics. It begins with a peremptory protest against experts who allegedly use unscientific methods to estimate the cost of equity capital in regulatory proceedings. In Copeland’s opinion, the ultimate proof of a fraudulent expert is one who appeals to his “informed opinion” as the final basis for making estimates. The author concludes that the resulting confusion leads commissions and courts to be more liberal in their awards than the facts warrant. He then develops his own solution to the problem of estimating the cost of equity, which he claims is both scientific and objective. This solution turns out to be a sophisticated econometric model designed to obviate the methodological and mis-specification errors of similar efforts by other distinguished students of financial theory.

Copeland arrives at his model by tracing the evolution of valuation models starting from a single-equation regression advanced by M. H. Miller and Franco Modigliani in 1966. He correctly notes that the
model produces answers that are intuitively too low, but fail to recognize that this important conclusion admits the key role of "informed judgment" when evaluating the output of econometric models. The technical deficiencies of single-equation models are pinpointed and lucidly explained. Among other things, such models are usually unspecified because they naively assume that the regression coefficients are equal for all sample observations. Even so, the models can produce empirical results that are statistically significant but at the same time quite meaningless.

Copeland claims all these difficulties can be resolved if we are prepared to make a number of simplifying assumptions, such as infinite constant growth and a market valuation process ensuring equity values are the inverse of dividend yield. Within this framework, his model attempts to relate the market-to-book value of any stock to the cost of equity capital and the expected rate of return. Presumably, the model assumes frequent disparities between the cost of capital and investor expectations about future returns. These expectations are claimed to be measurable by capitalizing dividends by a "normal" or expected payout ratio.

The author has generated yet another interesting econometric model which can undoubtedly hold its own against many similar competitors which continue to make their appearance. Even so, perhaps the biggest unintended contribution to scholarship made by this paper is that it serves as a warning about the overly enthusiastic use of statistical methods in finance.

Robert Wayland tackles three major topics in very general terms by emphasizing the deficiencies of demand forecasting, cost of capital, and corporate finance models. It is hard to quarrel with his observations because limitations of time and space prohibit him from expanding upon technical details or proposing alternatives. He is convinced most models fail to address relevant questions and that they are viewed with skepticism or outright hostility by the parties they are intended to serve. If this is true, it is difficult to understand how regulation works at all.

Wayland complains that demand forecasts are usually too limited in scope. From one point of view this may be true, because all major decisions are interrelated. However, comprehensive models are inevitably very complicated, and there is little evidence they are any more accurate than are simple extrapolations. A major bone of contention is management's reluctance to share its forecasts with regulatory authorities. It may be that some arrangement facilitating a joint planning effort by companies and regulators is needed, although this would lead regulatory agencies into the company decision-making process and oblige them to accept more direct responsibility for policies adopted. This is the sticking point that makes management extremely reluctant to share its own forecasts with outsiders. Wayland anticipates that this reluctance will induce regulators to make their own forecasts with which to encroach on managerial strategic planning. In short, management is likely to have its prerogatives eroded either way. It is also possible, of course, that regulators and other groups would like to deliberate upon a spectrum of plans without accepting any responsibility except the privilege of exercising hindsight.

Turning to cost of capital models, Wayland notes the dangers of using historic data to predict future yields. He feels too little attention has been given to the formation of expectations and the role they play in setting yields. Modern financial theory teaches that security prices are a random walk, which means that everything the market expects about the future is impounded in current prices. Thus, current expectations are included in most models, but by definition uncertainty cannot be quantified. As a result, the best any model can do is shed light upon what the cost of capital was and is rather than what it will be.

Much the same conclusion is applicable to financial modeling. The output of any financial model of the firm cannot, except by chance, predict future prices. Moreover, a model's output cannot be better than the accuracy of the data used. As noted by Wayland, with suitable skill and care two objectives can be achieved. First, models of the firm can provide a type of sensitivity analysis revealing what could happen to specific financial variables under a well-defined set of conditions. For example, if the capital budget is increased by a given percentage, the model can help to bracket the likely impact on external financing needs. Second, models can also trace the course of past events to explain how the firm got where it is. This type of audit exercise is useful in evaluating policy decisions and could serve as a learning tool. We must agree with Wayland that models cannot make social judgments and are therefore of much more limited value than some enthusiasts seem to think.
Comments

Joseph M. Quigley

The subject of the reasonable rate of return for the common equity portion of capital would have little romance if all agreed on its determination. There is no question that there has been a change in the nature of the capital markets in which utilities must attract large amounts of capital to continue providing good service. Investors are attaching a great deal of importance to regulatory procedures that will allow utilities a reasonable opportunity to earn their permitted rates of return.

A large portion of the problem in obtaining capital for new construction results from the inflation rate, which has tripled between the 1960s and 1970s. Until, if ever, inflation abates, utilities must raise more and more dollars for the same unit output of plant. Environmental requirements have also added significantly to the need for new capital.

Most of my comments will be directed to Basil Copeland's paper. He begins with the statement that regulation should permit regulated firms to earn a return on rate base equal to the cost of capital. I have difficulty with this statement since I have always been under the impression that utilities were to be permitted a fair rate of return on rate base. These statements are not the same. Simply because a utility earns its historical cost of capital does not automatically mean it is earning a fair rate of return. Ronald Melcher correctly referred to the Bluefield and Hope cases, which are still used in regulation today. Briefly stated, they conclude that a fair return should be adequate to maintain existing capital and thus must be commensurate with the returns earned by other firms with corresponding risks.

Copeland's thesis proposes that rate of return computation is a science. He states that all evidence submitted by witnesses on the use of comparable earnings tests, discount cash flow, risk premium approaches, price-earnings ratios, and so forth, is just so much arithomancy.

Regardless of Copeland's opinions, we seem to have muddled our way to rate of return conclusions even if they are determined in the incorrect way. I do not believe that the results of these rate of return determinations are appalling simply because all witnesses do not arrive at the exact same number. The use of judgment has been and will continue to be a significant factor in determining a reasonable rate of return, even though objective criteria are the primary factors used by rate of return witnesses in determining the return requirement.

After presenting a series of formulas, a conclusion is reached by Copeland that two firms with identical risks, expected earnings, and book value will have different market values if their earnings retention rates vary. This is extended to conclude that if two firms have the same expected profit rates, the earnings growth rate of the firm with the higher retention rate will exceed the earnings growth rate of the firm with the lower retention rate. This conclusion, in my opinion, may or may not be true.

It cannot always be concluded that the higher payout company will be capitalized in the marketplace with a lower price-earnings ratio. In fact, the opposite may be true. If the higher payout firm can sell additional shares in the marketplace at a price which is advantageous for the existing shareholders, these investors may benefit more from the sale of securities as a means of raising capital than by having the firm retain the earnings. Sometimes we tend to forget that management's job is to maximize the earnings for the shareholders out of any given amount of revenue. In theory, no shares should be sold by a company unless the shareholders are benefited. Of course, a utility must also remember its responsibility to serve its customers adequately.

Assuming new shares are sold only when it is advantageous to the shareholders, it becomes more difficult to maximize earnings for new shareholders once they purchase company stock. However, the job of management is to maximize their return, also, once new shares are
sold, even though the maximum price is sought by the company when the shares are marketed.

Copeland states that those who work in regulation are already familiar with the proposition that there is a positive relationship between market-book ratios and return on common equity. I believe this to some extent, but the market-book ratio more likely is determined by how the total market perceives the firm's future outlook rather than its past performance. Obviously, those companies with higher market-book ratios are also most likely to have higher returns on equity, but this is also because the market perceives their prospects are better.

A few words about models are in order. They have been developed for various aspects of business these days, and one may hope they are useful. Time and effort are spent in building them so that management can measure quickly and easily the effects of making changes in assumptions.

Robert Wayland stated that although most companies have rather sophisticated forecasting systems, the utility industry resists exposing its internal financial forecasts. Companies do accept the possibility that lawsuits may be filed against them for choosing to forecast their earnings. The demands to make forecasts are growing, and these come not only from intervenors in rate cases, but also from the Securities and Exchange Commission. Despite these pressures, my company does not make public such earnings forecasts. A company that chooses to do so has the responsibility to continue updating its forecasts once facts and circumstances make the previous forecasts obsolete. This means that most companies, including utilities, would issue new forecasts a number of times each year. Some observers request multiyear forecasts from companies. If these are provided, the problems referred to would be compounded.

It is true we are able to revise forecasts very quickly through the use of computer models. The primary benefit of such models is the ability to examine a number of "what if" situations. When forecasts were prepared manually, time constraints prevented us from examining all of these. However, a forecast prepared using a computer model with one set of assumptions does not necessarily provide a better projection of the most likely earnings for future years, only a faster method.

In reviewing the three papers, I was very impressed with their in-depth analysis. I would like to thank each of the authors for contributing to our knowledge of new regulatory concepts and tools.

Comments

Clement T. Loshing

Before I discuss Robert Wayland's paper, I would like to give my strongest endorsement to another, "Risk Measurement and Rate of Return under Regulation," by Ronald Melicher. It is an outstanding one for several reasons. First, it contains an excellent statistical quantification of the various types of risk as they relate to a fair return. It also points out that which is not currently quantifiable under known statistical techniques. The second and more predominant reason for directing attention to this paper is that it is a "sleeper." By that I do not mean that reading it will cure insomnia; rather, there is benefit to be gained from reading the paper that would not be obvious from its title.

I highly recommend that any individual of a nonmathematical bent, interested in understanding or obtaining a sound knowledge of alphas and betas as used by financial analysts, read this paper. It presents the clearest exposition on that subject that I have seen. Melicher eschews the statistical mumbo jumbo that so often befogs and permeates such writing. Even when he has to use Greek letters, he spells them out in English. I am confident that readers will find the time spent reading his paper well worthwhile.

I first knew Bob Wayland when he was Chief Economist, Public
Utilities Commission of Ohio (PUCO), a position he held for some time. We tangled in adversary proceedings and cooperated in implementing improvements in the Ohio regulatory process. The subject of Wayland's paper. Although we did not always agree, we could discuss our differences with mutual open-mindedness. He was very creative and suggested many innovations for improving the regulatory process. Many of his proposals were good. I regard his departure as a loss to the regulatory climate of Ohio.

Now that I have praised him, let me offer my critical comments. Once again, I am more in agreement with the content of his paper than I am in disagreement with it. Most of the disagreement arises from a difference in his reliance upon theory and my recognition of the realities of the real world. I shall point out this nuance of viewpoints by citing seven specific points raised by Wayland.

Internal Consistency

Wayland charges that forecasts are not internally consistent, that is, the rate level may not support the level of investment implicit in their forecasts. I do not believe this is a valid criticism of the entire electric utility industry. In an industry so vast, there are bound to be companies of varying levels of management competence. The symptoms that Wayland complains about are not generic to the entire industry; rather, they are the mark of an inadequate or poor forecast and probably indicate a poorly managed utility. In such a situation, regulators often apply the same regulatory logic to both ends of the spectrum of management competence. In fact, their actions often inhibit innovative and creative management by failing to provide an incentive, while financially bailing out the mismanaged utility by allowing each the same common rate of return. Furthermore, Wayland suggests that the regulators and utility managers should have a mutually participative single forecast. I shudder at this, and I will discuss it later.

Alternative Energy Sources and Experimentation

Wayland maintains that forecasts do not reflect the impact of alternative energy sources and rate structure experimentation. This is certainly partly true. I would like to give some of the reasons why this has occurred and why it should not be cause for concern. Until very recently, there has been a general unwillingness on the part of the industry to reduce long-term forecasts on the speculation that solar energy or wind power will be economically and sufficiently available or attractive to displace a significant amount of load within ten or fifteen years, which is our minimum planning horizon.

There are many reasons for justifying a bias in utility load forecasts toward the higher side of the probable range than toward the lower side. In general, there is no excess of power plant capacity. In fact, unless there is a slackening of even the reduced current rate of load growth, expectations are for a nationwide capacity shortage in the early 1980s.

The following are some compelling reasons for this forecasting posture. First, any temporary excess capacity provides some material reduction in operating costs. The newer plants have a better heat rate, are environmentally clean, and can displace high-cost fossil units and scarce oil generation. Second, because of the rather tight capacity situation around the country, partly caused by retrofitting for environmental purposes and other factors cited herein, any excess capacity can probably be sold at rates that compensate for that investment. Third, that capacity will be cheaper to install now rather than later because of skyrocketing construction costs at a time when the rate of inflation is often greater than the cost of money. Fourth, the risk of excess capacity is more economically digestible than would be the monetary, social, and customer-regulatory reactions to the increased costs and deterioration in service if we were to underestimate our load-capacity relationship and have a shortage. Fifth, there is an inevitable delay in getting capacity in service. Thus, excess capacity probably will not materialize.

Rate Experiments

Wayland claims that the impact of rate experiment changes is not currently being reflected in future forecasts. Until we see more solid evidence that U.S. consumers are willing to change their living patterns, prudence dictates a more cautious course, but developments should be monitored and considered. Furthermore, there may be some comparable lead offsets due to electricity replacing or substituting for other sources of fuel energy. I believe most utilities are taking a hard look at their forecasts in view of the realized conservation and cutback in use by their customers and will reflect these effects in good time.

Industry Knowledge

Wayland states that probably no industry of comparable size has as little knowledge about its customers as have the electric utilities. The government, not the industry, has taken the lead in load research and rate structure experimentation. Both of these comments are half-truths that need some clarification.
Until the early 1970s, an electric utility could forecast its future load by drawing a straight line on semilog paper through its historical peaks and project it forward into its planning horizon. If the correlation of the forecast with the actual materialized peak was not .88 or better, and the peak was missed by more than about one percent, then something was wrong. Thus, there was no driving need to study individual customer characteristics for system peak projections. When something had worked very well for a long time, there was no need to employ alternatives that failed to improve upon that system. This did not mean that utilities were not studying alternatives and using them as early leading indicators and checkpoints.

Furthermore, many utilities have been doing extensive load research on the characteristics of their customers, some for as long as four decades. This information, in addition to being used to synthesize the system peak, was used for cost-of-service studies and rate design and for selective selling to attempt to maximize profits while minimizing rate requirements through judicious load building and rate design. I cannot resist the temptation to call attention to the fact that during Wayland’s tenure at the Ohio commission, we entered such data in supporting certain rate structures, and our efforts were summarily ignored in the rate decisions.

What Wayland says is true — major recent efforts in initiating load and rate experiments throughout the nation have come from the government, not industry. But do not confuse frenzied activity with progress and discovery. A number of spinning wheels are being reinvented. I do think that this government-induced wave of research is a stigma on our industry, which should have realized that enlightened self-interest called for aggressive load research and rate studies. In defense of some of the utilities’ inaction, the jury is still out as to whether time-of-day rates and other forms will produce enough savings to warrant their costs. Because no action was taken did not necessarily mean that the subject had not been evaluated. It could mean that such evaluation showed that no action was justified at this time.

Furthermore, many utilities are able to achieve the economies inherent in fuller facility utilization through pumped storage hydro and interruptible rates for very large customers both on and off peak. Thus, many economies available due to optimum capacity use already are being captured without mass disruption for thousands of residential customers and without the attendant cost of millions of hard-to-find investment dollars for additional complex metering and billing.

**A Stronger State Role**

Wayland believes that to help eliminate the confusion resulting from many competing forecasts, the state’s role should be strengthened in preparing estimates of future demand of electricity. I am both comfortable and uncomfortable with this suggestion. There is certainly value in a diversity of approaches to future load forecasts. Provided they are both competent done and professionally objective, the benefits are obvious. I do have some severe misgivings about which one is controlling if there are irreconcilable differences. Management is responsible on a continuing basis to investors and customers, and its forecasts should be controlling. They are the ones who have the best insight into their customers’ plans and their service area, and the general measures of economic indicators that the state agency would have are also available to them. If the utility is consistently wrong in its forecasts, the state regulatory agency has it in its power to take recognition of such performance in setting rates for that utility. The last thing we want is central planning, which, if it is wrong at times (and it surely will be), can create problems for the whole state or region. If individual companies make their own forecasts, there is not this feast or famine consequence.

Furthermore, utility managers are generally career people who will be around to answer for the consequences of their forecasts. This is not true of government regulatory agency personnel. Generally, staff members, albeit intelligent and dedicated, have little work experience, and their average career expectancy in the regulatory agency is less than one-third to one-half that of the forecast planning horizon. Many will be long gone before the forecast they made begins to deviate measurably from actual developments. Any feedback is then delivered to a new vintage of inexperienced government employees. I do not mean to imply that there are not experienced career employees in government service, but there is a high turnover rate.

Another of my major concerns with government planning of forecasts is what I call the “Roman road” approach. A forecast is made, and that becomes the unswerving plan. In essence, the forecast becomes the actual load objective. All regulatory and corporate efforts are then directed at restraining free market demands to conform to this predetermined forecast. Although this system may be used in some countries, I believe it is not compatible with the free enterprise system that made this country economically great through abundant, reliable, and cheap energy.

**Historical Data Are Limited**

Wayland acknowledges that historical data can be useful in establishing relationships among variables, but he declares that reliance solely on historical data is clearly limited. I cannot agree more with
Wayland's discussion of the possibility of subverting the value of models through strict use of historical data to determine cost of capital. Furthermore, I fully agree with his observation that "many of the policies considered by regulatory agencies can affect the perceived riskiness of firms." Overstringent regulation begets economic strangulation and induces investor overreaction. Some well-intentioned regulatory actions may be counterproductive. I diverge from agreement when Wayland laments that the Discounted Cash Flow model is still only the second choice among regulators. My immediate response is: Thank goodness it is only second choice.

Through my more than one-quarter century in the rate regulation field, I have seen many methods of determining cost of capital come into vogue and then fade into semibliviobion. I believe there has not been, and probably will not be, a single financial model that will hold up over varying economic conditions as the sole determinant for cost of equity. Just as the ancients searched for the philosophers' stone, so modern economists and financial analysts are looking for a single magical formula that will suffice for all conditions. I do not believe this will be found, because the capital marketplace is a constantly changing mix of investors, limiting conditions, and investor expectations. I do believe that these models should be constantly used, reviewed, and improved so that they may be as responsive as possible to changing future conditions. This is best done by using several models that approach the problem from different directions.

Further support for my observation comes from Wayland's comment that "the utility industry's feelings about DCF were illustrated by its reaction when ... FERC recently proposed to determine 'zones of reasonableness' on equity returns on the basis of a DCF analysis." I would like to point out that many of those whose "feelings ... were illustrated" were public power and consumer advocates who felt the formula was too liberal. In fact, several segments of the FPC itself wrote position papers opposing the idea.

Irrespective of the data input to these models, there is often some all-controlling assumption that in large measure dictates the end result of these mathematical exercises. In the DCF model, it is the market-to-book ratio that affects the asymptote of return on equity. Choose your assumption, and get your answer. The most compelling proof with which I am most familiar is found in Ohio, where the company's cost of capital, at least in the past, has considered a mark of about 1.0 as parity and has cranked that number into the formula. In the past four years the bond ratings of every electric utility in Ohio have been reduced. One-half of the Ohio electric utilities now have Baa/BBB credit ratings, the lowest level of investor-grade securities.

Instead of dropping down to the alleged parity of a market-to-book ratio of 1.0, these utilities slipped from selling well above book to less than 90 percent of book value. Apparently, the investing public is not using the same formula as the regulatory body.

Fear of Disclosure

Wayland states: "Utilities have sometimes resisted the adoption of forecasts generated by corporate models in regulatory proceedings. One of the most common reasons is the fear of publicly disclosing their own forecasts lest they be liable for stockholders' suits. Many firms appear reluctant to expose their management planning to criticism, especially given the adversarial nature of regulatory hearings." Both the reluctance and fear of the utilities are based to some extent on the factors that Wayland mentions; however, there are many other underlying reasons for not airing our laundry in public.

The disclosure problems with the Securities and Exchange Commission can be significant. Particularly during periods of registration, there is need for full and equal disclosure to all interested parties, and once a forecast is made public, subsequent forecasts, which may vary materially, must be recycled to the investing community. Even if amnesty were granted on this disclosure requirement, the forecasts may take on more meaning and be considered more predictive than they really are. Changes among successive forecasts in such items as load, rates, and capital structure may be misleading because these forecasts are made on a "what if" basis under certain assumed conditions; they are current best estimates of future expectations. They are more in the nature of a series of extrapolative studies that define the boundaries of the company's possible progress rather than maps for unchangeable courses of action. The real value of forecasting models to management lies in the ease with which they can be changed to meet new conditions or new courses of action, challenges, and opportunities. They are used by management as analyses of the bottom line's sensitivity to various paths the management may decide to take. Exalting them to some official status of combined acceptance by the company and the regulatory agencies would crown them with an importance they do not possess.

Conclusion

I would like to conclude by commenting on some oral remarks Wayland made in summarizing his paper. He indicated that perhaps regulatory agencies should have some access to and participation in what goes on in the board room of utilities since the regulatory agency is affected by the decisions made therein. They must grant the rate
increase to support the capital program necessary to build the capacity and facilities needed to meet the forecast load. This seemingly places them in the difficult position in the public eye of bailing out the utilities financially by having no option but to grant rate relief as the last step in a chain of events over which they had no control.

Although there is seemingly some logic to participative involvement in this series of events, the regulatory agency does have a trump card in recognizing the competence and performance of the management in the rate of return granted. In fact, regulators can offer input at power siting and financing hearings, as well as at other generic hearings they may call. Through constant communication between the parties, a closer rapport can develop and provide early signals of differences of opinion. However, the ultimate responsibility for managing a utility rests with its management. I do not believe the regulatory agencies should be privy to all that goes on in the board room, nor do or should they have the ultimate responsibility.

Part Seven

New Issues in

Telecommunications Regulation
The Impact of Intercity Rate Restructure on Local Telephone Rates

Alan C. Hasselwander

Anyone even remotely connected with telecommunications is aware of the expressions of emotion ranging from grief and anguish to fear and anger occasioned by discussions of competition in the intercity transmission market. Whatever sympathy is evoked from makers of public policy seems to be largely directed toward small, rural telephone companies, perhaps because of the advocacy of OPASTCO (Organization for the Protection and Advancement of Small Telephone Companies) and NTCA (National Telephone Cooperative Association), as well as the REA (Rural Electrification Administration, which has a large stake in the outcome of the current public policy debate). The implicit concern of small telephone companies is not, it seems to me, over the loss of monopoly rents (if indeed they ever existed). Rather, their concern seems to be whether they can, in the long run, cover their total costs under competition. Not unpredictably, the public debate over competition in intercity transmission is couched in quasiethical or public policy terms — in this case universal service or, more specifically, the “right” of rural customers to service comparable to that in urban areas at prices comparable to or less than urban prices. Subsidies flowing from toll services, it is contended, should support local ex-
change service, while toll revenues on dense routes which exceed costs should support toll usage on less dense routes.

It is not my intention to express any opinion about the public policy debate. (In fact, I cannot think of any useful way in which to define policy.) But the quality of the debate could be improved by the existence of reasonable estimates of the real and potential effects of competition on common carriers and consumers.

I wish I could make a large contribution and report comprehensive estimates of these effects. My effort will be much more modest. I will discuss some aspects of the analysis and share some data that may be useful.

In thinking about the costs of toll transmission in the common carrier network, it is obvious that some costs are a direct function of intercity calling, while others are incurred in common with local exchange usage. Examples of direct costs are associated with investments in toll switching machines, intertoll trunks, toll connecting trunks, toll billing, and operator services. Examples of common costs are associated with investments in Class 5 central office switching and subscriber loop plant. It is this latter investment that is most relevant to toll revenue levels of small independent companies.

Toll revenues accruing to telephone companies reflect costs and return associated not only with direct investment, but also with an arbitrary allocation of common investment through toll separations procedures. (Any allocation of common costs without reference to demand, it seems to me, is arbitrary.) Both nontraffice sensitive and so-called allocative traffic sensitive common plant are reflected in the ratio of toll minutes of use to total minutes of use, weighted by a factor called SPF or DEM (Subscriber Plant Factor and Dial Equipment Minutes).

Prices faced by toll users do not generally reflect relative costs. Interstate toll prices, for example, are a function of distance regardless of the route, although off-peak costs and operator handling are reflected in price.

Loss of potential toll revenues by companies currently result from four distinguishable substitutes. First, point-to-point private line service without access to the local exchange network is a price discriminating message toll substitute, for example, from PBX to PBX. The private line may be provided by AT&T, a specialized common carrier, or the using customer. Before private line offerings, intercity traffic from one PBX to another was message toll, each message making a contribution to common exchange costs. From the customer's point of view, the substitute is perfect at a lower cost than message toll, and since price is not directly sensitive to usage but, rather, to capacity, usage and therefore utility are probably increased.

Second, point-to-point private line traffic can be indirectly connected to the exchange network on a dialized basis or through operator intervention. The additional price faced by the customer is the price of a local call, which in some cases is zero.

Third, with foreign exchange type service (FX), a customer subscribes to a line in a central office located in a distant exchange. FX service provides a close toll substitute for customers who originate and/or receive substantial traffic to and from a specific distant exchange. FX service can be provided by common carriers, or jointly by common carriers and a specialized common carrier, or a private user.

Common carrier FX pricing is generally based on circuit mileage and the local exchange rate.

Finally, executing-type service originates and terminates on the local exchange network. As can private services with common carrier network access, it can originate a toll call at a distant point.

Each toll substitute reduces toll usage of the network, and to the extent that the toll substitute traffic originates or terminates in the exchange network, it increases measured local usage. Because the substitutes are priced lower than toll, local usage is presumably increased more than toll usage is decreased. This is significant when one considers that loss of revenues to the telephone company under toll separation procedures is a function of both the decrease in toll usage and the increase in local usage.

In the future, intercity competition presumably will elicit continuing competitive response from AT&T in the form of prices which are cost related rather than average, and which generally are lower than present prices. To a small telephone company, this implies increasingly less contribution from toll to the common costs of the exchange plant. It also implies a higher price for toll traffic or generally less dense routes to and from small company exchanges.

I think the smaller company's view of intercity competition is that toll substitute services have already shifted a burden to local exchange revenues, that the rate at which toll substitute services are growing is increasing, that separations procedures for cost allocation will change and further shift burden, and that higher prices on lower density routes will further reduce toll usage and thus toll revenues. And these factors are viewed as independent of one another and totally negative.

But a local exchange pricing strategy seems to follow from the discussion. The suggestion of a charge sensitive to originating and terminating usage as a replacement for toll separations and applicable to all intercity services frequently has been made. But what should be the level of the charge? A quick answer may be to search for a price that maximizes the contribution of intercity service. But if the price is higher than that of originating local service, there will be an incentive
for some customers to shift to toll substitutes which provide indirect access to the network and cannot be identified as intercity traffic. That would suggest that the price of originating traffic, toll or local, and terminating toll calls be equivalent.

I said earlier that the typical small firm views the potential impact of intercity competition as unidirectional — and negative. In fact, it is impossible to know a priori whether the combined effects of intercity competition on a specific company will, in the long run, be positive or negative, if the telephone company has an access charge which does not in itself provide an incentive for toll substitute services.

Under competition, the price of an AT&T toll message to any exchange will be a function of competitors’ prices to that exchange, subject to the necessity to cover the total costs of the entire intercity enterprise. If competition is intense on a route, price could go low enough to just cover the direct costs of that route. Even if a cost differential exists between routes of various densities, it is unclear whether the resulting price would be higher or lower than present, without first knowing the attractiveness of the route to competition, the direct costs of serving the route, and the price elasticity of demand of calling to and from the exchange.

Unfortunately, there seems to be little public information concerning the relative costs of providing intercity service as a function of traffic density. We also do not know a great deal about price elasticity except, perhaps, in aggregated forms. Concerning the relative costs of providing intercity calling, I would have serious difficulties in analyzing those relative costs. First, toll traffic to various destinations can generally be carried in common over a large part of the route. Second, busy hours or peak periods are not necessarily coincident over the entire route, and thus the marginal cost of traffic on some piece of the route may approach zero. Third, in thinking about future costs it is not clear that the level of service, for example, the probability of blockage, needs to be identical to and from every exchange.

I would now like to report on a somewhat crude estimate I have made of the potential impact of intercity competition on local exchange revenue requirements. I have worked with data from so-called average schedule companies in New York and Wisconsin that serve fewer than 5,000 total stations and perform no toll traffic operator functions. For those uninitiated into the mysteries of toll settlement procedures, an average schedule company is one whose toll revenues are based on the average revenue requirements of a number of studied companies.

For purposes of the analysis, it is necessary to recognize the nature of SPF. The so-called subscriber plant factor can most easily be described as a coefficient greater than one which is multiplied by toll SLU (toll subscriber line usage) to allocate nontraffic sensitive plant investment. The average effect of SPF on toll revenues of average schedule companies without toll traffic operator functions is nearly 34 percent — 44 percent in the case of interstate traffic and 27 percent in the case of intrastate toll traffic. That is, if SPF is removed from the toll settlement calculation, toll revenues would decrease by the amounts mentioned.

For the companies studied, I reduced toll revenues (1977 revenues in the case of New York companies, 1976 revenues in the case of Wisconsin companies) by 30 percent and 40 percent to reflect the potential impact of eliminating SPF. I then increased local revenues by an amount equal to the toll revenue reduction. This method provides an approximation of revenue requirements under the condition that local and toll usage of the exchange plant is priced identically based on usage. It assumes zero price elasticity of demand, which is the major weakness of the analysis.

Table 1 shows the distribution of single party residential flat rates of the 30 New York State independent company exchanges fitting the study definition. The first column describes the rates as of 1 January 1978. The second column reflects what the rate would be if a 30 percent loss of toll revenues was recovered by local revenues through an equal percentage increase of all prices. The third column is similar to the second column, except that the toll loss is 40 percent. The fourth column shows the distribution of rates for the same exchanges on 1 January 1968. The final column shows the 1968 data adjusted to relative price levels on 1 January 1978, using the Consumer Price Index (CPI) of all items.

Table 2 analyzes the distributions described in Table 1 by using ordinary descriptive statistics. The average increase at a 30 percent loss of toll is 31 percent. The low end of the range increases significantly greater than the high end. The variation in rates increases under the condition of decreased toll revenues. Of particular significance to me are the CPI adjusted data in the last column and their similarity to the 30 percent column. The comparison suggests that, in real dollar terms, the price of residential telephone service in the companies studied is significantly less than in 1968. It further suggests that at a 30 percent toll less loss, prices would be approximately equal to the 1968 level.

Table 3 refers to the same telephone companies, but measures the local revenue effect per main station. The movement of the mean is similar to the effect on one party flat residential rates in Table 2. A 30
### Table 1. Effect of Toll Revenue Loss on the Distribution of Basic Service Rates of Certain New York State Telephone Companies

<table>
<thead>
<tr>
<th>Rates</th>
<th>1/1/78 30 percent loss (percentage of exchanges)</th>
<th>1/1/78 40 percent loss (percentage of exchanges)</th>
<th>1/1/68 30 percent loss (percentage of exchanges)</th>
<th>1/1/68 40 percent loss (percentage of exchanges)</th>
<th>CPI adjusted 1/1/68 (percentage of exchanges)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$5.50 - 4.49</td>
<td>3.3</td>
<td>0</td>
<td>0</td>
<td>32.1</td>
<td>0</td>
</tr>
<tr>
<td>4.50 - 3.49</td>
<td>13.5</td>
<td>0</td>
<td>0</td>
<td>39.5</td>
<td>0</td>
</tr>
<tr>
<td>5.50 - 4.49</td>
<td>16.7</td>
<td>6.7</td>
<td>0</td>
<td>14.5</td>
<td>0</td>
</tr>
<tr>
<td>6.50 - 5.49</td>
<td>26.7</td>
<td>13.3</td>
<td>20.0</td>
<td>14.5</td>
<td>10.7</td>
</tr>
<tr>
<td>7.50 - 6.49</td>
<td>20.0</td>
<td>16.7</td>
<td>3.3</td>
<td>0</td>
<td>46.4</td>
</tr>
<tr>
<td>8.50 - 7.49</td>
<td>16.7</td>
<td>16.7</td>
<td>16.7</td>
<td>0</td>
<td>10.7</td>
</tr>
<tr>
<td>9.50 - 8.49</td>
<td>5.3</td>
<td>13.3</td>
<td>26.7</td>
<td>0</td>
<td>17.9</td>
</tr>
<tr>
<td>10.50 - 9.49</td>
<td>0</td>
<td>30.0</td>
<td>3.3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11.50 - 10.49</td>
<td>0</td>
<td>3.3</td>
<td>26.7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12.50 - 13.49</td>
<td>0</td>
<td>0</td>
<td>5.3</td>
<td>0</td>
<td>14.5</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

**Note:** Shown are the distribution of single party residential flat rates of 20 New York State exchanges as of 1 January 1978, showing the "effect" of 30 percent and 40 percent loss of toll revenues, and the distribution of single party flat rates of the same exchanges on 1 January 1968, "adjusted" by the Consumer Price Index to 1 January 1978.

### Table 2. Statistical Summary of the Effect of Toll Revenue Loss on Basic Service Rates of Certain New York State Telephone Companies

<table>
<thead>
<tr>
<th></th>
<th>1/1/78 30 percent loss</th>
<th>Percentage change from 1/1/78 (percentage)</th>
<th>1/1/78 40 percent loss</th>
<th>Percentage change from 1/1/78 (percentage)</th>
<th>1/1/68 30 percent loss</th>
<th>1/1/68 40 percent loss</th>
<th>CPI adjusted 1/1/68 (percentage)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>$ 6.97</td>
<td>$ 9.16</td>
<td>31.45%</td>
<td>$ 9.89</td>
<td>41.96%</td>
<td>$ 4.98</td>
<td>$ 9.08</td>
</tr>
<tr>
<td>Median</td>
<td>7.25</td>
<td>9.23</td>
<td>27.3%</td>
<td>10.14</td>
<td>39.9%</td>
<td>4.50</td>
<td>8.20</td>
</tr>
<tr>
<td>Low</td>
<td>4.15</td>
<td>6.22</td>
<td>49.9%</td>
<td>6.91</td>
<td>66.5%</td>
<td>3.75</td>
<td>6.84</td>
</tr>
<tr>
<td>High</td>
<td>9.70</td>
<td>11.98</td>
<td>23.5%</td>
<td>12.74</td>
<td>31.3%</td>
<td>7.15</td>
<td>15.03</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>1.57</td>
<td>1.70</td>
<td>1.83%</td>
<td></td>
<td>1.02%</td>
<td>1.86%</td>
<td></td>
</tr>
<tr>
<td>Coefficient of variation</td>
<td>22.5%</td>
<td>18.5%</td>
<td>18.5%</td>
<td></td>
<td>20.5%</td>
<td>20.5%</td>
<td></td>
</tr>
</tbody>
</table>

**Note:** Given are various statistics describing single party residential flat rates of 20 New York State exchanges as of 1 January 1978, showing the "effect" of 30 percent and 40 percent loss of toll revenues, as well as single party residential flat rates of the same exchanges as of 1 January 1968, "adjusted" by the Consumer Price Index to 1 January 1978.
percent loss of toll revenues results in roughly a 30 percent increase in local revenue requirements, and a 40 percent loss results in roughly a 40 percent increase, reflecting an approximate fifty-fifty split in toll and local revenues among the companies studied.

Table 4 compares monthly one party flat rates in 1977 in each of the New York State exchanges studied with the single party rate "adjusted" for the effect of a 30 percent loss in toll revenues. Obviously, the effect differs from one company to another, depending on the firm's reliance on toll revenues as well as other factors.

Table 5 contains data from 63 Wisconsin independent telephone companies presented in a format similar to Table 5. The average effect of a 30 or 40 percent loss in toll revenues is to raise local revenues by roughly the same percentage (as in the case of the New York companies). As with the New York data, the distribution of local revenue requirements is widened after toll revenue loss.

Table 6 shows the proportion of toll revenues to local revenues for the companies studied. The proportion in both cases is nearly one, lower than a 1.15 average for the 757 independent companies reporting data to the U.S. Independent Telephone Association and representing 96.5 percent of independent company revenues.

What conclusions do I want to draw? Obviously, the data are not conclusive, but I think they tentatively suggest some things. In the case of the New York State companies, the relative price level of local telephone service was significantly lower in 1978 than in 1966, at least as reflected in residential one party rates. A 30 to 40 percent decrease in toll revenues has some support as an outer bound of toll revenue decreases.

### Table 4. Monthly One Party Flat Rates in Effect 1 January 1978 in 30 Independent Telephone Company Exchanges, Compared with One Party Flat Rates "Adjusted" for a 30 Percent Loss of Toll Revenues

<table>
<thead>
<tr>
<th>Monthly rate - 1FR</th>
<th>Number of exchanges</th>
<th>&quot;adjusted&quot; for 30 percent loss</th>
<th>Change</th>
<th>Percentage change</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.15</td>
<td>1</td>
<td>6.22</td>
<td>2.07</td>
<td>49.9</td>
</tr>
<tr>
<td>4.50</td>
<td>1</td>
<td>6.43</td>
<td>1.93</td>
<td>42.9</td>
</tr>
<tr>
<td>4.65</td>
<td>3</td>
<td>6.85</td>
<td>1.98</td>
<td>42.6</td>
</tr>
<tr>
<td>5.70</td>
<td>3</td>
<td>7.87</td>
<td>2.17</td>
<td>38.1</td>
</tr>
<tr>
<td>6.00</td>
<td>1</td>
<td>8.64</td>
<td>2.64</td>
<td>44.0</td>
</tr>
<tr>
<td>6.25</td>
<td>1</td>
<td>8.58</td>
<td>2.39</td>
<td>37.3</td>
</tr>
<tr>
<td>6.50</td>
<td>1</td>
<td>7.95</td>
<td>1.45</td>
<td>22.3</td>
</tr>
<tr>
<td>6.50</td>
<td>1</td>
<td>8.87</td>
<td>2.37</td>
<td>36.5</td>
</tr>
<tr>
<td>6.51</td>
<td>1</td>
<td>7.15</td>
<td>0.44</td>
<td>9.8</td>
</tr>
<tr>
<td>6.85</td>
<td>1</td>
<td>8.38</td>
<td>1.53</td>
<td>22.3</td>
</tr>
<tr>
<td>7.25</td>
<td>2</td>
<td>10.81</td>
<td>3.56</td>
<td>49.1</td>
</tr>
<tr>
<td>7.26</td>
<td>1</td>
<td>9.22</td>
<td>1.96</td>
<td>26.9</td>
</tr>
<tr>
<td>7.50</td>
<td>1</td>
<td>9.24</td>
<td>1.94</td>
<td>26.6</td>
</tr>
<tr>
<td>7.50</td>
<td>1</td>
<td>10.96</td>
<td>3.46</td>
<td>46.1</td>
</tr>
<tr>
<td>7.55</td>
<td>4</td>
<td>10.87</td>
<td>3.32</td>
<td>44.0</td>
</tr>
<tr>
<td>8.40</td>
<td>1</td>
<td>10.67</td>
<td>2.27</td>
<td>27.0</td>
</tr>
<tr>
<td>8.75</td>
<td>1</td>
<td>11.41</td>
<td>2.66</td>
<td>30.4</td>
</tr>
<tr>
<td>9.30</td>
<td>4</td>
<td>10.15</td>
<td>0.85</td>
<td>9.1</td>
</tr>
<tr>
<td>9.70</td>
<td>1</td>
<td>11.98</td>
<td>2.28</td>
<td>25.5</td>
</tr>
</tbody>
</table>

### Table 5. Local Revenues per Main Station per Month of 63 Wisconsin Independent Telephone Companies, 1976, "Adjusted" for a 30 Percent and 40 Percent Loss of Toll Revenues

<table>
<thead>
<tr>
<th></th>
<th>30 percent loss</th>
<th>40 percent loss</th>
<th>Percentage change from 1976</th>
<th>Percentage change from 1976</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>$ 7.93</td>
<td>$10.58</td>
<td>30.9</td>
<td>$11.20</td>
</tr>
<tr>
<td>Median</td>
<td>7.55</td>
<td>10.29</td>
<td>36.3</td>
<td>11.08</td>
</tr>
<tr>
<td>Low</td>
<td>3.94</td>
<td>6.65</td>
<td>68.5</td>
<td>7.13</td>
</tr>
<tr>
<td>High</td>
<td>15.30</td>
<td>17.56</td>
<td>14.8</td>
<td>18.31</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>2.3</td>
<td>2.18</td>
<td>2.18</td>
<td>2.18</td>
</tr>
<tr>
<td>Coefficient of variation</td>
<td>29%</td>
<td>21%</td>
<td>19.5%</td>
<td>19.5%</td>
</tr>
</tbody>
</table>
Table 6. Local and Toll Revenues per Main Station per Month

<table>
<thead>
<tr>
<th></th>
<th>Local</th>
<th>Toll</th>
<th>Toll as a percentage of local</th>
</tr>
</thead>
<tbody>
<tr>
<td>New York companies</td>
<td>$9.56</td>
<td>$9.14</td>
<td>95.6%</td>
</tr>
<tr>
<td>Wisconsin companies</td>
<td>8.15</td>
<td>7.99</td>
<td>98.0%</td>
</tr>
</tbody>
</table>

NOTE: Data are for 18 New York State independent telephone companies for 1977 and 65 Wisconsin telephone companies for 1976.

loss in the foreseeable future — assuming allocation of common costs based on usage and zero elasticity. Small telephone companies depend heavily on toll revenues. The loss of these revenues within the range of 30 to 40 percent can have a significant effect on local service prices, but in my opinion not to a degree that is politically inolerable, particularly in view of the improbability of the entire toll revenue loss occurring at one time. Viewed in constant dollar terms, subscribers may pay no more for local service than in the past.

In conclusion, I reiterate that much more must be known about the potential effect of competition in the intercity market before discussion can proceed at more than an emotional level.

The Continuing Role of Federal Regulation in the Transition to Competition in Communications

Walter G. Bolter

Over the last ten years the Federal Communications Commission (FCC) has made increasingly greater use of competitive market forces to supplement traditional regulation. These efforts have been focused on increasing the efficient use of communications resources, and on making services available which more closely match the needs of both business and residential users. Rather than exposing all communications markets to a new and disruptive environment simultaneously, the commission has attempted to foster competition gradually. That is, its decisions have centered on introducing new entry in discrete stages and in isolated market areas. In the course of implementing these policies, the FCC has unearthed a host of parties who have contrary interests. Critics have ranged from those wishing to slow or reverse the process of introducing competition, to those who feel that regulatory interference in the marketplace must be drastically curtailed or eliminated outright.

This article briefly reviews the evolution and bases for marketplace

NOTE: The ideas and conclusions expressed herein represent the personal views of the author and should not be interpreted as representing the views of the Federal Communications Commission or other members of the staff.
intervention that have been used to rationalize regulation in the past. It considers alternatives to the FCC's approach of "gradualism," particularly those that have been advanced by the telecommunications industry and in Congress. The need for continued federal intervention over the short run is then explored from the standpoint of the current stage of competitive entry, the national importance of communications, and the effect that FCC involvement might have on market processes. Finally, a review is made of industry and national economic factors that may warrant intervention even after recent market entrants have reached maturity.

Traditional Regulation

In the past, regulation has been advocated because of the natural monopoly conditions under which services have been provided, or due to the services' "essential" nature. When regulation has been imposed, commissions have been responsible for constraining utilities from earning excess profits by virtue of their special franchises, and for promoting maximum productivity and efficiency. Economic supply under natural monopoly conditions, however, has seldom been shown to exist in practical fact. Therefore, needs for governmental intervention on the basis of social requirements have had to be interpreted broadly. In Nebbia v. New York, for example, the Supreme Court found that "the State's power extends to every regulation of any business reasonably required and appropriate for the public protection."9

Specific rationales for regulation falling within the general guidelines of Nebbia have included the need for protection of the public from extortion, unjust rates, or discriminatory business practices. Furthermore, franchising of suppliers has been required, such as in the case of the motor carriers, when certain businesses affect the prices or services of existing utilities as to require regulation themselves. In these instances, there has been the practical fact or belief that competition will not work to benefit the public weal. Regulation has often been imposed as a substitute for market forces. That is, regulation has been expected to afford consumers protection from exploitation, and also to ensure that they benefit from technological change and industrial progress just as they would in a competitive system. Because of their special status, regulatees have assumed certain obligations and rights not commonly associated with firms operating in a free marketplace. Typical utility obligations have included requirements to (1) make service available to all who request it up to the utility's capacity, (2) provide such service on equal and reasonable terms, and

(3) render service that is safe, adequate, and available on a continuous basis. In return, utilities have usually been assured of (1) the general right of legal protection of private property, (2) the right to collect a reasonable price for services rendered to cover their costs, (3) the right to impose reasonable service rules and regulations on customers, and (4) the right of eminent domain.

Utility obligations and rights have two major implications for intercustomer cross-subsidization. These take on particular policy significance in the context of introducing competition into an erstwhile monopoly based industry. First, the obligation to provide service to all who request it on equal terms implies that cross-subsidization will exist within a particular service category. No matter how similar the subscribers to a service are, all cannot conceivably have identical cost characteristics. Thus even though a service's total revenues may equal its total costs, certain customers will be paying relatively more for service than others. If new firms are allowed to enter the market without similar obligations, they will choose to serve only the low cost or more profitable customers. The existing utility may then have no option other than to internalize these "obligation to serve" costs, and to earn less than a normal return on the service in question.4

Second, the utility's rights to legal protection of property and to price its services to ensure full cost recovery imply that cross-subsidization may also exist between services. Customarily, the right applied to all utility expenditures past and present and ensure that the firm will earn a normal return on its collective assets. If certain services are subjected to competition, it is likely that the return on these services will fall. This may be the resultant of a utility's obligation to serve all customers. However, it could also be due to imprudent past investment decisions, unforeseen changes in business conditions, or other events that would cause divergences between historical (book) and current costs. Whatever their source, the utility would have the right to adjust the prices of other services in order to recover all of its costs and earn a normal return on its property as a whole. Most likely, the utility would find it expedient to overcharge those customers having the fewest alternatives, namely, monopoly service subscribers.

Regretably, commissions have embarked on procompetitive policies in the past without fully addressing these policies' implications regarding cross-subsidization and utility obligations and rights. In particular, it has not been recognized that when all services are subjected to new entry, hopeless inconsistencies are likely to arise between traditional regulation and procompetitive policy objectives. In theory, there would be no available targets for cross-subsidization, and the utility may find that it is impossible to recover residual past costs plus
current costs of service. Thus the firm could be precluded from earning the same level of profit as that enjoyed by new entrants and could be effectively denied legal protection of its property.6

The specific procedures and economic bases of traditional regulation also place impediments in the path of procompetitive policies. Commissions have uniformly performed the translation of utility rights and responsibilities into practical procedures by relying on a revenue requirements standard. This standard involves calculations of depreciation and other operating expenses, capital costs, and rate base. It provides for coverage of the operating costs of rendering service, plus a return on rate base or property devoted to the regulated enterprise.6

In communications, regulatory procedures have been developed in accordance with this general model. However, before the introduction of competition, the revenue requirements standard was applied on a company-wide basis rather than to particular services. In the case of the Bell System, for example, only the company’s overall expenses, rate base, and profit were subjected to regulatory scrutiny. The FCC was not concerned with the rates Bell charged individual services or the company’s rate structure. Moreover, the commission allowed Bell to justify its incurred, allocation, and recovery of costs separately. Thus the reasoning that the company used for the justification of construction plans could be different from that used for distribution or recovery of these costs from individual services. Under this aggregate approach, the commission did not have the ability to detect the presence of cross-subsidization, between or within services, and could not prevent its occurrence should the agency embark on a policy of introducing competition into the industry.

Entry of Competition

To have viable competitive entry at least three preconditions must be satisfied: (1) the existence of financial or technological advantage for new suppliers, (2) demand for the services of these suppliers and their access to available markets, and (3) regulatory sanction. Prior to the 1970s, competitive entry in communications was most likely to have succeeded in the private-line interservice transmission and terminal equipment markets. In each market, however, entry was not substantive. In part, financial problems and the vagaries of demand were to blame. However, the lack of a strong regulatory mandate for competition was also evident. In the pre-1970 period, the FCC did not espouse a strongly procompetitive posture, nor did it establish pricing and costing standards which would reflect all carriers’ relative service costs and permit full competition. Instead, commission decisions favoring competition appeared to be reactive to the business and technological pressures of the 1950s and 1960s, or simply complied with external legal requirements.7

During the late 1960s and particularly the “fabulous 1970s,” this has all changed. Technological advancements have favored new entry, and major industrial corporations such as IBM and Southern Pacific have provided needed financial backing for new suppliers. Emerging needs for innovative business data, voice, facsimile, and other services have provided customer demand and ready markets for new suppliers. Finally, the FCC’s terminal equipment registration program and such decisions as Specialized Common Carrier and Domsol in the private-line market have signaled the commission’s intent not only to sanction competition, but also to provide competitive advantages to firms which enter the industry.8

The courts have generally upheld the FCC’s procompetitive decisions and, in some instances, have broadened their impact. For example, in the ExecuNet case the U.S. Court of Appeals effectively extended the FCC’s competitive experiment into the message toll area of interservice transmission.9 Other forces are at work which will extend competition to the very frontiers of communications itself. For instance, proposals like the Xerox Corporation’s XTEN service and the Postal Service’s ECOM offering will permit direct substitution between the products of the postal and communications industries. Substitute in-kind and interindustry substitution is evident between data processing and communications. On balance, approximately ten years after the FCC embarked on its procompetitive program, these policies have become the watchword throughout communications. Moreover, they have had significant “spillover effects” on the frontier between communications and the products and services of other industries.

The entry of competition has raised a host of policy issues. Many of these reflect the schism between traditional regulation and competitive entry discussed heretofore. For example, in order for new entrants and existing carriers to compete “fully and fairly,” common carrier obligations and rights must be reconciled with the new communications environment. The most immediate issue that must be faced is the distribution and recovery of established carriers’ past expenditures for plant and equipment. These expenditures, which were made under the “old rules” of common carriage, are substantial. Furthermore, if not properly dealt with they will distort price competition and the market signals provided to communications investors and users. A similar issue is that of cross-subsidization. The FCC addressed this issue in Docket 18128 and found that full and fair competition could be assured if costs were allocated according to “cost causalional” principles.10 However, if this system is to work, the traditionally isolated
Role of Federal Regulation

processes of incurring, distributing, and recovering costs will have to be tied together, and the FCC’s revenue requirements mode of regulation will have to be dramatically improved. It is not yet clear whether the commission will have the resources and political support needed to implement its findings.

Several policy issues stemming from the introduction of competition relate to its spillover effects in the interindustry area. Of particular concern are the rules under which common carriers can participate in the provision of mixed or substitute services. For example, should the terms and conditions under which communications firms are permitted to compete be any worse than those faced by outside concerns which offer substitutes for communications services? Another issue involves the processes of valuing and distributing public resources such as radio spectrum. These processes should not lessen communications companies’ inherent economic advantages over competing entities, such as the Postal Service, whose major activities lie outside of the industry. Contrarily, they should not cause excessive substitution of communications for the products and services of other industries.

Regulatory Alternatives

Over the last few years, there has been a wide range of responses to the issues raised by competitive entry. At the commission, of course, it has been “business as usual.” For the most part, the FCC has continued along the path it embarked on nearly ten years ago, namely, of furthering competitive entry in selected market areas. Unfortunately, several other factors have recently been placed on the FCC’s plate. In the past, commission decisions were premised on continuation of the established telephone carriers’ de facto monopoly on services offered over the intercity public switched network. As a result of the Execusnet decision, however, no such constraints on competitive entry can now be drawn. Thus the commission cannot continue to limit competitive carriers’ facility authorizations to the private line area. While this does not necessarily imply that the commission must depart from its traditional policy of gradually fostering competitive entry, it will limit the FCC’s role as a catalyst of market change. Indeed, over the near term, major new policy directives must await the outcome of Docke 78-72, the MTS/WATS market structure inquiry. In Execusnet, the court found that the FCC’s decisions which limited competitive entry lacked record support. Docke 78-72 will provide the record evidence found lacking in Execusnet decisions, and will underlie the FCC’s major common carrier policy determinations in the 1980s.

Telephone industry proposals dealing with competitive entry and the future structure of telecommunications have undergone relatively violent metamorphoses over the last few years. The first was the proposal for a “Consumer Communications Reform Act” (CCRA) in 1976. Under CCRA, the basic thrust of the FCC’s procompetitive policies in the terminal equipment and private line areas would have been reversed, and the institutional relationships and structure of the industry would have been shaped to resemble those of the precompetitive era. CCRA seemed to reflect the view that new entrants and the traditional telephone industry were intrinsically incompatible, and were thus locked in a “mortal struggle.”

In 1977, after passage of CCRA became increasingly remote, the established telephone carriers advanced the “industry bill” proposal. Although not as extreme as CCRA, the industry bill would have resulted in the permanent restriction of new entrants to certain discrete service areas. It would have also established pricing, interconnection, and other market rules which were highly advantageous to the established carriers. For example, under the industry bill a specially designed connection charge for access to the nationwide public switched network would have been imposed on all competing carrier services. Sharing or interconnection of these services with MTS and WATS would also have been prohibited.

The established carriers have recently given indications that a “second generation” of industry proposals may be forthcoming. In a speech before the Chicago Commercial Club in November 1978, chairman-elect Charles L. Brown indicated that the Bell System no longer sees “the telephone industry as locked in a mortal struggle” with its competitors.31 In figuratively inviting new competitors to dinner, Brown indicated that he could envision an industry which would encompass the established telephone companies, specialized carriers, and interconnects. Brown did not dwell on the negative aspects of competition, but, instead, emphasized requirements for “customers to be served” and for resolution of policy issues in order to permit the “full application of our industry’s energies to the nation’s needs.” However, certain cost increases would accompany acceptance of the specialized carriers and other competitors. For example, it was indicated that “home telephone service is going to go up faster and farther than it would otherwise,” and that nationwide rate average pricing would eventually have to be abandoned. The reasoning involved was that regulators “extend to others the right to compete with us on routes that profit them most while holding us to our obligation to serve everybody regardless of the cost of doing so.” Thus Bell’s second generation proposal might be that if the company is to compete fully and fairly and also to cover the obligation to serve costs of common carriage, local service rates will have to be adjusted to ensure that Bell’s right of full cost recovery is satisfied.
A recent discussion paper by R. M. Alden, president of United Telecommunications, may represent another telephone industry tentative proposal.13 As would chairman-elect Brown, Alden would be willing to accept competition on certain conditions. These include requirements for an established carrier monopoly on local exchange services. Such services would have to be state regulated, and the prices charged to intercity users would be determined according to value (rather than cost) based standards. Deregulation of all terminal equipment and intercity services would also be mandated, and established common carriers would be permitted to enter any competitive market. Furthermore, these carriers would be assured of the right to set access charges and to plan in concert, to refuse interconnection to competitors except through local (state regulated) services, and to maintain their present corporate structures.

Although the conditions for acceptance of competition delineated in Alden’s paper are much more specific than those contained in the remarks of Brown, certain parallels can be drawn. For example, both seem to recognize an intrinsic need for isolating local exchange services as a residual cost recovery area. In comparison to past proposals, both appear to favor the access change element of the industry bill. Of course, this common feature is predictable, since all proposals after CGRA accept the existence of competition. Thus they must include a technique for charging new entrants with costs that the established carriers feel they should bear. Finally, the views of Brown and Alden seem to reflect the idea that some control over the FCC’s “competitive experiment” must be established. Although the latest thinking is that competitors can be “lived with,” the industry still appears intent on segregating certain service areas from market entry.

Whereas the telephone industry seems to have concluded that the FCC’s pace of introducing competition has been too fast, the sentiment of Congress appears to be just the opposite. In recent years, those who maintain that competition and regulation do not mix or that communications should be deregulated have generated considerable support. Of course, the Communications Act Rewrite (HR 13015) represents the most prominent deregulation proposal that has been sponsored to date.13 Under the rewrite there would be a presumption that competition is in the public interest and that the rates charged in “workably competitive” markets are equitable. Thus there would be no need for a proceeding such as FCC Docket 78-72 to determine whether competition should be excluded from traditional monopoly markets. Regulation would be necessary only “to the extent that marketplace forces are deficient,” and regulatory powers would be drastically diminished. For example, there would be no commission powers to suspend or reject discriminatory or predatory tariffs in workably competitive markets. Moreover, regardless of the degree of competition, carriers could construct new facilities or extend existing ones without regulatory approval.

Other provisions of HR 13015 include requirements that carriers which provide noncompetitive services must divest their equipment manufacturing affiliates. The rewrite would establish a universal service compensation fund, financed by intercity carrier access charges, which would be used to make voice telephone service universally affordable. HR 13015 also appears to contain various market and carrier distinctions. Indeed, firms offering services under both workably competitive and noncompetitive market conditions would be subject to two sets of operating rules. These firms would be obligated to extend service to all customers of noncompetitive services, but not to customers of competitive services. However, it seems that only firms classified as “common carriers” would be subject to depreciation prescriptions.

The Transition to Competition

Throughout the period in which it has attempted to encourage competitive supply, the FCC has lacked the support of industry as well as the guidance of many who now espouse competition and instant deregulation. The question facing the commission today is whether the policy alternatives that have been advanced over the last few years are preferable to continuing to carry out an orderly transition to a competitive marketplace. As noted, the industry has opted for the reversal or imposition of permanent strictures on the FCC’s “competitive experiment.” However, current prospects for adoption of these proposals do not appear promising. For example, the proposed reversal of FCC policies contained in CGRA seems passé when viewed in terms of existing political support for new competitive entry. If the industry bill had been advanced in the late 1960s or early 1970s, it may have been considered only slightly less radical than the commission’s policies of that period. Unfortunately, today many would consider it to be even more conservative than an already too conservative (FCC) policy approach.

Although packaged differently, many of the second generation views of Alden and Brown are faithful to the underlying principles of the industry bill. Both seemingly give up considerable ground in conceding the intercity service area to competition. However, they do so on the industry’s terms. Under the industry bill, the established carriers would be assured of 100 percent recovery of their expenditures and flexibility in setting rates and distributing costs to services and juridic-
tions. They would also enjoy a “favorable” regulatory climate (that is, one in which regulatory surveillance is fractionated but is still available for use as a countergovernmental policy based on the views of Brown and Alden would satisfy these conditions. Complete recovery of expenditures would be assured by maintenance of an established carrier monopoly over local exchange services as a residual cost reservoir. Deleveraging, recognition of value of service, and the ability to collude in setting prices would maintain flexibility in distributing costs and ratemaking. Finally, there would be strictures on corporate re-structuring, and requirements for (limited) state regulation of local exchange services and the intercity-local exchange interface. These provisions together with minimal federal regulation elsewhere would achieve a favorable regulatory climate. The industry’s terms for acceptance of competition are not as yet reminiscent of relationships between suppliers in a purely competitive market. Furthermore, they would not result in a marketplace that is “workably competitive,” if this is defined as an environment in which market rules do not operate to the advantage of any carrier.

Assuming that competitive market conditions exist, deregulation in the context of HR 13015 may offer many institutional and industry structure advantages. Indeed, the rewrite’s presumption that a competitive industry structure furthers society’s goals would avoid need for the considerable record that must be amassed in Docket 78-72. One cannot, however, simply presume (or legislate) away the currents of competition that exist or will exist in the industry. Under today’s conditions the risks are high that if established carriers are not regulated, they will engage in discriminatory or predatory pricing. Moreover, the difficulties that are inherent in the detection and prevention of such behavior would not be mitigated under HR 13015. The rewrite would permit established carriers to operate in both competitive and monopoly markets and would thus perpetuate present difficulties of ascribing cost causation. It would, in effect, require some cross-subsidization since established carriers would retain their obligations to serve. Finally, the rewrite would give established carriers a vehicle for financing cross-subsidization since it would require all intercity suppliers to contribute to a fund for recovery of local exchange costs. This fund is intended to make local service affordable. However, it would also permit the intercity carriers to “left hand” established carriers to take losses in meeting competition, which would then be made up by payments to these same carriers’ “right hand” subsidiaries which operate local exchange facilities.

An examination of Bell’s market power or the concentration of such power in telecommunications vis-à-vis other industries seems to indicate that near term deregulation may be premature. For example, within communications the Bell System accounts for over 80 percent of the industry’s total assets, revenues, and employees. Although competitive entry has had regulatory sanction in the private line transmission and terminal equipment areas for nearly ten years, one would hardly classify even these market areas as being “competitive.” To date the degree of penetration in each of these areas still amounts to less than 6 percent of total sales revenues. Further, in comparison to the entirety of new companies operating in such markets, Bell has more than 35 times these firms’ total assets and 150 times their combined revenues.

Those who argue for communications deregulation often support their positions by citing similar experiences in the air carrier and trucking industries. Whatever the intrinsic merits of such comparisons, it appears that at best only imperfect analogies can be drawn. For example, the financial risks that are involved in precipitous deregulation of communications are hardly comparable to those of the airline industry or trucking. Bell alone has over six times the airline industry’s total assets and over eight times those of trucking. Moreover, no reasonable facsimile of Bell’s dominance in communications exists in either of these other industries. Whereas in communications, AT&T accounts for 84 percent of industry revenues, the largest airline and trucking firms account for only 18 and 9 percent of industry revenues, respectively. On the basis of total assets, these figures are 81 and 30 percent for Bell, but only 15 and 8 percent for the largest airline and trucking concerns.

On balance, it seems that if a workably competitive industry were awaiting its regulatory unshackling, the rewrite might provide an appropriate vehicle for unleashing market forces. In reality, however, communications is an industry dominated to an exceptional degree by one firm. Moreover, there is a high probability that this domination would continue and perhaps even spread to other industries in the wake of deregulation. Therefore, there appears to exist a need for a changeover or transitional period before legislation such as that proposed under HR 13015 can have an efficacious effect. FCC Chairman Charles Ferris has observed that “the domestic communications industry is presently in an era of transition towards greater consumer choice and broader, more diverse technological applications. … [However], this industry remains dominated by a single firm with resources and market shares many times those of recent and potential new entrants.”

Fortunately, there may be some lead time before pressures for a new regulatory approach in communications become inexorable. For
example, the FCC's ENFIA negotiations have established interim access charges. By virtue of these negotiations, three years of lead time may exist for determination of proper fees for connections to local exchange competitors to the established carriers’ local exchange facilities. Similarly, in the area of intersity services AT&T has indicated that no major restructuring of rates would be forthcoming, at least in 1979. In the time that remains, the FCC will have several important tasks in managing the transition to a new industry environment. During this period, a first order of business will be continued enforcement of provisions of the existing Communications Act of 1934 against practices discriminatory or anticompetitive behavior. This will provide a surrogate for the discipline of the competitive market forces that have yet to evolve. The FCC will also have to fulfill the objective of “clarifying market rules” which it espoused in Docket 18128. That decision established a costing standard and ratemaking guidelines to permit full and fair competition and to determine whether established carriers' rates are reasonable and nondiscriminatory. However, to be effective the standard adopted will require the support of a modern record keeping and reporting system based on a mixture of economic, engineering, and accounting data. Therefore, in the transitional period such proceedings as those involving revisions to the Uniform System of Accounts and specification of a costing manual will have to be completed in a timely manner. If suppliers are to act rationally in future ‘workably competitive’ markets, they must have proper data and decision rules for market entry, ratemaking, investment, and other purposes.

In addition to the foregoing, the FCC must develop methods of distributing and recovering the industry’s past facility costs and of valuing radio spectrum. The industry’s embedded costs currently total over $120 billion, but only about 20 percent have been recovered through past subscriber charges. Historically, the FCC has prescribed the rate of recovery of these costs, and, in part, current reserve levels reflect the commission’s past depreciation practices. During the transition to competition, the FCC will have to continue to administer the recovery process, particularly in view of evidence that the net book value of existing plant substantially exceeds its economic value. The commission will have to devise methods for recovery of these embedded costs which do not distort the true worth of the established carriers’ services. If this is not accomplished, substantial customer shifts will occur between these services and those of new entrants. Meanwhile, recovery must be accomplished within a relatively short period of time, or these costs will provide price advantages for use of incumbent ‘stand-alone’ systems which do not reflect these systems’ underlying economic advantages. Of course, unlike other competitors, stand-alone systems would not contribute toward recovery of such embedded costs if an access charge were imposed on all firms providing services through the existent carriers’ local exchange facilities.

The issue of determining a proper valuation and distribution of radio spectrum has long been treated as merely an engineering concern. However, this resource is becoming increasingly significant for the efficient usage of all factors of production, both within communications and between communications and other industries. The commission must devise a process to provide for the economic valuation and distribution of spectrum during the transitional period. Like all scarce productive inputs, this resource clearly has an “opportunity cost.” If no value is attached to spectrum, it will be employed at artificially high levels. Of course, increased usage will lessen the need for other resources, and may even cause a shift in production techniques. In an intraindustry context, this would make spectrum intensive systems such as those proposed by SBS and Xerox economically more attractive than they would be otherwise. In an interindustry sense, it would make productive processes using communications more attractive than substitutable processes which use the products of other industries. For example, false incentives would be created for shifts to such communications using services as electronic mail as substitutes for conventional mail delivery.

Finally, the FCC must continue to scrutinize established carrier service offerings and pricing practices and establish means of eliminating existing market segmentations and the potential for discriminatory or predatory pricing. Clearly, the commission’s resale and sharing decision in the private line area represents a major step in this direction. Therefore the public interest benefits that would accrue from generalized application of that decision should be seriously considered. Moreover, the commission should review tariff imposed restrictions on service availability. Removal of restrictions on such services as WATS, for example, would enhance competitive conditions. This would permit more efficient customer substitutions among a given carrier’s array of services and also between such services and those provided by other carriers.

**Regulation in the Long Run: A Few Afterthoughts**

Over a sufficiently long period, communications competition may reach a stage where no firm dominates pricing practices, the services available to any particular group of customers or region, technological development, or other market aspects. That is, suppose that at
Role of Federal Regulation

some point in the future competition matures and that the premia underlying deregulation are satisfied. In such an environment, would the efficient role for federal intervention in the industry? Moreover, if direct intervention was in fact unnecessary, would this imply that federal policy toward the industry should adhere to the principles of laissez-faire?

In general, communications policy might be concerned with only intraindustry (microeconomic) issues such as industry structure and pricing practices, or it might be directed toward interindustry (macroeconomic) issues such as communications' proper share of the nation's resources. In the past, debates over appropriate telecommunications policy have focused almost entirely on its intraindustry aspects. This emphasis may be reversed in the future. That is, given the increasing importance of communications products and services to the nation's economic health, policymakers may direct their attention toward planning macroeconomic intervention, under the assumption that microeconomic issues can be adequately settled by market forces.

Of course, planned macroeconomic intervention is hardly a novel idea for Western democracies, nor should it be considered alien to "capitalistic" enterprise. Such Western European nations as France, the Netherlands, and Sweden have engaged in national planning on an interindustry basis since World War II.12 These nations have retained private ownership of important sectors of their economies and have relied on market forces within industries to further public interest objectives and, in the case of Sweden, have actually established new firms in industries in which competition was lacking. In the area of interindustry resource distributions, however, the planning or national view is prevalent. In the interest of furthering national goals, these countries have directed financial capital, raw materials, and other resources toward industries which promote parastatal goals over those of competing nations. While this might result in the decline of certain industrial sectors, it has been expected that these declines would be more than offset by gains in other areas.

As the United States' natural advantages in timber, oil, and other raw materials diminish, it may be necessary to direct vital resources between industries much as the Western Europeans have done.13 These efforts may be crucial if we are to realize our national goals, balance of payments equilibrium, stable prices, and maximum employment and growth. A need for special emphasis on communications may exist because of the disproportionately large national investment that has been made in communications in the past and its high level of growth. As noted, industry investment totals over $120 billion and is currently growing at a rate of over $15 billion per year. Bell's share of this investment alone totals nearly 6 percent of the national sock of private corporate capital assets. Bell's investment has doubled between 1950 and 1960, 1960 and 1970, and 1970 and today.

From a planning perspective, one can question the efficacy of this extraordinary concentration of resources in communications in terms of the effects it has had on other industries. For example, has it been in the national interest to have the world's best telephone system, but also to have a comparatively antiquated rail system? More important, given that we have made this investment (and have chosen to wait for a train rather than a dial tone), are we making use of communications in ways which will work for our nation's best comparative advantage? It would seem that these policy concerns will provide ample opportunity for fruitful macroeconomic, rather than microeconomic, intervention in the future. Federal initiatives may be needed to ensure that full utilization is made of the communications network and that free substitution of communications for the products of other industries will not occur. Such substitutions are already possible in the "frontier" areas between data processing, mail services, and communications. Federal macroeconomic intervention can be useful to ensure that these substitutions are both timely and pervasive, and also to promote the considerable opportunities that exist in other frontier areas such as transportation and information or physical product storage.

Other instances of ongoing federal intervention in the "long run" will be frequent. These include radio spectrum allocation when franchise rights involve substantial issues of public safety and affect interindustry product substitution. Other means of allocation, such as market "auctions" to the highest bidder (AT&T) or engineering-based distributions, appear to be intrinsically flawed or inferior to a governmentally administered allocation based on national priorities. Besides spectrum allocation, continued federal intervention may be useful as a means of rectifying more persistent market imperfections or legal constraints (for example, the AT&T 1956 consent decree), for specification of reporting statistics and accounts, and for administration and surveillance of engineering standards. Specification of reporting requirements is important for the informational needs of investors, suppliers, and subscribers alike. Properly generated, such information is vital to efficient decision making, and thus should be kept out of the control of those with vested special interests. Engineering standards, of course, are important from a public safety and engineering economy standpoint. Federal administration of these standards will also prevent them from being used as anticompetitive devices.

Notes


4. Cross-subsidization and depression of a utility's rate of return would not occur after entry under all possible industry circumstances. However, those that have been present in communications would be likely to produce such effects. This industry's traditional milieu has been characterized by average pricing, often at the behest of regulatory authorities. Technological evolution has provided lower unit production costs, which have been available at the lower operating levels at which entry has occurred. Finally, lower cost customers generally have exhibited more elastic demand characteristics.

5. As in the case of single service category, it is conceivable that circumstances exist under which cross-subsidization and depression of a utility's rate of return would not occur.

6. The revenue requirements standard in algebraic terms can be expressed as follows:

\[ R = O + (V - D) r, \]

where:

- \( R \) = total firm revenue requirements,
- \( O \) = operating costs,
- \( V \) = gross value of tangible and intangible property serving the public,
- \( D \) = accumulated depreciation, and
- \( r \) = (allowed) rate of return.

7. FCC decisions in the terminal equipment and intercity transmission areas, as well as decisions affecting the "boundary" between communications and other industries, reflected these factors. See the *Hugh-A-Phone Corp. v. AT&T* (22 FCC 112, 115 (1967)) and *Garterone* (13 FCC 2d 420 (1968)) terminal equipment cases, *Allocations of Frequencies in Bands Above 890 Mc* (27 FCC 359 (1959)) in the intercity transmission area, and the Notice of Inquiry in Regulatory and Policy Problems Presented by the Interdependence of Computers and Communications Services and Facilities (7 FCC 2d 11 (1965)).

8. See *Specialized Common Carrier Services*, 29 FCC 2d 870 (1971), and *Domestic Communications — Satellite Facilities*, 35 FCC 2d 844 and 36 FCC 2d 665 (1972).

9. See (Executive I) *MCI Telecommunications Corp. v. FCC*, 561 FCC 2d 365 (D.C.Cir. 1977), and (Executive II) *MCI Telecommunications v. FCC* (D.C.Cir. No. 76-1635, decided 14 April 1978). Also see the 11 May 1978 order of the U.S. Court of Appeals.


Competition and Regulation in Telecommunication Services

Katherine E. Sasseville

No system for the provision of social wants is value neutral. It manifests the underlying value structure of the society within which it exists. Because our society is founded on basic values which are often in conflict with one another, its market systems will reflect that conflict. In other words, our telephone system reinforces certain social and political values, and it is inimical to others, but it represents in its present form a balancing of those values and the interests they represent, in a way which is acceptable to the body politic.

As the system changes, the balance changes, and we must come to grips with the social and political consequences of that change. Unfortunately, we are ill equipped to do so if we do not understand the underlying scheme of values and the relationship of those values to the present system.

Regulators and the regulated industry have failed to examine critically the function of the present system in terms of its real, that is, its value-directed, purposes in our economic and political society. Perhaps this failure has been unimportant in the past because, after all, it worked rather well, and we were not faced with the need to tinker with it to get it going.

But today, as we all recognize, rapid technological change and the competitive forces it has spawned are drastically altering the economics of the industry and requiring concomitant changes in regulatory approaches and methods.

We cannot possibly hope to deal intelligently with these changes without a basic understanding of the relationship of market systems to social values. It is up to state regulators, I think, to analyze the changes we face in terms of that relationship. Then we must agree on which new market structures and regulatory methods will, on balance, best support and perpetuate those social and political values which are of greatest importance to the American people.

Let me try to describe what I think those values are and how they conflict.

The Bill of Rights says, among other things:
Congress shall make no law respecting the establishment of religion or prohibiting the free exercise thereof or abridging the freedom of speech or of the press or the right of the people peaceably to assemble and to petition the government for a redress of grievances.

The right of the people to be secure in their persons, houses, papers and effects against unreasonable searches and seizures shall not be violated. No person shall be deprived of life, liberty or property without due process of law nor shall private property be taken for public use without just compensation.

That language makes it clear that the protection of individual liberty against the state, the right to hold and use property free of governmental interference, is of fundamental importance to the American people. And that notion finds expression in a whole complex of related values: the right to use property for private gain; freedom of individual thought; values that I will simply term individual rights.

Unfortunately, my rights as an individual cannot be maximized without interfering with yours. My right to swing my fist is limited only by the presence of your nose. And so we have learned that the other great principle of American thought — equality of rights under the law — is at the opposite end of the continuum. To the extent that we seek to maximize notions of equality, we must limit the rights and opportunity of the individual to be more than equal.

All of our laws seek, ideally, to balance those inherently conflicting values, to allow that extent of individual rights and freedom which is consistent with our desire to promote equality. The point at which the proper balance is reached, I submit, is the societal consensus as to what constitutes "justice," "fairness," and "equity."

I have used the word continuum, but the process is more comparable to a teeter totter. Society’s notions of justice are the fulcrum, and a
precious balance is established between individualism, on the one end, and egalitarianism, on the other.

Regulators in a time of change must understand the relationship between the economic system they are regulating and the underlying social values it manifests. If the value system may be seen as a teeter totter, the economic system that reflects those ideas may be seen in the same manner. At one end of the teeter totter we may put unfeathering, private, free enterprise, free to maximize its own interests, its private gain in providing some service or social want. At the other end we may put the counterpart of egalitarianism: public or government ownership, guaranteeing every person an equal right of access to the publicly provided service.

This is not intended to be a lengthy discussion of different forms of economic market systems. Let me briefly characterize, and simply try to dichotomize, the conditions which describe these alternative market systems.

Private enterprise is still considered to be the norm in our democratic capitalist system, although totally unregulated business enterprise, I suspect, is a thing of the past. But individualism still exists and creates a vigorous competitive market structure for the production and delivery of most goods and services, and it is commonly believed that this system contributes to efficiency, innovation, range of choice, a growing economy, and incentives to develop resources and allocate them efficiently. It is accepted, I think, that a free enterprise, competitive system is the "system of choice," absent some public policy reason militating against it.

In contrast, public or government ownership is the preferred system when there is a perceived public policy need to protect the welfare, safety, or health of the citizenry. When a service is deemed essential, that public policy reason exists. The fundamental duty of the state to protect its citizens and provide equal access to all to an essential service provides the justification for public ownership, and so does the existence of a natural monopoly for providing essential services.

Let me try to explain where regulation fits into this picture. Our American ideals of individualism, and especially of freedom of thought and freedom to use private property as we see fit, are so highly valued that the collectivist system as a means to enforce notions of equality is generally abhorrent to us, despite our deeply held and comparatively embedded notions of equality of opportunity and equal rights. Nevertheless, just as we limit the swinging fist, so we limit free enterprise when it impinges inimically on those notions of equality.

And that is exactly what happens when a market monopoly in an essential service exists! There is the potential for denial of service to certain groups. Thus the American system has developed methods to control the consequences of monopoly when it could not eliminate them. The existence of the "natural monopoly" in providing public utility services is necessary, we believe, for reasons of economy and efficiency; since the service being provided is essential, American values favored the development of government regulation. It is interesting that our strong commitment to the individual rights of investors limited us to regulation, rather than public ownership, which has been the chosen system in most other Western democracies, even those with strong private capitalism systems.

Regulation and regulators serve as the fulcrum, the balancing point, between maximization of the individual rights of the owner-investors and the rights of society to have equal access to the essential utility service.

We are familiar with the traditional view of regulation, that it is a substitute for competition and that its purpose is to force regulated enterprises to behave as they would under competition. That view is so widely held by regulatory economists that regulators can expect a Pavlovian response when they ask an economist: Why should we price on the basis of marginal costs? The inevitable answer: Because to do so will maximize the efficient allocation of resources in society. Efficiency seems to be the only tune to which the economist marches.

The traditional view of the purpose of regulation is not so much inadequate as it is incomplete. It is shallow and superficial because it totally ignores the other function of regulation. It focuses only on one side of the teeter totter. It misses the point that regulation is also a substitute for public or government ownership, and its purpose is to constrain a utility to behave as it would if it were owned and managed by the public it serves.

Viewed in this way, we can see the importance of American notions of equality and justice. When a service is essential to the public welfare, the government makes it available to all. Equal access to roads, education, and police and fire protection, without respect to ability to pay, is a fundamental tenet of American values.

The inherent conflict between the desire for equality and the notion of individual rights emerges again. The goals of one cannot be maximized without restrictions on the goals of the other.

Thus, if we truly sought to price every service at its marginal cost to replicate the pricing of the ideally competitive free market, we might allocate society's resources efficiently, but we would have to deny service to many people. That, in fact, is the genesis of the rural telephone system in the United States. On the other hand, if we were to guarantee to every single potential user access to the same system
and equivalent facilities, the cost to all would rise to prohibitive levels. Balancing such competing interests and societal values is the job of the state utilities regulator.

Before I try to deal with specific regulatory problems, let me make one more point about the inadequate way in which we have analyzed the existing system: Efficiency is the key to the balancing act. Efficiency is one subject of a cost-benefit analysis of competing values and goals, and no position can be accepted as just, fair, and equitable unless it is reasonably efficient. But efficiency has been described by most commentators on utility regulation not as one aspect of justice and equity, but as a necessary determinant of competition and an end in itself.

I suspect that this basic mistake has been made because so many economists live in a perfect world, where everything else is theoretically equal. And it can be agreed that under such perfect conditions, the effects of competition will force both the entrepreneur and the consumer toward efficient allocation of resources. But economists are the only ones who live in that perfect world. The rest of us, and especially regulators, live in a political maelstrom of change in which we are constantly dodging and ducking, seeking to meet the real needs of flesh and blood people who have very unequal resources.

In the real world, efficiency does not automatically result from competition. Consider corn flakes, for example, sold for $1.50 a pound, or about $84 a bushel. Since the price the farmer gets for that bushel is $1.79, that means that more than 97 percent of the price we pay for that cereal is going to the middlemen — processors, advertisers, distributors — and for taxes and regulation. If the competitive free enterprise system has added value 32 times as great as the value of the farmer’s product, then I could call it an efficient system. But in my view the added value of processing and distribution is not greater than the original value of the farmer’s efforts in producing the corn, and I can only wonder where the efficiency has gone.

But in the long run, neither a competitive system, a regulated system, nor a government-owned system can tolerate gross inefficiency, because inefficiencies and resulting high costs and prices destroy the perception of fairness and the legitimacy of the system itself. The results in the competitive market come quickly: ‘The consumer switches to another product. And if the entire market is subject to the same inefficiencies, the consumer limits or stops buying entirely.

An essential service raises different problems, but even in that case, and even with government-supplied services, consumers have recourse, and that recourse is in political action. It takes time, and people will long tolerate injustice before they will take forceful action. But our history provides numerous instances of citizen action, even uprisings, to right the perceived wrongs of an inefficient and unfair economic system. From the Whiskey Rebellion to Proposition 13, Americans have found ways to change the system to make it reflect, as it must, their underlying values and a just balance between competing values.

In fact, that is exactly what has been happening in this country during the last decade, and it is exactly why the effect of competition is a topic of discussion. The “people” have begun changing a system which does not meet their needs in a way that they perceived as equitable, and so the system is being changed. And those who are part of that system are being taken along for a wild ride and are struggling to regain control.

The public interest is hard to determine because there are so many publics. But economic self-interest is a forceful tool for separating those publics, and the relevant public has identified itself. Furthermore, it has begun reshaping the communications industry to suit its interests. The relevant public here is simply that group of users who can most benefit by the new technologies, and their existence is the stimulus that the competitive services and providers need.

Unfortunately, the other publics to not understand what is happening and are unable to protect their interests. Recognizing that fact, the industry and many regulators have fought a rear guard action to eliminate competition for these last several years, but industry has finally accepted the inevitability of competition and is now engaged in meeting it.

Unfortunately, many state regulators have been watching from the sidelines, unaware that the battlefield for competitive struggle is going to be in the back shops and the hearings rooms of the state public service commissions. Few are prepared to deal with the reality of that struggle.

I do not claim to have any solutions, but I do have, I think, some basic principles to guide me. I suggest that as we are faced with increasingly difficult regulatory questions brought about by competition, we may at least set our sights in the right direction, if we remember that our goal is still to weigh the conflicting needs and values of different interest groups so as to balance them in the public interest.

The obligation of the state regulator today is no different than it was under monopoly. The structure of the new and developing system must continually be analyzed in terms of the underlying values which support it, and the inherent tensions and conflicts must be identified and understood. There will always be a role in regulation for common sense. When the debate ascends to levels of abstraction beyond the ability of mere mortals to understand, it is still useful to ask whose ox has been gored.
Whatever the debate at the federal level, state regulators will still be sworn to uphold state laws which in words or substance usually call for the provision of service ability to all (universal service); to maximize efficiency and minimize costs (through just and affordable rates); and to provide continuity of reliable service (by compensating investors fairly). And they are always adjured to prevent undue discrimination (justice again), but this statutory obligation also serves the purpose of competition because it can be said to require that we price services reasonably closely to their identifiable costs in order to prevent undue price discrimination, and that, in theory, should stimulate innovation, new technology, and freedom of choice.

In fact, our regulatory goals and fundamental values have not changed to guarantee to the individual the right to use his property to maximize his individual interests, as investor or user, but to balance the right in an equitable manner with the rights of society, so as to protect the public welfare by making an essential service available to all and maximizing equality of the right of access to that service by requiring the efficient and nondiscriminatory recovery of the costs of providing the service.

The problem is not that goals or values have changed, but that rapidly developing technology and competition with its complexity of alternatives require that we find new approaches and methods to serve our permanent goals. But we cannot do that until we agree on what the terms mean, and thus the first step in addressing the problem is to define our terms. Perhaps this is the hardest step of all, because defining "universal service" requires us to perform our balancing function. Do we really mean every single home in the United States must have telephone service? Of course not! Think of the cost! So we balance the right of the individual against the cost to society. And we have to define who is included. And we have to define the basic, or least cost, service for which we will continue to protect monopolized territories and to average, or subsidize, or guarantee the cost to all those that we have included in our universe.

Whether we use a separations pool, or a universal access pool, or a state equalization aid pool, there will be a limit to its revenues and both competitive and political pressures to keep it low. And so we will have to define for ourselves a set of priorities consistent with reaching an equitable balance between the rights of the individual and society vis-à-vis: urban customers; rural customers; remote customers; residential; business; public-cord phone users; high volume users; low volume users; improving and upgrading service with new facilities or extending facilities; measured versus flat rate service; and so on.

The choices to be made about intercity traffic will require thinking through the meaning of local exchange and interexchange service and a new definition of the relevant community of interest with respect to "local exchange" costs.

The more broadly we define the "who" and the kinds of services to be included in our definition of local monopoly service, the higher its cost, and that is going to be the crux of the problem state regulators will face in the next decade. If I have not been explicit enough up to now, let me state my assumptions about the kind of telephone system state regulators are going to be regulating in the future.

I assume that competition will proceed inevitably in every sector of the market where it can develop a technology or a pricing scheme to save some consumers money while making a profit for the provider. And I assume that in balancing the rights of individual investors and consumers against the interests of society in providing basic communications service to all, legislators and regulators should seek to regulate and protect from competition only that irreducible minimum of service which is truly necessary and essential to the public welfare.

For that basic irreducible minimum (BIM) of service, we will have to continue to average costs, but for all other aspects of service we will seek to price at the cost to the user. Assuming away the biggest problem state regulators will face, that is, assuming that we know the limits of the universal service concept and that we agree on the BIM which we will continue to make available to all (even at subsidized rates), I want to raise some other problems with which state regulators must deal.

In the real world — the world of politicians, regulators, public officials, inflation, fuzzy-headed economists, and rate payers who rip telephones out of the walls — things do not work because they are supported by unimpeached testimony, or by theory, or even by brute force. Things work, systems work, laws work, if and only if they represent, more or less accurately, that delicate balance between the rights of all the individuals seeking to maximize their own vested interests and the rights of society as a whole to force individuals to conform to the norm, or the average, for the greater good of us all.

And in the regulation of public utilities, it will continue to be the job of the state regulator to find and support that delicate balancing point. I would say further that that point can be described in many ways depending on which set of concerns one is addressing. I have spoken thus far primarily about balancing social concerns. When we do that, correctly, our balancing point is called equitable. If our concerns are legal rights, and we succeed in finding that balancing point, we call it justice. And if our concern is between the rights of today's users to use resources and tomorrow's users to have resources available to use, the balancing point is called conservation.
But the concerns I want to address now are economic concerns, and if regulators correctly balance the rights of the individuals against the rights of society in matters of economic considerations, that point will be termed efficiency. (And if we achieve it, we will be called miracle workers.)

There is another way to describe what I am saying about the job of the state regulator: It is no different than it has been. Within the same context of social goals and statutes with which we have always worked, we have to ask the same kinds of questions, but we are going to be asking them about new methods and new technology because competition is here to stay. We still must perform the crucial role of the generalist, evaluating and translating the "public interest" and continuing to do our balancing act.

In the practical work of regulation we still have two fundamental questions to ask about every proposed new approach, methodology, or technique: (1) How do we keep the revenue requirements down to the lowest possible level consistent with adequate and reliable service and fair treatment of investors? (2) How do we fairly allocate costs and prices among different classes of customers?

Regulators have always had an obligation to promote efficiency by disallowing rate base items that were excessive and unnecessary, or not "used and useful in providing service," as our statute puts it. But new technologies and competitive forces mean that we have to start asking those questions anew. For example, do we have companies (probably cost setting companies) that have had incentives to invest in unnecessary equipment or plant under the present requirements process, and how should that "excess" investment be treated? Can operating efficiencies be improved by exchange consolidation? With more equipment being developed which is subject to rapid obsolescence, such as computer equipment, how can we develop appropriate depreciation rates and reserves so as to treat management, the investor, and the ratepayer fairly?

State regulators for the foreseeable future are going to continue to ask those kinds of basic, dull, revenue requirement questions so long as there is any aspect of monopoly service left to regulate. In my view that day will continue unless and until technology develops a cheaper way to provide BIM service. (Remember when they called it pots?)

The more exciting kinds of questions that state regulators will need to explore and examine appear to arise not so much in the revenue area as in the arena of the pie. How do you cut it? Who pays?

The central question is the one about separations and settlements, and I would like to throw out a thought to stimulate the debate.

It seems to me that the criticism about separations formulas is valid. We should seriously question the continued existence of these procedures and ask whether it would be more fruitful to abandon the formulas and retain the result.

Let us assume that intercity toll calling is to become unregulated, completely competitive at both inter- and intrastate levels. And assume that state regulators, desirous of minimizing local calling rates, approve tariffs that would charge an access fee to every local user who uses the local exchange to facilitate intercity calls. It would not matter whether that local user was AT&T, another telephone company, the local exchange company's "long lines division," MCI, the local high school, or my mother. Obviously, we could devise a variety of tariffs to reflect the varying costs associated with traffic volume, density, holding time, and so forth, that different users impose. But the maximum revenue those access charges could produce would be equal to the cost of duplicating the existing local telephone system. Because if they were higher, presumably we would develop a competitor for the local system itself.

What I am suggesting is that the local system's duplication cost is the maximum cost that should be paid for local switching, not only by the local users, but also by the interstate user and by the intrastate, intercity user. All three user groups have the same opportunity cost. Why not allocate the costs of the local exchange thus: one-third to local users, one-third to intrastate toll use, and one-third to interstate use? This theory has a great virtue: It is simple. It has another: It would not change over time as SFP and SLU factors changed. It has a final and present advantage that is not the same as the first two: it can represent a radical departure from present separations practice.

While I am tossing out outrageous and unworkable schemes for discussion, let me suggest another: Suppose we retain the concept of interstate rate regulation but in a modified form. We retain average long haul toll rate concepts only for low density routes without significant competition. The FCC or its successor would allocate the access pool for interstate calling to the states on a basis that included consideration of the difference between the actual costs and the assumed "average" costs. Finally, the states would distribute the revenues from the interstate and intercity access charges to local exchanges on a basis that included both considerations of cost and of need. It would function somewhat analogously to a state's education board of equalization in allocating "foundation aid."

Let me mention a few other state areas of concern. First, I agree absolutely with those who point out that all toll substitutes — EAS, FX, CCSA — must pay their share of the costs of local service, and that the ENFIA agreement is far from the final act in this passion play. And as
I hope my position on the opportunity costs of basic exchange service made clear, I do not consider that these access charges are a "contribution" to local service costs. They are the actual, fair share of the cost of service. Consider the law of real property: Although owners in common may each be wholly liable for the total costs of necessary maintenance to their commonly held property, owner A has an action against owner B for his fair share of those costs. In joint ownership of property, each party is jointly and severally liable for all necessary costs. The basis in law and in economics is that each party has total responsibility for costs. It is a theory of equity that reduces this total liability to a liability only for that share by which one actually benefits. In law, in equity, and in economics, each class of user, I would assert, should pay one-third of the joint and common costs of providing total telephone service.

Regulators cannot overlook the scut work in store for them: The need to separate costs fairly and assign them to the competitive or to the monopoly service. We have already seen filed a tariff for selling station sets and equipment which would have put all the revenues from this competitive service below the line, and all the "successful marketing" costs related below the line, and all the cost of unsuccessful marketing attempts and overhead above the line. This is only one example of what is to come.

State regulators, do not expect to get out of the trenches. Having insinuated that it is their job to protect the people against the monopolist and from the competitor, let me suggest in closing that state regulators have another continuing obligation. They must ensure a continuity of reliable service. What that requires is that we "protect the provider," that we assure fairness to the investor. The freedom to compete means the freedom to go broke, but it also means the hope of competing successfully enough to earn high returns. In other words, the introduction of competition means the introduction of significantly greater risks than have existed in the past.

What is going to be the effect on the cost of capital? Surely, we should be as concerned about minimizing the cost of that input factor as of any other. But how? Are we to guarantee not only BIM telephone service to the user, but also the right of the prudent monopolist who provides it to continue to exist? Should we make that an explicit social contract to reduce risks and cost? If so, how do we decide what is prudent when a company finds itself with an expensive and underpriced stock of "primary instruments" on hand and technologically obsolete? Are we going to have to require the corporate disestablishment of competitive and monopoly services? Can we count on traditional accounting systems to allocate costs adequately and distinguish between competitive and monopoly expenses and revenues and investment? Or do we need the FCC's proposed economist's version of a new system of accounts? Or will it only confuse us and lead us into a great accounting debate?

I have no answers to the questions I have asked. I am only suggesting that questions like these are the ones state regulators will be asking in the coming years as they try to fulfill their oaths to carry out the laws of their states and to ensure that basic telephone service continues to be available to all at affordable and nondiscriminatory rates.

In conclusion: The king is dead; long live the king. The system goes on, and it will go on, not just because the system is the solution, but because it is a social and political imperative. It meets the needs of the people in a way that has been largely accepted as fair. But the new king is already ascending the throne, and it is up to us to tailor the king. After all, none of us wants someone to say: Look, the king isn't wearing any clothes.
In contrast, many regulators and other decision makers prefer to rely on cross-subsidies precisely because they are hidden. They are much more comfortable regulating the firm in the conventional manner — involving an “obligation to serve” accompanied by the *quia pro quo* of protection against competitive entry — and letting the firm make the trade-offs itself. As long as the firm discharges its obligations to serve all comers at “reasonable” rates and avoids earning patently excessive profits, this regulatory scheme can persist indefinitely. Lack of clear competitive benchmarks to judge performance, those favoring the benefits of competition can quickly be dismissed as armchair theoreticians; those who pay the hidden subsidies cannot easily protest, since they cannot estimate, or in many cases cannot even be conscious of, the burden they bear; and intended beneficiaries of subsidies have no way of knowing how much they are, in fact, receiving. Both the firm and its regulators can enjoy a quiet life.

This regulatory scheme was pursued in the telephone industry for many decades. However, in the 1990s striking technological advances took place, largely outside the telephone industry, in the computer and electronics fields, and these fueled irresistible demands for market entry. In response to a long series of decisions by the Federal Communications Commission and the courts, competition is vigorously emerging in terminal and transmission markets, so that today lines are difficult to draw between services that should continue to be regulated and those that should be left to the full forces of the marketplace. Thus, regulators, legislators, and others no longer can avoid the painful task of dealing directly with trade-offs between efficiency and equity.

Fortunately, making these trade-offs may be less difficult in telecommunications than, say, in the energy field. In current policy debates, the impact of competition on local telephone rates is of concern, particularly with respect to residential subscribers unable to take advantage of competitive offerings. Using a crude but useful methodology, Hasselwander examines the potential impact of competition in intercity markets on local rates. He assumes that competition would reduce toll revenues to local companies to reflect the relative use of toll calls and other services using local exchange facilities, in contrast to the current situation, in which toll calls pay additional amounts given by the SPF factor in toll settlements calculations. With this reduction, he shows that local rates would not rise in real terms above those that generally existed in 1968. His tables illustrate that since general inflation, per capita price rises have far outpaced the increases in local telephone rates during the last decade, a substantial reduction in toll revenues would be required to make local users, as a class, worse off than they were a decade ago. His illustrative examples show persuasively, in my...
view, that in today's inflationary economy telephone service is one of the best bargains and is likely to remain so in the foreseeable future. This situation stands in marked contrast to that in the energy field, where heating and power bills for many consumers rival their monthly mortgage payments, and the policy debate is being conducted on far more slippery ground.

In fact, one has more reason for optimism about coping with conflicts between equity and efficiency than even Hasselwander suggests. An emerging consensus within government and industry supports the notion that an access charge should be imposed on all intercity services that use local networks. Until recently, only message toll service (MTS) and wide area telephone service (WATS) entered directly into the settlement procedures that Hasselwander mentions. Private line services competitive with MTS and WATS, including Bell's own CCSA and FX offerings, contributed nothing. In late 1979 an interim agreement was reached within the industry whereby specialized common carriers that compete with Bell and GTE will pay a charge for use of the local exchange facilities to ensure at least rough comparability with the contributions made by their MTS and WATS. Today, attention is being directed to questions of how Bell's own private line services can best be included as part of a longer term agreement. With broad coverage afforded by access charges, basic local rates could be insulated to a substantial degree from whatever competitive pressures emerge in intercity markets. Thus, local telephone rates, alongside a highly competitive intercity market, might remain lower in real terms than they do in Hasselwander's calculations.

Moreover, by changing the way local telephone service is priced, the industry may achieve more efficient use of local exchange facilities, while assuring continued easy availability of service to residential users. By adopting usage-sensitive pricing, in contrast to the prevalent flat-rate and extended area service, companies could impose a low access charge to the network, perhaps lower than current levels, and cover remaining revenue requirements by additional charges on users. Such pricing techniques involve additional costs of measuring equipment. Moreover, this approach is resisted by heavy telephone users who are able to transfer some of their costs to others under flat-rate pricing. However, we will likely see increased adoption of measured pricing as telephone companies convert to more modern switching equipment adaptable to measuring local usage. Debates about equity and efficiency will continue, but given the relatively small dollar magnitudes for individual users (again in contrast to the large individual transfers at stake elsewhere), I surmise that decisions at the federal and state level will increasingly reflect considerations of efficiency.

Boiter discusses a number of issues raised by the moves of the Federal Communications Commission to permit competition in intercity services, moves that particularly affect the Bell System. I will concentrate on three of the many topics he treats. First, Boiter has good reason to be concerned about problems of cross-subsidization that arise when a firm with a strong monopoly position in some markets is able to compete in others. Many fear that Bell, with its huge and lucrative MTS market, will price its competitive private line services at predatory levels in order to dominate competitive markets as well. Designing against cross-subsidization is extremely difficult, much more so than assuring reasonable rates for basic local service. To protect against cross-subsidization, Boiter urges that we continue striving to develop cost allocation standards and better accounting systems that effectively segregate costs between monopoly and competitive services. As he emphasizes, "to be effective the standard adopted will require the support of a modern record keeping and reporting system based on a mixture of economic, engineering, and accounting data."

I would have been more sanguine about the prospects of meeting these requirements a decade or so ago. Unfortunately, ever since Bell filed its Telpak tariffs in the early 1960s (its first major competitive response in interstate markets), the FCC has been unable to mandate successfully accounting procedures and sound cost standards that would assure adequate protection against cross-subsidization. Despite massive studies, prolonged hearings, decisions, appeals, and new decisions, the outcome of Telpak remains in doubt. At this writing, the decision by the commission to employ one particular cost methodology (so-called FDC-7) has been appealed to the courts, where it is pending.

The most basic problems with using the cost accounting approach to ensure against cross-subsidization are that it must deal with questions of allocating common costs among multiple services and forecast market and cost data. In seeking answers to these questions from the regulated firm, and in reviewing them for accuracy and reliability, the FCC is inevitably placed in the position of trying to second-guess entrepreneurial decisions. The 17-year-old Telpak and other prolonged proceedings surrounding Bell tariff filings attest to the difficulty, if not the impossibility, of that task. Continued pursuit of this cost accounting approach will, I fear, embroil the commission and the industry in more years of controversy, delay, and uncertainty. Without going into details here, I am becoming increasingly persuaded that a structural rather than an accounting approach will be needed to cope satisfactorily with cross-subsidization as competitive pressures continue to grow.
Second, Bolter points out the difficulty of moving toward a more competitive market structure because of the current overwhelming dominance of the Bell System. His observation that "Bell has more than 350 times these firms' [new market entrants'] total assets and 13 times their combined revenues" provides good reason for pause among those who envision rapid erosion of Bell's dominance. It is true that the FCC's recent decision to permit resale by outsiders of Bell's bulk private line offerings (such as Telpak) will reduce opportunities for price discrimination. Reinforced by possibilities of resale, the threat of continued market entry may reduce Bell's actual market power, measured in terms of price elasticities of demand, below the level that one would infer by looking only at relative market shares. Nevertheless, as Bolter points out, "despite regulatory sanction of market entry in the private line transmission and terminal equipment markets for nearly ten years," competition has so far nibbled largely around the edges. Barriers to new entrants may still be too high to spur widespread predatory pricing and inordinate delays within the FCC and the courts, we can expect increasing competitive pressure, with outsiders accounting for a progressively larger market share. But it remains to be seen whether at the end of any particular period—ten, fifteen, or twenty years—Bell will be regarded only as a large, and one of many, competitors.

Finally, Bolter is concerned that total investment in the industry totals over $120 billion and is growing at a rate of over $50 billion per year. He notes that "Bell's share of this investment alone totals nearly 6 percent of the national stock of private corporate capital assets." In response, he suggests that macroeconomic controls may be justified in controlling levels of investment in this industry, in contrast to our nearly exclusive reliance on microeconomic regulation—a point of view that frightens me. If the government tries to decide investment priorities and levels for major industries, keeping a special eye on telecommunications, we could be opening ourselves to the enormous potential of mischief by central planning authorities. It is true, as Bolter says, that "planned macroeconomic intervention is hardly a novel idea for Western democracies." But I see no successful model drawn from Europe or elsewhere that could be applied with confidence in the United States to further national goals. In fact, it is the very difficulty that nations have had in setting national priorities and in forcing particular industries into a predetermined mold that gives me pause.

Bolter asks whether it has been "in the national interest to have the world's best telephone system, but also to have a comparatively antiquated rail system?" But we could just as well ask whether it has been in the national interest to have a comparatively antiquated rail system, on one hand, and clogged, smog-ridden freeways, on the other. With appropriate economic, health, safety, and environmental regulations, plus appropriate antitrust safeguards and a tax structure that avoids perverse behavior—all operating at the macroeconomic level—resources will flow into areas where they are most valued. If, under those circumstances, a substantial portion of total capital resources goes into telecommunications, or into any other industry, so be it.

In short, let us by all means continue macroeconomic intervention in tackling such problems as general price inflation and unemployment. We will have more than enough to keep us busy in the telecommunications field by perfecting and using wisely the full set of microeconomic tools.
Comments

Jeffrey Rohlfs

Let me say that I am not a spokesman for the Bell System. The comments I am expressing are my own views, not a consensus of Bell System policy makers. My remarks are on the transition to competition. I want to give some historical perspective on the issue and look at the economic performance of various industries in the U.S. economy.

One can reasonably argue that competitive industries have generally served consumers well. In agriculture, the costs have declined dramatically over the years. Similarly, in electronics the hand calculator is a dramatic example of benefit to the consumer through competition of many firms.

At the opposite extreme, regulated monopoly has also served the public well. I think it is fair to say that the telephone industry has performed well through the years. Similarly, electric power has had an excellent record of productivity gains (although it has faced serious problems in recent years, aggravated by the entry of players other than the public utilities and their regulators into the game, namely, the environmentalists and, more recently, the energy czars).

In contrast, a third form of industry structure, regulated competition, has been an unmitigated disaster. In the railroad industry we observe that a large fraction of the firms are in receivership, and industry performance is an international disgrace. In airlines, we are just now learning what we could have had all along if the industry structure had been more fortunate.

I feel that this point cannot be overemphasized. As I see it, the telecommunications industry is being pointed directly down the path the railroads followed. The key issue in the transition to competition — indeed, in the telecommunications industry as a whole — is how to get the industry off this track.

I want to address this issue, but let us first consider why regulated competition has worked so badly. I think it is easy to understand. Suppose there are many firms in an industry with rate-base regulation. The regulator may feel obligated to protect these firms from downside risk because rate regulation limits the firms' ability to reap windfall gains. This protection may, in fact, be necessary to attract new capital into the industry. Furthermore, the failure of a sizeable regulated firm, with the concomitant dislocations, is a great embarrassment (not to mention political liability) to regulators.

Now, suppose that one of these firms develops an innovation which, if carried out, would greatly increase its market share and thereby hurt the other firms. The regulator would be strongly tempted, if not legally obligated, to protect the other firms. He can either bar the innovation altogether, or allow it but otherwise protect competitors and thereby prevent the originator from profiting by his innovation. In either case, the regulator has taken the first step on a long downward path. I do not want to say where it ends, but one of the milestones on the road is a large number of firms in receivership and a disgraced industry.

In telecommunications, the structure that is evolving is a bit different from that of the railroads. Nevertheless, there are still very serious problems. As I see it, Bell is likely to continue to be regulated, but the competitors will have substantial freedom in their pricing.

Suppose a competitor innovates. Bell would expect to be protected so that it could continue to earn a fair rate of return, as previously discussed. The concern at Bell would be that this protection might be lacking. Indeed, this was often the case with the railroads. The concern of competitors would be that they would be prevented from benefiting from their innovations. Perhaps both are right. In the railroad industry, there were no long-run beneficiaries to regulated competition, and there is no reason to expect that there should be any in the telecommunications industry.

Suppose that Bell innovates. That is, it offers a new service or a lower price based on cost-saving technology. It will inevitably face the challenge that this new service or lower price is being subsidized by
monopoly services. Competitors may well find that the best way to maintain their competitive position is not to offer low-cost quality services responsive to customer needs, but to have a good staff of lawyers to make trouble for their enemies.

I want to emphasize that the problem is not that regulators are bad intentioned, shortsighted, or dull witted. They are most often the opposite. The problem is that the merger of competition and regulation creates difficulties that are very hard to handle in the political environment of regulatory commissions.

There has been a statement of the problem; let us discuss a possible solution. I want to emphasize again that this is not a Bell System proposal. My suggested solution follows directly from my previous remarks — that the problem is not competition or regulated monopoly, but the merger of the two. The solution is economic deregulation of the competitive part of the telecommunications business. That is, Bell and other telephone companies, as well as anyone else, would be free to offer whatever services they choose in the competitive sector and would be entitled to whatever profits they could earn in the free market. (This entails, among other things, releasing Bell from the 1956 Consent Decree.) Noneconomic regulation, analogous to that of the FAA need not be affected. This would include such activities as defining and enforcing standards for electrical interconnection among products sold by competitive firms.

There are three aspects of economic deregulation that I want to discuss in detail: The boundaries between the competitive and regulated sectors, the transfer prices they charge each other, and the transition to this long-run solution.

It is fairly easy to define a boundary with respect to a competitive terminal equipment sector. For example, we can simply say that the jack is in the regulated monopoly sector, and whatever terminal is plugged into the jack is part of the competitive sector. This is essentially the same boundary as in the latest version of registration.

For long distance, defining boundaries is considerably harder. However, among the innumerable ways of doing so, each is better than regulated competition. One possibility would be to designate regions, each defined precisely. All telecommunications traffic between any pair of these regions would be competitive and deregulated. Other traffic would be served by regulated monopoly.

A second set of accounting boundaries between the monopoly and competitive sectors must also be defined. Bell and other telephone companies would be operating in both regulated and competitive markets. Therefore, one must determine what part of the total expenses of these companies is to be recovered in the regulated sector. Also, on what part of their total investment are they entitled to a fair return in the regulated sector? Telephone companies would not be entitled to any particular amount based on their remaining (competitive) costs. In fact, no firm operating in the competitive sector would be "entitled" to anything from that sector, not even the right to stay in business.

A good way to define these accounting boundaries is as follows. When constructed, a facility is designated as entirely regulated or entirely unregulated. This decision would be based on whether it is expected to serve primarily regulated or primarily competitive markets. If the facility is regulated, all expenses and returns associated with it are recoverable in the regulated sector. Otherwise, none are.

This method of defining accounting boundaries advantageously avoids the most troublesome allocation problems. There would still be some allocation necessary — for example, joint administrative costs. For these, any reasonable, albeit arbitrary, allocation method (perhaps FDC7) would be satisfactory, provided that these costs are a relatively small part of the total.

The second aspect of this proposed solution involves transfer prices. These would work two ways. First, the competitive sector would sometimes use the local network to complete toll calls. It would have to pay a charge for this usage, and that charge would be subject to regulatory review. The charge might be greater for toll than for local usage.

Andy Margeison and Alan Haselwander have pointed out elsewhere in this volume, toll substitutes create difficulties in the long run if the differential between local and toll charges is substantial. In any event, a network access price would provide a short-run stopgap in making the transition to stable competitive pricing. Obviously, Bell's competitive sector and other competitors should pay the same charge, whatever it is. Second, transfer prices would work the other way. Currently, AT&T uses the toll network for local tandem switching. In the proposed solution this would involve a transfer price that would not be subject to regulatory review. However, regulators would ensure that that is the cheapest way for the regulated monopoly to handle this local traffic — given the transfer price.

Such a system of transfer prices greatly facilitates network planning. Separate facilities need not be constructed for competitive and regulated markets. Rather, the telephone companies can plan so as to minimize total costs.

The third aspect of the proposed solution is the transition period. Perhaps its most important characteristic is that it must have a specified termination date. Otherwise, political forces might cause it to last indefinitely. During the transition period three functions must be
performed: The public must be protected, Bell must be protected, and other competitors must be protected.

The public must be protected against the possibility of Bell's raising prices in the competitive sector and thereby realizing short-run profits before competitors can gear up and set the situation aright. I recommend the CAB solution to this problem, that is, allowing rates to be raised only to a limited degree per year without regulatory approval. Thus, before any substantial change could occur, competitors would have plenty of time to move into the market and prevent further increases.

Bell must also be protected during the transition period. It must have good prospects of earning a fair return on already embodied capital. I believe that this is simply a question of providing fair ground rules and allowing Bell sufficient pricing flexibility during the transition period.

Third, competitors must be protected. There are two aspects to consider: predatory competition and the infant industry argument. Let me say that I am bearish on the use of public policy to prevent predatory competition. I think that in the past fifty years there have been very few cases in which successful predatory competition has been demonstrated. But there have been many instances in which zealous businessmen have prevailed in the courts and protected the consumer against lower prices. In any event, preventing predatory competition is the responsibility of the Department of Justice, and no one has accused it of being too cordial to the Bell System in recent years.

The second aspect of protecting competitors involves the infant industry argument. Although it has internal logic and consistency, some historical perspective is useful. The application of this argument has a terrible record. It has been used to support the most awful economic policies — most notably to provide tariff protection for trusts that were already exercising substantial monopoly power.

In telecommunications, I think that protecting inefficient firms would set a very bad precedent. Competitors might incorrectly assume that the protection would continue to exist in the long run; worse, they might correctly assume that the protection would continue in the long run. That is, political forces may perpetuate the protection of inefficiency.

This is the end of my proposal. Let me emphasize again that it is a first draft. I intend it as a focus for discussion. It is not a complete solution and almost surely not the best solution. However, it is vital that we work out some solution to get us off the well-worn track the railroads traveled.

In conclusion, let us recall that the FCC introduced "controlled" competition as a regulatory tool. However, this tool has had a terrible record in U.S. history. Through the years, it has not shaved much wood, but it has cut off a lot of fingers and thumbs. One can sensibly regard competition, not as a regulatory tool, but as a vehicle for deregulation.

I earlier mentioned a long downward path. Perhaps it is paved with the good intentions of regulators who are trying to make the industry perform just a little better and cannot let go.
Part Eight

New Issues in Gas Regulation
Incremental Pricing and the National Energy Act

John E. Holtzinger, Jr.

One of the more controversial sections of the Natural Gas Policy Act of 1978 (NGPA) is Title II, which establishes incremental pricing rules. Unlike some other areas of the NGPA, in which Congress is quite specific, the provisions in this section are relatively general and allow the Federal Energy Regulatory Commission (FERC) substantial latitude in terms of time and the final shape of the regulations to be adopted. Nonetheless, Congress has determined, as part of the natural gas compromise legislation, that natural gas sold for certain industrial uses will be "incrementally priced."

Exactly how the incremental pricing provisions will work is unclear. There were no legislative hearings concerning the incremental pricing provisions that emerged in the compromise. There are no data readily available indicating the current volumes of sales that will be subject to incremental pricing provisions. Complicated rate and administrative problems would be involved in implementation of any incremental pricing provisions, and those outlined in the statute are no exception. For example, how does one trace incremental costs incurred by a pipeline supplier when two or three wholesale transactions may occur.
before a retail sale? How does one treat volumes that are cycled through storage? Perhaps the most serious problem is the compilation of a valid database for implementing incrementally priced sales. In the pipeline curtailment cases, assembling end-use data on individual pipelines took years, and some major questions are still unresolved.

As pipelines will undoubtedly incur substantial incremental gas-purchase costs, it is likely that pipeline rate proceedings will become heated contests about who will pay them. Assume, for example, that a pipeline has incurred incremental costs of $10 million and sells gas to two distributors with equal volumes of incrementally priced industrial sales. Assume also that the first distributor has already priced those sales at the alternate fuel cost, whereas the second is making industrial sales at well below the alternate fuel cost. The explanatory statement of the con-

fers, which accompanied the NGPA, indicates the second distributor will be required to price his industrial sales at alternate fuel prices, and in so doing he will absorb a portion of the pipeline's $10 million expenditure. In effect, allocation of the burden of incremental costs will be based on the varying ability of distributors to pass such costs along to industrial customers in the incremental pricing category.

Cost allocation and rate design historically have been hotly contested in pipeline rate cases. Incremental pricing issues, revolving around which pipeline customers are in the best position to pass on increases to their industrial markets, will add much fuel to the fire.

Background

The issue of incremental pricing is not new to the gas industry. In 1959 it arose in Trunkline Gas Co.,1 in which a natural gas pipeline company purchased its supplies in a new production area at significantly higher prices than previously. One customer objected to the pipeline's proposal to roll in the higher priced gas with other lower priced purchases because that customer would not receive any additional supplies as a result of the new purchases and did not want to be assessed any of the burden. The Federal Power Commission (FPC) approved the rolled-in pricing concept, which decision was affirmed by the Court of Appeals for the District of Columbia Circuit in the landmark Battle Creek case.2 The Court stated that rolled-in pricing had "many apparent advantages" and "avoids the enormous impractical burden of having to assign a different portion of the cost to each of a large number of customers."3 Moreover, use of rolled-in pricing "serves the interest of equal treatment for customers receiving equal service."4

In the early 1970s, when base-load liquefied natural gas (LNG) pricing became an issue, the commission changed its approach and specified that the LNG imports, which then appeared to be priced at a level well above conventional domestic supplies, should be priced incrementally. In Opinion No. 622,5 the commission went so far as to order that the LNG should be priced on an incremental basis at the burner tip, even though the commission then did not have jurisdiction over retail sales in most instances. Upon applications for rehearing and strong protests by almost every party, the FPC abandoned the burner-tip incremental pricing proposal and provided, in Opinion No. 622-4, that the gas should be incrementally priced at the city gate, where sales by the pipelines were made.6 This decision was ultimately overturned by the Court of Appeals for the Fifth Circuit primarily because the commission's innovative approach was unsupported by substantial evidence.7 On remand, the commission approved rolled-in prices for the project, largely because its policy was to encourage all forms of supple-

mental supply.8 In a later LNG import case, the FPC first required incremental pricing, then retreated when faced with arguments that the project could not be financed unless costs were rolled in.9

Since these early cases, the incremental pricing issue has been a significant factor in LNG import applications and has generally surfaced whenever any high priced supplemental supplies of gas are discussed.

The 1975 and 1976 congressional attempts to deregulate natural gas included several versions of incremental pricing.10 The recent compromise resulting in the passage of the Natural Gas Policy Act of 1978 included yet another. Although strong objections were raised about the workability of these pricing provisions,11 it is fair to say that incremental pricing probably was a key to the adoption of the bill.

The legislative version of incremental pricing radically diverges from pure economic concepts. For example, the NGPA specifically excludes transportation costs of Alaskan gas from incremental pricing. The available supply of Alaskan gas will not, therefore, depend upon a comparison of the incremental cost of that gas with expected revenues. The legislation also imposes a ceiling on incremental pricing equal to the price of alternative fuels, in order to limit the loss of customers. In that sense, incremental pricing does not serve to dampen demand. In fact, incremental pricing is a misnomer; the legislation actually provides for a species of end-use pricing. The intent of this paper is to focus on some of the problems associated with implementation of the complex legislative provisions and to identify some of the crucial decisions that must be made.
Incremental Pricing

Incremental Pricing Provisions of the
Natural Gas Policy Act

Purchases Subject to Incremental Pricing

Sections 201(a) and 203(a) of the NGPA allow the interstate pipeline purchaser to pass through certain costs associated with acquisition of several types of natural gas: (1) new natural gas above the threshold price; (2) new, onshore production well gas above the threshold price; (3) new LNG imports above the threshold price; (4) new NG imports above a specified quantity and above the new gas price; (5) stripper well natural gas above the new gas price; (6) high cost natural gas above 130 percent of No. 2 fuel oil cost; (7) first sale acquisition costs of Alaskan gas above the ceiling price; (8) increased state severance taxes; (9) purchases from an intrastate pipeline above the threshold price; and (10) surcharges paid to other pipelines.

The NGPA calls for rolled-in pricing of transportation costs and certain first sale acquisition costs of natural gas production from Alaska. The reasons for this exemption from the incremental pricing rule were clearly stated in the Conference Report: “The conference agreed to provide rolled in pricing for natural gas transported through the Alaska Natural Gas Transportation System and for the cost of transportation because they believed that private financing of the pipeline would not be available otherwise. Rolled in pricing is only Federal subsidy, of any type, direct or indirect, to be provided for the pipeline.”

The exclusion of Alaskan gas transportation from the incremental pricing rule poses several interesting questions, especially the extent to which this treatment will result in preference of Alaskan gas over other sources. For example, transportation costs of new supplies from Mexico and western Canada presumably would be incrementally priced, even though Canadian gas might be transported in the same facilities as Alaskan gas.

Sales Subject to Incremental Pricing

Section 201(b) provides that incremental pricing shall be applied initially to the boiler fuel use of natural gas by any nonexempt industrial facility. Exempt facilities, as defined in Section 206, are: (1) small boiler fuel users (less than 300 Mcf per day average); (2) agricultural use (for agricultural production and for feedstock to produce fertilizer chemicals, animal feed, or food); (3) schools, hospitals, or similar institutions; electric generation; electric cogeneration; and (4) any other facility or category provided for by FERC.

Implementation of the incremental pricing provisions of the NGPA will occur in two stages. Within twelve months of enactment, the commission must promulgate regulations for passing through costs to large industrial boiler fuel users of natural gas. Within eighteen months of enactment, the commission must amend its regulations to expand the category of industrial facilities subject to incremental pricing. The amended regulations are subject to a veto by either house of Congress.

Conflicting Goals

Incremental pricing, in theory at least, would force users of natural gas to make a rational economic choice between the cost of gas and the cost of alternative fuels. Much of the opposition to incremental pricing in FERC proceedings and elsewhere has been the concern by gas marketing companies that imposition of marginal pricing would result in the loss of industrial markets to alternative fuels and would result in higher costs to residential and commercial users because the contribution to fixed costs previously made by the industrial customers would no longer be available.

Some advocates of incremental pricing approved of these results. It was desirable to eliminate sales to industrial users that could use alternate fuels and would do so if they were required to pay the full cost of new gas supplies, and it was just as desirable (although less expedient) to increase the price of gas to higher priority users to the levels indicated by the true cost of new gas supplies. In his textbook on utility price concepts, James Bonbright discusses the use of long-term marginal cost as a pricing tool to avoid construction of new capacity that is not needed by the market. The merits of marginal pricing in the utility field have been widely disputed. Much emphasis has been placed on the need for prices to send the correct “signal” to the market, and various concepts have emerged, such as the “doctrine of inverse proportionality,” to adapt the economic theories of marginal pricing to the practical aspects of rate making. The incremental pricing provisions of the NGPA do not settle these issues; they attempt to agree with all sides.

In the NGPA, Congress legislated higher new gas prices and a gradual phasing out of gas price controls because, in its view, higher prices were needed to encourage increased supply. There can be no doubt that Congress wanted to augment gas supplies, eliminate the dual market structure that had developed under regulation, and place gas on an equal footing with other fuels, primarily oil. It should follow,
then, that Congress would not be concerned if the pricing of gas resulted in decreased industrial use. The whole purpose of deregulation seems to be to allow supply and demand to reach equilibrium at whatever price is required.

It is clear, however, that legislators had an overriding desire to shelter residential and commercial customers from the higher gas prices that were the intended results of the NGPA. Congress, therefore, designed the incremental provisions to require industrial users, beginning with those having alternate fuel capabilities, to absorb the price increases. Congress had no intention of permitting incremental prices to allocate or restrict demand. Indeed, both Senator Henry Jackson and Congressman John Dingell, two highly influential backers of the NGPA, made it plain that the loss of industrial sales was not a goal of the incremental pricing provisions and would not result from their implementation. They stressed that the provisions would reduce residential and commercial gas costs below the price produced by current pricing mechanisms because the industrial market would absorb much of the price increase.

The congressional conclusion that both goals — higher gas prices without loss of industrial sales and without a large increase in residential and commercial rates — were attainable may not be well founded. What Congress may have overlooked is the actual state of the gas market. From 1970, when curtailment of interstate gas pipeline sales commenced, until 1978, there has been a major increase in the price of natural gas. Table 1 illustrates that this change has been dramatic at the wellhead, the city gate, and the burner tip. Although average prices to industrial users generally are less than the average prices to residential and commercial users, there is no way to determine the extent to which further price increases to industrial users are economically feasible.

Most curtailment cases were based on end-use data for the early 1970s, when boiler fuel markets were at a peak. Both price increases and curtailment of so-called low priority boiler fuel sales caused part of this market to disappear. Because the size of the market subject to the first incremental pricing rule has been diminishing since collection of the end-use data, those statistics do not indicate the actual size of the boiler fuel market today. Furthermore, under the Fuel Use Act, another component of the National Energy Act, new plants dating back to 20 April 1977 cannot use natural gas without a strong showing that coal is not a feasible choice. Even though no numbers are available on the size of the current market for boiler fuel gas, it is certain to be far below the level of the early 1970s, and addition of new major fuel-burning installations will not cause it to rise.

As the market for incrementally priced gas shrinks, the ability to pass price increases on to industrial users will diminish, and the probability of price increases to residential and commercial customers will rise. It is possible that the effect of incremental pricing will be to nullify the basic objectives Congress sought to achieve in passing the NGPA.

**Implementation**

The FERC is required to adopt a rule within one year of enactment implementing incremental pricing. The implementation problems are formidable, and the one-year timetable may be insufficient, as the statutory framework for establishing incremental pricing does not fit well with the structure of the industry or with existing regulations. Some of the problems the commission staff can expect to encounter in the next twelve months are discussed below.

**Data Problems**

**END-USE DATA.** The NGPA provides that incremental pricing will apply initially to industrial boiler fuel facilities and subsequently to other nonexempt industrial facilities. One of the more complex problems involved in implementing these pricing provisions will be establishing an adequate survey of incrementally priced industrial facilities or users. In curtailment proceedings, this problem was approached on a pipeline-by-pipeline basis. Since the commission did not assert jurisdiction at the retail level, data collection from industrial users was indirect, through survey of gas distributors. The incremental pricing
provisions, however, will require more specific identification of incremental facilities before even the first rule can be placed in operation. An interagency task force at the Department of Energy (DOE) is now developing the reports that will form the data base for the FERC's incremental pricing rules.

In addition to identification of incrementally priced facilities and users and establishment of the criteria for determining facilities that are statutorily exempt, some provision for periodic reporting of end-use data probably will be needed. Data collection with this degree of sophistication has never been achieved. End-use data collected for curtailment cases covered a single, unchanging base period. Each distribution company purchasing gas from a particular pipeline assembled an end-use profile, which showed the priority of use for each customer's market during the specified base period. The sum of all base-period requirements reported by end-use priorities by the distributors was balanced against supplies then available, with curtailment of lower priority volumes ordered first. The base-period approach was used only to allocate supplies among pipeline customers and did not attempt to control the actual end use at the boiler tip. A distributor could, and many did, continue to make sales to users of lower priority than was indicated by the depth of the pipeline's curtailment.10

Incremental pricing is likely to require continuous reporting of end-use data, for it is doubtful that incremental costs could be allocated properly using stale base-period data. Distributors and state agencies active in pipeline rate matters are not likely to acquiesce in an approach that allocates costs on any basis other than actual sales. This is particularly true in the common situation in which competing industrial users purchase gas from the same ultimate pipeline source. One industrial customer is not likely to accept incrementally priced gas if his competitor can acquire gas at rolled-in prices merely because the latter is served by a distributor with a better base period.

The accumulation of data on actual use will not be easy, however. Industrial customers have been free to use their gas supplies for any category of end use, from boiler fuel to process, and thus it has not been essential to record actual use. In fact, there is no evidence that industrial users can determine exactly how much of the gas delivered to a large plant with various fuel uses goes into its boilers, as distinguished from gas used for process or other nonincrementally priced loads. Individual metering of boiler-fuel sales may be a necessary adjunct to incremental pricing.

ALTERNATE FUEL CEILINGS. A crucial FERC decision required to implement the incremental pricing rules of the NGPA will be determinination of the proper ceiling price applicable to a given pipeline's incrementally priced market. The commission's ruling is of pivotal importance because, as a practical matter, such ceilings will be reflected in pipeline and distribution company rate structures for industrial sales. Instead of the well-worn arguments with respect to cost allocation and rate design, rates will be set on the basis of what the traffic will bear, given the alternate fuel situation, and will be subject to change in tandem with a change in prices of the alternate fuels.

Detractors of the incremental pricing provisions contended that passing through incremental costs would cause industry to convert to foreign oil or move to gas-producing states. In his speech supporting the gas compromise, Senator Jackson argued that this was a myth:

This statement evidences a complete lack of understanding of the operation of the incremental pricing provision in the conference report. The legislation requires boiler fuel users of natural gas to pay the increased costs of more natural gas supplies if the interstate pipeline's acquisition cost exceeds $1.48 per million BTU's plus inflation. This mechanism must be put in place within a year. The legislation also requires FERC to come up with a plan to provide a similar mechanism for other industrial users. This second phase will not go into effect if Congress disapproves.

It is critical to understand that incremental pricing will operate only until the price paid by a facility covered by the rule is equal to the BTU equivalent of the cost of substitute fuel. FERC has a very flexible mechanism for determining the cost of substitute fuel on a regional basis. They can set the price anywhere between the cost of No. 2 fuel oil and the cost of No. 6 fuel oil. They can recommend an exemption for virtually any facility, with the exemption, which may be complete or partial, being subject to congressional review.

The incremental pricing mechanism is in no way intended to drive [away] industrial users of natural gas. Nor will it do so. It will make them [ cognizant] of the price they are paying for natural gas and it will make pipelines acutely aware of the prices producers ask them to pay. Nor is the incremental pricing provision likely to move factories to gas-producing States. In fact, the only hope of keeping factories in nonproducing States is to enact new legislation which is the only way they will get new desperately needed gas supplies.14

Difficult policy decisions and factual determinations must be made if the alternate fuel ceilings are to fulfill the function of pricing industrial gas at the highest level consistent with retention of the load. The commission must decide whether to approach the question solely on a geographic basis, or through end-use analysis, or by a combination of the two. For example, while the NGPA contemplates use of No. 2 fuel oil prices as the alternate fuel price ceiling, it also permits the commis-
sion to use lower priced No. 6 fuel oil if a user can prove that conversion would result, unless the commission determines that making its determina-
tion of the ceilings, the commission will need to consider factors such as size, location, mode of shipment, transportation costs, and storage capabilities. The alternative fuel cost ceilings will have to be updated frequently to ensure that gas will not be replaced if alternate fuel costs decrease.

In making these decisions, the commission will not lack critics. Oil jobbers seeking to compete with gas, industrial users, gas pipelines and distributors, and consumer groups, all will have a stake in the alternate fuel ceilings, the procedures by which they are set, and the factual determinations involved.

Passing Costs Through

Tracing costs of incremental purchases to incremental sales for industrial use is certain to be complicated. The statute specifies that pipeline sales, both direct and indirect, to boiler fuel users subject to incremental pricing will be required to absorb the incremental costs up to the level of alternative fuels. Major difficulties may be encountered in associating costs with sales. While there appears to be no particular problem in determining the portion of the sales to be incrementally priced when the pipeline makes a direct sale, the tracing problem becomes much more complex when pipeline sales are made indirectly, after two or three wholesale transactions involving other pipelines and distributors. In many instances, the retail distributor has more than one pipeline supplier and has its own supplemental supply sources, such as local production, propane, or even LNG; underground storage and LNG storage are also common. Identification of sales made at the retail level with supplies from a particular pipeline source will inevitably require complex and controversial allocation mechanisms.

In addition to allocation or tracing problems, timing problems arise in tracking incremental costs. The statute requires pipelines to record their separate account the portion of the cost of gas supplies subject to incremental pricing. This account would probably function in the same manner as a purchased gas adjustment deferred account; the pipeline would accumulate incremental costs, along with carrying charges, which it would periodically pass on to its customers by a surcharge. By the time a surcharge is reflected in distribution company rates, however, the incremental sales volumes may have changed significantly, due to the total or partial loss of incrementally priced sales, addition of new industrial sales, or changes in industrial activity. With the consumption by individual users subject to significant change, it may not be equitable to attribute incremental fuel costs incurred in a prior period to the sales made six months to a year later. Nor is it likely to be feasible to bill industrial customers for incremental costs associated with prior deliveries without piercing the alternate fuel ceilings for that period of time.

The Impact of Incremental Pricing

In response to requests by Senator Jackson and Congressman Dingell, the Energy Information Administration (EIA) made a two-stage analysis, which attempted to quantify the effect of incremental pricing of natural gas upon the new price provisions in the proposed Senate and House legislation, as well as the conference bill. The EIA studied the effects of various price provisions on the residential, commercial, raw material, and industrial sectors of the market on two bases: (1) assuming that incremental pricing would stop at the pipeline level, and (2) assuming that incremental pricing would be carried through to the retail level. It was expected that the pricing mechanism would be a combination of the two approaches, and the impact would fall somewhere between the two extremes outlined in the studies. While those studies showed that incremental pricing would result in rates to residential and commercial consumers lower than under rolled-in pricing, the basis for the EIA studies is open to question. In the mass of material that was presented to the Congress and that has been presented to the commission over the years, reliable and widely accepted data on the elasticity of natural gas supply have not been available.

It is generally agreed that higher prices will mean more gas, but no well-established basis exists for determining how much more gas will result from any particular level of higher prices. There is an equivalent lack of widely accepted data concerning the elasticity of gas demand, analysis of which is greatly complicated by curtailments in the past several years. In addition, incremental pricing will be imposed in the context of existing curtailment tariffs, and some industrial markets subject to incremental pricing will not be supplied with natural gas as a result of curtailment policies.

It remains unclear whether the incremental pricing provisions will provide the leeway Congress envisions, for example, whether higher prices will encourage greater supply and yet incremental price provisions will prevent loss of the industrial sales necessary to shelter resi-
dential and commercial users from such price increases. Congress has given the FERC considerable flexibility in setting alternate fuel ceilings and also has required the commission to develop an expanded rule that
would reach other industrial users and soften the cost impact on residential and commercial users. While there is a substantial question concerning the size of the boiler fuel market now available for application under the first rule, since electric utility boiler use is exempt, there can be no question that the remaining industrial market is large. According to American Gas Association data, about 57 percent of all gas sold in the United States in 1977 was for industrial users exclusive of electric generation.26

The expanded application of incremental pricing will undoubtedly be even more difficult and controversial than the initial application to boiler fuel sales. One of the provisions of the act exempts "small" industrial boiler fuel use of less than 300 Mcf per day on an interim basis, but provides authority to the commission to lower the 300 Mcf threshold to ensure that exempt "small" industrial boiler fuel use does not exceed 5 percent of the U.S. interstate boiler fuel use of natural gas.27 The 5 percent determination is to be based upon the aggregate amount of natural gas used in 1977 for boiler fuel and transported through facilities of interstate pipelines.28

Exemption from incremental pricing of certain uses, for example, agricultural purposes and electric generation by utilities, makes the impact of incremental pricing on the gas industry unpredictable and may create serious conflicts within the industry. A pipeline may find it difficult to choose either to acquire incrementally costed supplies or to continue curtailment of some loads in low curtailment priorities that are exempt from the passing through of incremental costs.

Another potential source of controversy is the uneven need of distributors for incrementally priced supplies. For distributors unable to market all the gas the pipeline is offering, the acquisition of incrementally priced new supplies at higher costs could further reduce their market. When a pipeline serves such distributors, as well as others who do not have all the gas they need, a conflict is bound to develop if the pipeline seeks to add new supplies at a significant incremental cost. This problem has already arisen in some curtailment cases, when the pipeline has sought to purchase emergency supplies on an incremental basis, and some customers have refused the additional supplies as too costly. Nor is the pipeline likely to receive advance approval of the incremental costs of new gas purchases.29

Conclusion

At this stage, any attempt to draw specific conclusions about the incremental pricing provisions of the NGPA is premature. The key issue — whether there should be incremental pricing — has been decided by Congress. The provisions do not follow pure marginal pricing theories, and their objectives seem to be in conflict with other parts of the NGPA. The effectiveness of the first incremental pricing rule is seriously in question because no solid data are available as to the size of the market that may be affected by incremental pricing or the extent to which price increases under the NGPA can be absorbed by this market within the ceilings imposed by alternate fuels.

There are, however, areas in which reaction to the act can be predicted in advance of any regulations. It seems clear that the conferees contemplated that direct and indirect customers subject to incremental pricing will be treated in a similar fashion. This means that the incremental costs passed through to a pipeline's customer, direct or indirect, will be the difference between the alternate fuel ceilings fixed by the commission and the prices charged to that customer. Because the allocation of incremental costs among a pipeline's customers will be a function of the rates those customers are charged incrementally priced facilities, it would seem to follow that industrial boiler fuel rates will immediately increase to the level of the alternate fuel ceilings.

Another important area will be the segregation and classification of loads to establish what volumes should be incrementally priced. In this respect, not only the commission regulations, but also the EIA data collection, could have important substantive effects and should be carefully analyzed.

The accumulation of hard data on alternate fuel prices is crucial. Most pipelines and distributors involved in marketing industrial gas do not need to be told what the competition is, but proving it may be another matter. There are also questions as to whether such ceilings should be based on No. 2 or No. 6 fuel oil, and whether they should be determined by class of customer or on the basis of an overall regional approach. It is not too early to define loads that will be exempt and to develop the case for creating and supporting additional exemptions, as well as contesting those that will be sought by others.

It has been my good fortune to survive the area rate proceedings, which lasted roughly a decade, and then live through the curtailment cases, which some would agree were the progeny of the area rate cases. The latter were complex and involved massive records, several carloads of data, and no end of controversy over technical issues of cost allocation involving the exploration, development, and production of natural gas. The curtailment cases were almost as global as the area rate cases because, for some of the larger pipelines, the impact was felt by the gas distribution industry in almost the entire eastern half of the United States. The curtailment cases likewise were exceedingly com-
Incremental Pricing

tplex, not only because of the unsettled state of the law, but also because of the huge volumes of data that needed to be processed. By and large, the curtailment cases are on the wane, although some continue to be litigated in the commissions and the courts, and the damage cases may have just begun. Now comes the incremental pricing phase. While it is too early to tell, it promises to be an even grander arena for the legal and consulting businesses that have flourished along with the natural gas industry.

Notes
3. Ibid., p. 46.
4. Ibid.
7. Columbia LNG Corp. v. FPC, 491 F.2d 651 (5th Cir. 1974).
10. The Senate bill, H.R. 9463, passed on 5 February 1976, prohibited new sales of gas for boiler fuel use after 1 January 1976 (unless alternative fuels were unavailable or their use was impractical), and phased out over a ten-year period old sales of gas for boiler fuel use. The Senate bill passed on 28 October 1975, S. 2310, established a priority for sales of old gas to distribution companies for service to residential and small users, thus shifting the burden of higher new gas prices to lower priority users. The 1976 Senate bill, S. 3422, reported out of the Senate Commerce Committee on 19 May 1976, eliminated new gas sales for boiler fuel use after 12 May 1976 and phased out over a ten-year period old sales of gas for boiler fuel use for the purpose of generating electricity. That bill imposed incremental pricing on boiler fuel users, except that no boiler fuel user would be required to pay in excess of 120 percent of the average price of heating oil.
11. The letters of Charles J. Cicchetti, Chairman, Wisconsin Public Service Commission, to the teleconferences to Congressman Henry Reuss illustrate the criticism leveled at the incremental pricing provisions.
12. The threshold price is $1.48 per million BTUs for March 1978, adjusted for inflation in subsequent months.
15. For pipeline curtailment cases, a pipeline might show curtailment of all loads in Priorities 5 through 9, whereas distributors served by that pipeline might be making sales in all priorities.
17. Given the DOE's policy to capture as much as possible of the market now using foreign oil, it appears that the commission will take advantage of the statutory flexibility in setting the alternate fuel ceilings. In a recent press conference concerning the promulgation of coal conversion regulations, Secretary James Schlesinger stated that the department strongly supports use of gas rather than foreign oil in existing facilities.
19. The FEA studies used the Project Independence Evaluation System (PIES), which may not accurately assess the current condition of the gas industry.
21. Section 206(a)
23. For example, see Texas Eastern Transmission Corp., Docket No. RP78-32, Order Denying Request for Advance Approval (27 January 1978).
Financing Large-Scale Gas Supply Projects

John F. Curley, Jr.

The continuation of the level of economic growth needed in this country to provide jobs in the 1980s and beyond depends on energy, particularly gas supply. Many of the projects to supply this gas will provide expensive energy and will be costly and risky. U.S. gas companies do not have the financial strength and resources of firms in the industrial sector who have recently or are about to face similar substantial capital commitments. Examples of such would be IBM, Xerox, General Motors, and the partial utility, AT&T. Current financial conditions and the capital requirements of the U.S. gas industry will require new regulatory responses to the problems of risk and return. An obvious, logical, and reasonable solution to these problems is the increased use of the all-events, full-cost-of-service tariff.

The special problems of large-scale gas projects all relate to risk. In particular, what are the risks involved in a given project? Who should bear these risks? How should the risk-bearers be compensated?

The risks of a project generally are of three types: economic, credit, and security. Lenders and sponsors are concerned with all of these, and regulators should be.

In considering the economic risk of a project, lenders and sponsors will consider three aspects. The first is technological. There is a risk that the technology on the scale proposed for the project will not perform as expected or will become prematurely obsolete. The second concerns reserves. The risk is that the gas (for transport or liquefaction) or the coal (for gasification) needed for successful operation will become depleted or unavailable during the life of the project. The third involves demand. The cost of the gas may be so high that there will be insufficient demand to provide the needed return on the project. Regulatory and environmental factors most directly affect this risk aspect.

In the credit and security areas, the principal requirements are generated by the lenders providing the debt money for the project, which might typically be 75 percent of the total cost. Lenders to large projects are unwilling to take any equity-type risks and require that they be given security against three kinds of risks by credit-worthy parties who are able to bear them. The first is completion. They wish to be protected against the risk that the project may not be completed for any reason, including the lack of funds to meet cost overruns. The second is debt service. Even though the project is completed and operating satisfactorily, it may not generate enough money to pay interest on and principal of the debt. The third is force majeure, the risk that after completion there may occur a curtailment or interruption of operations for a long period.

Protection against these risks is usually provided by the companies sponsoring the project through various contractual commitments. The sponsors also put up the equity funds (typically the remaining 25 percent of the project costs) to further protect the debtholders.

For today's large projects, the problem of risk-bearing is a double-edged sword. Lenders are concerned that the sponsoring credits are not strong enough to provide meaningful support, and sponsors are finding it impossible to justify economically the direct and indirect guarantees required by the lenders.

Current large-scale gas supply projects include all of these risks and to a degree greater than ever before. Projects differ considerably from the promotional pipelines built in the 1950s which connected large and proven fields of very inexpensive gas to markets where demand was tremendous. In addition, two or three decades ago the technology of pipeline and gas transmission was relatively simple and established, and it posed no new major technological problems. Construction periods were short and certain, and costs for materials were predictable and not subject to high rates of inflation. The economics of the project were so obvious that the gas provided the credit, and security arrange-
ments with sponsors were not as ironclad as would be necessary today.

Because large-scale gas supply projects no longer provide nearly as much support against economic risks, it is crucial that good credits provide the security. Furthermore, the larger the project, the more likely it is that a full-cost-of-service tariff would be needed as security.

This tariff would be one of a series of security arrangements which usually are included in the typical financing technique for large undertakings: project financing. This technique permits flexibility in a locating risks among the parties providing support. Discussion of project financing has been extensive in recent years; suffice it to repeat that it is a method of providing credit support to a financing — it is not a means of financing without credit support.

Allocation of risk among a number of supporters is necessary for large-scale gas projects because the risks are such that no one entity has been willing to bear them. The major reason for this unwillingness is that the size of the projects requires commitments that are not compatible with the size and capital structure of the sponsoring companies. The risk is often compounded by the presence of some (although small) technological risk.

An extreme example is the Alcan pipeline proposal. The estimated cost of $15 billion is about three times the combined equity of the identified sponsors. In essence, each sponsor would be committing all its assets to be able to transport gas reserves constituting less than 7 percent of annual U.S. consumption. This would be done against the backdrop of what has happened in regulation of the Trans-Alaskan (TAPS) oil pipeline. Although the TAPS sponsors were Aa and Aaa rated oil companies, the pipeline had the additional risk that it does not have a certificate of convenience and necessity, which would provide, in effect, a monopoly franchise. A bitter fight continues regarding the TAPS tariff, with a particularly dark note being the argument by some that the completion risk borne by the TAPS sponsors can now be ignored, since the project has been completed.

Another example of an imbalance in the risk and reward of a large-scale gas project is the proposed $1 billion Great Plains coal gasification plant. Although the amount being risked by the sponsors is only 17.4 percent of their combined equity, they are each risking $200 million for the right to 125 mcf per day of $7.25 gas. The cost is justified, in the long run, if the research and development of this technology on such a scale proves useful to the American people.

It is clear that the gas companies are not considering assuming these risks for the benefit of their shareholders; the expected returns are not in balance with the risks. The projects are being undertaken for the consumer. It is not surprising, therefore, that the rate payer is being asked to assume some of the risk.

This assumption of risk may take one of two forms. First, as a rate payer, the consumer may assume risk through tracking agreements whereby the distribution companies collect money from their customers to pass on to the project to repay the obligations of the project. Second, as a taxpayer, through U.S. government guarantees which might have to be called upon at some future date, the consumer may assume risk. This use of taxpayer money can be justified as being in the common interest. Additional gas reduces the need for imported oil, thereby enhancing national security and reducing the balance-of-payments deficit. Furthermore, the supply of the most environmentally acceptable fuel is increased.

The involvement of the federal government, distasteful as it may be to many, will be required if decisive and positive action cannot be taken to gain the support of the consumer as a rate payer. This federal involvement, which could be through a program of completion guarantees, debt guarantees, or insurance, would only be used for new gas supply projects which are deemed too large or too risky for the private sector.

Before concluding that the federal government is the only answer, however, it would be worthwhile to examine the other alternative — risk sharing by the consumer as a rate payer. This course seems more equitable and avoids a possible radical shift in the place of private-owned utilities in the capitalist system. The acceptance by state regulators of the all-events, full-cost-of-service tariff contract would be a major step in achieving risk sharing by the rate payer. This is a legally binding contract which will assure investors, at the time their funds are committed to the project, that all operating costs and capital costs will be punctually recovered under all circumstances through rates to gas consumers. “All circumstances” include noncompletion of the project and any force majeure interruption which would prevent the delivery of gas — both of which risks, as I mentioned earlier, are unacceptable to lenders.

In stressing the need for acceptance by state regulators of the all-events, cost-of-service tariff, I did not mean to imply that this tariff has enjoyed complete acceptability at the federal level. I emphasize state regulatory acceptance because the distribution companies must be able to collect from their rate payers the charges of the pipeline companies before those companies can be sure they have a chance of collecting their tariffs. This interaction of federal and state jurisdiction has caused some to express concern that the federal government will be usurping the rightful jurisdiction of the states. Certainly, a gas distribu-
Financing Gas Supply Projects

John F. Curley, Jr.

The operation company sponsoring a very large project must seek approval from all the states which it serves — to avoid the unhappy possibility of consumers in one state subsidizing those in another. It seems possible that at least part of this concern over federal interference is a smoke screen created by states which want to avoid risk. These same states may also be counting on federal allocations if their failure to support large projects results in a gas shortage. We can only hope that such allocations, if they occur, will be priced to consumers in the nonrisk states at a rate that fully compensates the consumers who help bring the gas to market.

Acceptance by state regulators of an all-events tariff is not an abdication of public responsibility if it results in completion of a larger number of gas supply projects than would otherwise be possible. Development of an adequate gas supply is clearly more in the public interest than is maintenance of artificially low prices, which would worsen the gas situation and ultimately lead to curtailment of residential customers and unemployment.

We expect it will be very difficult for regulators to ask consumers to pay for a project even if no gas is being delivered, however unlikely that may be. It is critical that they do so.

A related issue is whether the cost of this gas (or risk sharing) should be borne by one group of customers (incremental pricing) or by all (rolled-in pricing). Incremental pricing is not the answer; the cost of project gas would be so high that other forms of energy would be substituted. The lack of demand from new and industrial customers would bring the new gas back to those who switched to substitutes. Residential customers have been paying less than their fair share for gas and have been subsidized by commercial and industrial users for years, so it seems fair that these same customers now assume their share of the risks of new gas supply projects.

If the rate payer were to assume some of the risk, the critical remaining problems are understanding the risks taken by the sponsors and the proper return. The principal benefit (return) for the rate payer is more gas at a rolled-in price lower than that for other forms of energy. The risks to be assumed are those from which lenders must be protected: completion, debt service, and force majeure.

To digress briefly about lenders, the main source of money for large project financing has been the life insurance industry. Full participation of this industry in project financing is related to the legal investment test for New York insurance companies. Under New York law, investments in new entities, such as those used as vehicles for project financing, which cannot meet historical interest coverage tests are legal if they are "adequately secured." This term usually is taken to mean that the three risks previously described have been assumed by credit-worthy parties other than the lenders. As was mentioned, no sponsoring company could risk the amounts of funds required for very large projects, but even if it were willing to do so, insurance companies would not deem the debt to be "adequately secured." The failure to attract funds from major insurance companies would make large project financing extremely difficult, if not impossible.

The commercial banking industry is also an aggressive and imaginative source of money for large projects. Bank money, however, is often available at a floating rate and for shorter terms. This adds economic risk to the project, since capital costs are less certain and must be recovered over a shorter period. Bank financing contracts may be less ironclad than those required by insurance companies, but it should be noted that the exact limits of this flexibility are still being tested in the financing documents for such current projects as the Dakota coal gasification plant and the Alaskan gas pipeline.

The all-events, full-cost-of-service tariff at the wholesale level, with similar tracking agreements at the retail level, would usually satisfy lenders' reservations about completion, debt service, and force majeure risks. This tariff would normally be in the form of a demand charge, for which the rate payer receives the right to the gas if it is delivered. If no gas were delivered, the charge would still be paid.

Less than this full support might allow some smaller projects to be financed, but the new gas back to those who switched to substitutes. Residential customers have been paying less than their fair share for gas and have been subsidized by commercial and industrial users for years, so it seems fair that these same customers now assume their share of the risks of new gas supply projects.

If the rate payer were to assume some of the risk, the critical remaining problems are understanding the risks taken by the sponsors and the proper return. The principal benefit (return) for the rate payer is more gas at a rolled-in price lower than that for other forms of energy. The risks to be assumed are those from which lenders must be protected: completion, debt service, and force majeure.

To digress briefly about lenders, the main source of money for large project financing has been the life insurance industry. Full participation of this industry in project financing is related to the legal investment test for New York insurance companies. Under New York law, investments in new entities, such as those used as vehicles for project financing, which cannot meet historical interest coverage tests are legal if they are "adequately secured." This term usually is taken to mean that the three risks previously described have been assumed by credit-worthy parties other than the lenders. As was mentioned, no sponsoring company could risk the amounts of funds required for very large projects, but even if it were willing to do so, insurance companies would not deem the debt to be "adequately secured." The failure to attract funds from major insurance companies would make large project financing extremely difficult, if not impossible.

The commercial banking industry is also an aggressive and imaginative source of money for large projects. Bank money, however, is often available at a floating rate and for shorter terms. This adds economic risk to the project, since capital costs are less certain and must be recovered over a shorter period. Bank financing contracts may be less ironclad than those required by insurance companies, but it should be noted that the exact limits of this flexibility are still being tested in the financing documents for such current projects as the Dakota coal gasification plant and the Alaskan gas pipeline.

The all-events, full-cost-of-service tariff at the wholesale level, with similar tracking agreements at the retail level, would usually satisfy lenders' reservations about completion, debt service, and force majeure risks. This tariff would normally be in the form of a demand charge, for which the rate payer receives the right to the gas if it is delivered. If no gas were delivered, the charge would still be paid.

Less than this full support might allow some smaller projects to be financed, but the new gas back to those who switched to substitutes. Residential customers have been paying less than their fair share for gas and have been subsidized by commercial and industrial users for years, so it seems fair that these same customers now assume their share of the risks of new gas supply projects.

If the rate payer were to assume some of the risk, the critical remaining problems are understanding the risks taken by the sponsors and the proper return. The principal benefit (return) for the rate payer is more gas at a rolled-in price lower than that for other forms of energy. The risks to be assumed are those from which lenders must be protected: completion, debt service, and force majeure.

To digress briefly about lenders, the main source of money for large project financing has been the life insurance industry. Full participation of this industry in project financing is related to the legal investment test for New York insurance companies. Under New York law, investments in new entities, such as those used as vehicles for project financing, which cannot meet historical interest coverage tests are legal if they are "adequately secured." This term usually is taken to mean that the three risks previously described have been assumed by credit-worthy parties other than the lenders. As was mentioned, no sponsoring company could risk the amounts of funds required for very large projects, but even if it were willing to do so, insurance companies would not deem the debt to be "adequately secured." The failure to attract funds from major insurance companies would make large project financing extremely difficult, if not impossible.
situations which by themselves have increased the risk to the equity holder. This perversity is justified as a control measure; the FERC needs to give the sponsors of the project an incentive to complete and operate the project at the lowest possible cost.

We would argue that this type of regulation does not follow the practices of unregulated industry. Who pays for the failure of projects or products in the consumer goods market? Who paid for the Edsel? Who is paying for the recall of millions of cars and trucks? The answer is the consumer. He pays in the high profit margins of the successful products, which pricing practice is not available to regulated industries. The regulated company not only is required to absorb more losses, but also is precluded from offsetting losses by higher profits on other projects.

Nonetheless, project sponsors will be able to live with this control if (1) base cost estimates are reasonable and include inflation adjustments to allow for regulatory delay, which is not within the control of management; (2) the outages which may cause a lower return are limited and within the control of management; and (3) the base equity return reflects the risks of the project, which return probably will be at least 15 percent.

The particular risk sharing arrangement will determine the returns that are needed by the sponsors to offset their risks. The minimum overall return for a large energy project, however, should be substantially higher than the 10 percent which is frequently allowed to conventional pipelines. In recent testimony before the FERC, we noted that a project such as the TAPS pipeline should be viewed as an all-equity investment. The appropriate return should consider, among other things: (1) the colossal size of the project (which was smaller, however, than the planned Alaskan gas pipeline), (2) the undiversifiable risk of a "stand-alone" project which committed funds to one source of reserves and one pipeline, and (3) the economics of the project as a means of assessing the demand for the product at its expected cost.

The major misunderstandings about financing large gas supply projects could be cleared up if such projects were considered as they should be—as all-equity projects. Regulators could then start from an appropriate return level and consider how to reduce that required return by allocating risks to the consumer. The addition of debt to a project would not lower the overall required return but would increase the equity return to offset at least its subordination. The risk of equity participation in a project would be increased by the addition of debt which has a prior claim to the revenues from the project.

In summary, large gas supply projects will be financeable once regulators, particularly at the state level, appreciate the risks inherent in these projects and the returns required to make the risks bearable. Regulators will then have to allocate the risks and returns in a fair manner to all parties. Large projects will require a greater commitment by consumers prior to their completion. Perhaps regulators do not see the benefits of these supplemental gas supply projects as being sufficient to offset the political problems of changing the timing of consumer commitments to pay for gas. This imbalance will probably not be altered in the near future if the new Natural Gas Policy Act results in increased supplies of gas from conventional sources. We must begin, however, to evaluate the risks of large projects so that, when the time comes to finance such projects to provide supplemental supplies needed for economic growth, we will have the ability and the will to allocate and compensate these risks.
Public Policy Considerations in the Pricing of New Gas Supplies

David S. Schwartz

determining whether or not the supplemental project in fact does serve the broader public interest. Policy makers must examine the broader public need, not on the basis of the project-by-project justification, but within a broader context in order to determine whether the project assures reliability of supply, is the lowest cost alternative, and conforms to overall energy policy. Policy makers must examine the basis for resource optimization in choosing one project over another. In contrast to the firm’s efforts to find a means of justifying a project by risk minimization, the regulatory body should find the alternative which will pressure the firm to the greatest efficiency in operation, and the lowest price alternatives in certifying any new supplemental project. The framework for public policy decision making is clearly distinct from that of the firm, and regulators must draw this distinction very clearly in developing public interest standards.

A Fundamental Approach for Public Policy

In order to determine the place of supplemental gas supplies in the overall energy balance, there must be an examination of total energy requirements, and then a determination of the role of gas with respect to them. This requires a sensitivity to overall estimates of energy demand as influenced by the general rate of inflation, the impact of inflation on energy prices, and the effect on gross national product as influenced by population growth, higher energy prices, and environmental standards. In order to ascertain the need for supplemental gas supplies, one first must determine the overall energy requirements which are dependent upon some of these macrofactors which are basic to the determination for energy needs in general.

The consumption of all forms of energy in 1978 amounted to 77.7 quads and reflected a 1.8 percent increase over 1977. The 1978 consumption growth rate was down from the 2.5 percent growth rate in 1977 and 5.3 percent growth rate in 1976. The ability to predict the future growth rate in energy consumption given this very marked shift between 1976 and 1978 presents very serious problems. These are reflected in the very wide range of energy growth rates, varying from 3–4 percent to 1.5–2 percent. The estimates by industry representatives project a growth rate in the near term, that is, 1978 to 1985, of 2.8 percent, and in the longer term, 1985 to 1990, of 2.7 percent. If the growth rate is substantially below these projections, for example, one-half of what the industry projects, then the overall energy requirements will be considerably less, and the need for supplemental natural gas supplies will be considerably lower. The gas industry estimates that by the year 2000 we will require 130 quads in contrast to the 77.7 quads
consumed in 1978, and that the role of gas supply will increase from approximately 20 quads in 1978 to 32 to 35 quads in the year 2000.\textsuperscript{3}

The factors that must be closely examined by policy makers are these:

1. the need for natural gas in general and the specific need for supplemental supplies other than wellhead gas; in other words, natural gas must be seen in terms of overall energy requirements;

2. the need for supplemental supplies in relation to overall requirements depends upon the availability of conventional wellhead supplies of natural gas; and

3. a national approach requires that the need for conventional supplies as well as supplemental supplies be viewed as an integrated approach which does not examine each project on a piecemeal basis.

Examining the Role of Supplemental Gas Supplies

The volume and type of supplemental supplies required in the future depend upon the level of production from conventional sources of natural gas. The future availability of natural gas depends upon the already found proved reserves as well as the resource base from which proved reserves can be developed. Another significant consideration is the relative cost of alternative supplies of fuel which can substitute for natural gas. Some additional factors that will influence the need for supplemental gas supplies are the success of exploratory drilling on the North Slope of Alaska; the impact of higher energy prices upon conservation of public policies affecting the past and use of natural gas; the relative cost and availability of alternative forms of energy, such as solar, electric power, and petroleum products; and the export policies of Canada and Mexico.

The difficult task faced by policy makers is best illustrated in examining the estimates of the American Gas Association in contrast to the Department of Energy as to total gas requirements and estimates of supplemental natural gas requirements in the future. In the near future, AGA estimates that the demand in 1985 for natural gas will be 24.2 trillion cubic feet (TCF).\textsuperscript{4} The Department of Energy (DOE) estimates that in 1985 the overall requirements for natural gas will total 19.6 TCF. The difference of 4.6 TCF is approximately 23 percent. Before a rational decision can be made on the need for supplemental gas supply projects, the significant differences between the two estimates must be reconciled. Tied to these overall estimates is the need for supplemental gas supplies in 1985 of 5.2 TCF forecast by AGA, in contrast to the DOE’s estimate of 2.5 TCF. The difference is more than 100 percent and poses some fundamental problems for policy makers before the proposals that are filed with the Federal Energy Regulatory Commission (FERC) can be assessed.

In the longer term, AGA estimates the need for 28.6 TCF in 1990; in contrast, the DOE estimates 20.3 TCF. The AGA estimate is approximately 29 percent greater. Supplemental supplies of 8.5 TCF are included in the overall gas requirements by AGA for 1990, while DOE estimates the need for 5.3 TCF. This is more than a 250 percent difference.\textsuperscript{5}

The factors underlying the AGA estimate will have to be examined closely by public policy makers. For example, AGA gives no effect to price elasticity on the demand for natural gas generally, or with respect to the need for supplemental supplies. AGA assumes that all gas produced will be sold, and that the only constraint is on the supply, not the demand, side. AGA sees its competition with electric power and therefore assumes the price of electricity will determine the ceiling affecting the demand for gas.

The limitation of this view is illustrated by the fact that some uses for electric power, such as lighting and home appliances, do not permit substitutability. In addition, gas cannot be made available in rural areas because the necessary pipeline facilities do not extend to these remote regions. Finally, gas competes with other fuels in addition to electric power, for example, home heating oil, and gas rates must be examined in terms of substitutability in the use of other fuels, which must be considered as potential constraints on the demand for gas. AGA takes no account of the impact of higher prices on conservation efforts, which will reduce the demand for gas in the future. These efforts will include better insulation of individual residences and more efficient use of gas in commercial and industrial operations.

Supplemental Supplies: Public Policy Considerations

It should be emphasized that policy makers must view the need for natural gas and supplemental supplies of gas in a total energy context and within an appropriate time frame. In addition, supplemental supplies fill the void of unavailable conventional wellhead gas. There is an important difference in using supplemental supplies for peaking purposes as opposed to base load, and regulators must view the need
Pricing New Gas Supplies

for supplemental supplies differently in each instance. Finally, the primacy of conventional wellhead gas must be accepted before considering how much supplemental gas and what type are required.

Recent estimates of natural gas reserves indicate a formidable reserve base from which future gas of a conventional nature is obtainable. The Potential Gas Committee recently issued a report which included a number of estimates from various industry sources as well as the United States Geological Survey. It indicated that the resource base ranged from 500 to 700 TCF of remaining recoverable reserves. These estimates were for the most likely value, or 50 percent probability for undiscovered gas reserves in the probable and possible category. If the 200 to 700 TCF of proved reserves are added to the 500 to 700 TCF of probable and possible reserves, the range of existing resources is between 700 and 900 TCF. By any conservative objective level of consumption, this could provide conventional natural gas supplies for 35 to 45 years. If one includes reserves from the so-called speculative category, this would increase the potential span to 50 or 68 years. Given the availability of natural gas from domestic sources, policy makers should be very cautious in certifying high cost supplemental supplies and should concentrate on the policy formulation required to develop the existing resource base.

While the industry stressed the supply eliciting function of higher prices, very little attention has been paid to the downside pressure on demand. In addition, as gas prices continue to increase, other substitutable fuels become competitive alternatives. With the passage of the Natural Gas Policy Act in November 1976, the nature and scope of regulation was substantially changed. In contrast to the former utility type regulation, we now have pricing of natural gas premised upon market criteria and an upward thrust which will have a very marked impact on demand elasticity for natural gas. In comparison to the $1.52 per Mcf price for new gas that obtained at the passage of the Natural Gas Policy Act, the rates set for December 1978 under the new legislation were over $2.07 per Mcf. The price of unregulated intrastate gas was in the range of $1.82 to $1.86 at the time the legislation was passed. The result was administered prices far above the market price that would have existed if the new legislation had not been enacted. With prices at current levels, it is only reasonable to recognize the downside pressure on the demand for gas and the potential shift to cheaper alternative fuels. Regulators will have to determine whether these high prices will sharply contract the demand for natural gas; they certainly will lessen the need for higher cost supplemental supplies of natural gas.

Intramarginal Supplies

In view of the various applications for imported liquid natural gas (LNG) from Algeria, the Economic Regulatory Administration (ERA) of the Department of Energy has developed a new ranking when considering such matters as the security of supply, the national and regional needs for gas, the proposed import price, and the effects of the imports on the balance of payments. The ERA has placed the highest priority on conventional wellhead gas supplies in the lower 48 states. It has categorized intramarginal supplies, such as Alaskan gas, as the second most important, and the marginal or so-called LNG imports as the least preferable.

Alaskan Gas

A current proceeding before FERC involves the certification of gas from the Cook Inlet in Alaska. The Pacific Alaska Project (Pac Alaska) is sponsored by Pacific Gas and Electric Company and Pacific Lighting Company. It was initially filed in November 1974. The wellhead price proposed was $49 per Mcf, and transportation and regasification in Southern California was projected at $3.20 in the first year, decreasing to $2.36 per Mcf when full delivery commenced.

The proceedings were stalled because of procedural problems, and the producers refilled the application in 1977. The renegotiated wellhead price was increased to $1.46 per Mcf, and the delivered price to Southern California increased to $3.51 per Mcf. Original projected deliveries of 200,000 Mcf a day, developing to 400,000 Mcf a day when full deliveries commenced, were reduced to 130,000 Mcf per day.

One of the controversial aspects of the Pac Alaska proposal concerns the location of the terminal facilities in Southern California. The terminal is also to be used for imports of LNG from Indonesia. The sponsors of the proposal had planned to locate the terminal at Oxnard, California, but because of the density of population (a statutory prohibition by the California legislation), the project was proposed for Point Conception, California. The FERC staff, the Hollister Ranch interests, and the Santa Barbara Citizens Association opposed the location at Point Conception. They proposed an alternative that would ship Cook Inlet gas currently going to Japan to Southern California via the Alaskan pipeline, and the Indonesian gas which would have been imported to Southern California would be sent to Japan in lieu of Cook Inlet Gas. The sponsors would not consider this exchange, fearing that it would kill the entire Indonesian as well as Pac Alaska Project. There is an additional interesting aspect concerning the proposed pricing of Cook Inlet gas: The applicants proposed rolled-in rather than incre-
mental pricing. This issue will have to be resolved by the California commission, because the gas is all intrastate.

Prudhoe Bay — North Slope Gas

In the initial applications by Arctic Gas, Alcan Gas Pipeline (now Northwest Alaska Pipeline Company), and the El Paso Proposal, the administrative law judge in the initial decision in February 1977 raised some very significant but unsettled questions. He pointed out that the major producers — Exxon, Arco, and Sohio — had not entered into any sales contract, delivery volumes and prices had not been agreed upon, and these major impediments would normally result in the denial of the application. The Natural Gas Policy Act settled the price level by providing for $1.45 per million BTU's as of April 1977, with monthly inflation adjustments. The price anticipated by DOE in 1983 for Alaskan gas is $3.28 per million BTU in 1978 dollars at the wellhead. The delivered price for Alaskan gas is estimated to be above $7.00 per million BTU.

The administrative law judge also indicated that there were no field production agreements providing for a delivery schedule, and that the question of financing was unresolved. Financing of the Alaskan pipeline has persisted ever since the Alcan project was backed by the President and approved by the Congress.

At a press conference in November 1978, representatives of Northwest Alaskan Pipeline (formerly the Alcan Pipeline) and Footlong Pipeline contended that rumors of the project's demise were "ill founded." The sponsors insisted that the project could be privately financed without government guarantees if the Internal Revenue Service would approve a billion-dollar tax exemption for the issuance of debt by the state of Alaska. They stressed the need for an incentive rate of return if private financing is to be obtained, given the magnitude of construction costs and the costs that are involved at prevailing interest rates. They pointed out that the initial cost estimate of $10 billion was now projected at $12 billion. In addition, they stressed the need for rolled-in pricing for the wellhead as well as transportation costs in order for the project to succeed.

FERC has conditionally approved the construction of the southern segment of the Alaskan pipeline. The "prebuilding" of the western leg is to serve California. The project provides for 240,000 Mcf a day of surplus in Alberta. This is to be purchased from Pan Alberta Gas Company over a six-year period, with an option to renew for another six years. The western leg will deliver gas to California in 1980, and the eastern leg will deliver gas to the Midwest in 1981. The availability of Alberta gas will be discussed subsequently, when the need for supplemental supplies is examined in greater detail. It should be noted that the California producers appealed to FERC not to certificate the southern segment of the Alaskan line on the basis that they had surplus gas available and could provide 240,000 Mcf a day currently from California gas supplies. The challenge brought by Midwestern Pipeline and Michigan Wisconsin Pipeline to FERC's authorization of the southern segment of the Alaskan line was upheld by the Circuit Court of Appeals of the District of Columbia.

In September 1978, FERC approved in principle an incentive rate of return for the Alaskan gas pipeline. The explanation was as follows: "The Incentive Rate of Return Mechanism is desirable because it is the traditional tool for cost control, disallowance of imprudent cost, in pipeline construction is a blunt instrument that counteracts only extreme cases." FERC adopted Order No. 17 (December 1, 1978), which provided for an incentive rate of return in the certificate that was issued for the Alaskan pipeline. They pointed out that the incentive rate of return concept was required in the President's decision, which selected an applicant and route for the pipeline. While this is true, there is still a great deal of latitude concerning the approach to and conditions that can be imposed on adopting an incentive rate of return. They could have adopted a policy that would have pressured the firm to perform better, in contrast to the more generous approach actually chosen. They could have tied the allowed rate of return to stringent performance criteria so that the two, the allowed return and the firm's performance, were correlated.

Instead, FERC started with the concept of an incentive rate and the nonincentive return which just covered the normal risk of the firm, plus the added risk of the special construction and completion of the Alaskan line. It then developed the idea of a center rate of return, which was an allowance above the general risk and special risks that were allowed in the so-called nonincentive rate of return. In addition, the commission developed the concept of a cost-performance ratio which correlated the relationship of actual and projected construction costs, and then related the return that would be allowed to the cost-performance ratio. For illustrative purposes the commission indicated the incentive rate of return schedule and provided the following percentages: nonincentive rate of return, 15 percent; center rate of return, 17 percent; and marginal rate of return, 9 percent. FERC is providing a premium over the normal return which would cover the risk of the enterprise plus the special risk of construction in Alaska. Supposedly, this premium is being provided in order to induce the
firm to control costs. The Office of Regulatory Analysis in FERC calculated the impact of the 2 percent premium above what would normally be allowed after giving effect to the investment tax credit and the leveraging involved because of the debt; it then determined an average rate of return on equity of 6.12 percent and a marginal rate on equity of 31.8 percent. This generous return to the project’s sponsors, because of the added cost of 2 percent when applied to a $12 billion project, is a questionable burden to be borne by consumers. While it is difficult to enforce the prudence criteria, the basic question still persists as to whether this, in essence, is the proper responsibility of the commission.

FERC further eroded the tight reign on cost control when it decided a number of issues in favor of the sponsors in response to a petition for reconsideration and clarification.⁷

Concerning the allowance for funds used during construction component in the cost-performance ratio, the commission made four changes which benefited the sponsors of the Alaskan project. They characterized these changes as follows: “The net effect of the adjustments to the IROR [incentive rate of return] mechanism is to greatly reduce penalty for delay.” In contrast to the generous approach promulgated by FERC, the other approach is to hold the firm to a prudence standard, irrespective of the difficulties involved. The possible use of a more stringent standard is reflected in a study of the Trans-Alaska Pipeline prepared by the General Accounting Office (GAO) in 1978.⁸

The report points out that in 1968 the feasibility study for transporting 1.2 million barrels of oil per day projected a cost of $1.46 billion, but construction costs were $7.94 billion when it was completed in 1977.⁹ The GAO report indicates that the cost control system was inadequate from the outset, and at every point thereafter. It stresses that in order to control cost overruns, the following steps are necessary: (1) Initial and subsequent estimates of the cost for the project should be viewed with skepticism; (2) there is a need for specific site inspection and site data as well as an investigation of the geologic and technical uncertainties of the project; (3) government approval should be contingent on detailed plans for management control, which includes budgetary control; and (4) the project should have an ongoing government audit to protect the public against overruns.

Perhaps FERC could utilize the lessons learned from the Alaska oil pipeline and implement regulatory procedures in overseeing the construction of the Alaska gas pipeline.

David S. Schwartz

SNG — Coal Gasification

In the DOE framework, coal gasification falls within the intramarginal supply category, just as does Alaskan gas. It is preferable to imported LNG, but not as high a priority as conventional wellhead gas. Coal gasification projects can be characterized as high capital cost, or capital-intensive, projects involving as yet not fully developed technology. The first step is to use the Lurgi process to produce low BTU gas. Problems still exist with respect to the methanation process, which converts low BTU gas to high BTU gas of pipeline quality.

The original coal gasification projects proposed by El Paso and Transwestern have long since disappeared from the scene. Despite the efforts to include a full cost-of-service tariff, minimum bill, rolled-in pricing, and loan guarantees, these projects evidently were not viable. The latest coal gasification proposal before FERC is the Great Plains project.¹⁰ It was filed in March 1975 by a subsidiary of the American Natural Gas Company, and it proposes to gasify coal from lignite in Mercer County, North Dakota, and calls for the production of 275,000 Mcf per day. Because of the capital costs, which amounted to approximately $1.5 billion, American National phased out the project and proposed building a plant to produce 137,500 Mcf per day initially. In addition, they acquired a partner — Peoples Gas Company (Chicago) — as co-sponsor. Finally, the project was restructured to provide for 125,000 Mcf per day, and three additional sponsors were added. These were the Columbus Gas System, Transcontinental Gas Pipeline, and Tennessee Gas Pipeline.

The major departure from traditional regulation and the various forms of subsidization far exceed the original proposals by El Paso and Transwestern. The application includes the following provisions: (1) an all-events cost of serving tariff; (2) a surcharge to cover interest on debt and the return on equity during the construction of the project (in effect, construction-work-in-progress in the rate base); (3) rolled-in pricing; and (4) debt guarantees for the full recovery of debt capital which amounts to 75 percent of the total capital cost.

The DOE intervened and supported the Great Plains project in general. Specifically, it agreed with the following provisions: (1) full recovery of debt in all circumstances, even in the event that the project is not completed and has to be abandoned; (2) a surcharge to cover interest on debt, but it took no position on return on equity; (3) the support of a cost-of-service tariff and adjustment in equity return for load factor; (4) the support of an automatic pass-through, that is, the use of a purchase gas adjustment for all pipeline charges; (5) the support of rolled-in pricing at all levels and to all categories of custom-
ers; and (6) although DOE took no position on rate of return relative to equity, it recommended that FERC consider the extent to which risks are precluded by the consumer, given the specific procedures that DOE supported.

The Ohio Consumer Council and the FERC staff took the position that the coal gasification project was not required at this time. The witness for the Ohio Consumer Council indicated that the resource base was adequate to provide gas supplies from conventional wellhead sources, and that current deliveries were increasing from Intrastate sources.

In September 1978, the FERC staff filed a motion with the administrative law judge to dismiss the application because the sponsors failed to provide the necessary financial information relative to the project. The motion stated that the sponsors "willfully and deliberately failed to produce evidence pursuant to the discovery." Subsequently, a newspaper article, which was confirmed by American Natural, reported that design and engineering work had stopped on the project. It went on to state that the failure by the applicants to resolve the controversy surrounding the financing of the project was the major cause for delay; that $10 million had been spent in design and engineering work, and that an additional $2 million a month was obligated through April 1, 1979. American Natural indicated that it was having difficulty raising money by traditional means because the coal gasification project depended upon an untried process. The sponsors went on to state they could afford to provide the 25 percent in equity capital that was required. The article further pointed out that the sponsors were looking for guarantees in the event that the project was not completed. One such guarantee was a possible surcharge on present consumer bills during the construction period.

Canadian and Mexican Gas

On the positive supply side, there appears to be a significant amount of surplus gas available in Alberta, Canada. The border price for Canadian gas is currently $2.16 per Mcf, which was escalated from a border price of $1.00 per Mcf in 1975. At that time, the Canadian Energy Board adopted a commodity value basis for pricing Canadian gas, and this appears to be the price which is competitive with other fuels in Northwest markets.

On June 30, 1978, the Alberta Conservation and Energy Board released a report indicating that reserves were adequate to permit a removal of 14 TCF of additional gas, after a reserve-to-production ratio of 25 years with respect to Canadian requirements. Alberta's productive capacity in 1977 was over 2.6 TCF, although actual production was 2.3 TCF. The estimate of Alberta reserves vary from a conservative 110 TCF to a more likely 130 TCF. Alberta currently supplies 300 to 400 billion cubic feet a year to the United States, and it is expected that this can be doubled in the near future.

An article in the trade press indicates that Canadian producers have submitted evidence to the National Energy Board of Canada showing an exportable surplus of one billion cubic feet a day, which is likely to continue growing. Canadian producers maintain that the current gas surplus ranges between 221-430 billion cubic feet a year. In supporting increased exports from Alberta, the briefs to the National Energy Board indicated the gas supply from established southern basin regions would continue to increase from about 3 TCF per year currently to 3.4 in 1985. They concluded that there are approximately 139 TCF of gas reserves within economic reach in western Canada.

On the southern end of the North American continent, six major pipelines recently negotiated to purchase gas from Mexico approximating 2 billion cubic feet a day at full delivery over a six-year period. Subsequent reports indicate that Mexico and the United States could not reach agreement as to price; therefore, these original negotiations have not reached fruition. Mexico wants to price the gas on the basis of No. 2 fuel oil in New York Harbor, and the Carter administration opposes the use of this benchmark.

The Executive Director of Pemex (the Mexican government oil company), in a speech before the American Petroleum Institute, indicated a very significant increase in oil reserves and associated gas reserves of 40 TCF of gas in place. In a recent speech, Secretary of Energy James R. Schlesinger indicated that the Mexican government and the Carter administration were moving closer to agreement on the price of natural gas. He suggested that a gas price tied to residual oil rather than No. 2 fuel oil may be the basis for a compromise.

Marginal Gas Supplies

**LNG Supplies**

The first importation of Algerian LNG was approved by the former Federal Power Commission in June 1972. The FPC authorized the importation by El Paso of one billion cubic feet a day for the Columbia System and Consolidated Gas Supply System at Cove Point and the Southern Natural System at Savannah, Georgia. The first deliveries occurred on March 1, 1978, and full production and delivery took
Pricing New Gas Supplies

place in the first quarter of 1979. The price of Algerian LNG increased from the original certificated price of 77¢ per million BTU to $1.25 at Cove Point. A rate increase has been filed with FERC to increase the price to $1.46. With respect to Southern Natural, the price increased from the original 85¢ per million BTU to $1.51. Currently, rates have been filed to increase the price to $1.56. These increases came about because of the provisions approved in the application, which allow a 12 percent inflation rate for base expenses and a 10 percent contingency provision and increases in transportation costs.

The second LNG project approved by the former Federal Power Commission was the Trunkline Gas Co. LNG proposal.14 The FPC permitted a price of $3.37 per Mcf and insisted upon the sale of the gas on an incremental pricing basis. Unless there was 100 percent delivery of the contract volumes, the commission prohibited a return to equity investors or any return of the capital invested, although the commission did approve a modified minimum bill. Ultimately, the commission reversed itself and permitted rolled-in pricing instead of the original incremental pricing of this gas.15

Under the National Energy Act, the final approval of LNG projects is under the Economic Regulatory Administration (ERA) of the DOE. Recently, the ERA denied the application by Tennesco Atlantic Pipeline Company (TAPECO) to import LNG from Algeria. The TAPECO proposal would have provided approximately one billion cubic feet a day for over 20 years, and the capital cost of the project was estimated at $5 billion. TAPECO proposed to build terminal facilities in St. Johns, New Brunswick, and then transport the gas over the New York border to hook up with its main line in Milford, Pennsylvania. The price of the gas in Algeria was to be $1.30 per million BTU as of July 1, 1973. The delivered price in 1983, when it was estimated the project would be completed, would be close to $5.00 per million BTU. The ERA denial stated: “We cannot conclude that the large, long-term commitment to an LNG project is now needed.”16 They went on to cite the availability of potential conventional supplies from the lower 48 states, including the Baltimore Canyon and other portions of the Continental Shelf. In addition, they pointed to the Alaskan North Slope supplies and the availability of domestic gas as well as synthetic gas from domestic sources. The ERA emphasized that, in the case of an LNG import proposal, national policy dictates a most cautious, even skeptical, assessment of each gas importation on the basis that LNG “generally represents a marginal natural gas supply for the USA at the present time.” The opinion concludes that TAPECO failed to demonstrate a critical national or regional need for the gas and noted that the project sponsors did not include a “contingency plan covering possible supply interruptions.”

Immediately after the ERA denied the TAPECO plan, they also denied a proposal by the El Paso Pipeline Company to import gas from Algeria. El Paso II originally proposed to construct at an East Coast site, but later changes suggested that terminal facilities should be built at Madagora Bay on the Texas Gulf Coast. The project had an estimated cost of approximately $4.5 billion and involved an importation of one billion cubic feet of LNG a day. The base price was to be $1.50 per Mcf in Algeria, and the price escalations were tied to No. 2 and No. 6 fuel oil in New York Harbor. The application indicated that the estimated delivered cost to the terminal in the third year of delivery would be $2.84 per million BTU, based upon 1975–1976 cost data. El Paso would have retained 650 thousand Mcf a day, and the remaining 350 thousand Mcf a day would have been purchased by United Gas Pipeline. Approximately 50 percent of the imported gas ultimately would have been provided for the California market.

The ERA opinion stressed that there are substantial supplies available for residential, commercial, and industrial needs currently from domestic sources as well as potential supplies from Alberta, Canada. It pointed to recent legislation which has increased prices substantially and, therefore, should have a significant impact on gas supply. Finally, the ERA concluded that the applicants failed to demonstrate an overriding national or regional need for LNG.17

Another important LNG project involves the importation of LNG from Indonesia. The sponsors of the project, Pacific Gas and Electric and Pacific Lighting, are the same companies behind Pac Alaska. The environmental questions of locating the terminal facilities in Oxnard or Point Conception were previously discussed. The project would require a capital investment of approximately $2 billion and would provide imports of 539,000 Mcf per day. The price in Sumatra is initially proposed at $1.25 per million BTU. An administrative judge of the former Federal Power Commission approved the project in July 1977 and indicated a delivered price of $3.59 per Mcf.

ERA, in Opinion Number One, issued in December 1977, disapproved of the Indonesian project. ERA objected to theescalator clause providing that future prices above the base price of $1.25 per million BTU should be tied to world oil prices for 50 percent of the escalation, and that the other 50 percent be tied to the Bureau of Labor Statistics consumer price index energy prices. The ERA strongly objected to the OPEC oil price impact on the price of gas, as well as the general inflationary rate relative to the cost of domestic energy prices. In

David S. Schwartz
addition, it rejected the currency reevaluation as a basis for modifying the price of LNG.

In Opinion Number Two, ERA approved the modified proposal and the importation of Indonesian LNG.\textsuperscript{99} ERA accepted the 50 percent escalation of the price tied directly to the movement of world oil prices, but would only allow an annual flow-through of 15 percent as a ceiling to the annual price increase. Anything above 15 percent would be carried forward and applied in future years to the extent permitted. Instead of the 50 percent escalation tied to the Bureau of Labor Statistics consumer price index, ERA accepted the use of the wholesale price index of all commodities, not just fuel costs. Finally, ERA accepted Pacific Indonesia's explanation that 25 percent was the limit on any adjustment for currency reevaluation. By accepting these modifications, ERA resolved all the issues in favor of the Indonesian sponsors. No weight was given to the intervenor's contention that the OPEC prices which would escalate LNG prices already reflected currency reevaluation, and thus this was double counting. No weight was given to the FERC staff contention that the devaluation of the dollar was already reflected in the wholesale price index, and that ERA was permitted a double adjustment for inflation. The FERC staff contended that the world oil price, or the Indonesian price which conforms to the world oil price, already reflects inflation in the wholesale price index which is used as a basis for escalating the price of LNG. It is rather anomalous that ERA would accept these adjustments in its second opinion, whereas it explicitly rejected them in its first opinion. A close reading would indicate that the rationale provided in the second opinion does not even qualify as a cosmetic effort to justify the Indonesian LNG import.

With the exception of the Indonesian project, which many attribute to political considerations, it appears that the ERA conceives of LNG as a very marginal source of supplemental supply which will only be permitted under the rarest circumstances. As for other LNG plants, such as Tenneco's North Star project to import gas from Russia or the Columbia System project to import gas from Iran, they appear to be nothing more than conceptual approaches to supplemental gas supply.

The Common Nexus of Supplemental Gas Supply Proposals

An examination of the various supplemental gas supply proposals indicates that there is a common thread running through all projects. All attempts to minimize risk to the firm by shifting it forward to the rate payer or to the taxpaying public. There is an effort by foreign suppliers to rely either upon commodity or market value of higher cost oil as a basis for the price of the gas supply. Another form of risk minimization is the use of rolled-in rather than incremental pricing for the supplemental gas supply. The Alaskan gas pipeline, the Pacific Alaska, the original El Paso and Transwestern coal gasification, and the current Great Plains proposals all rely on rolled-in pricing. This is true of the LNG plans from the original El Paso I import and the Trunkline proposal, and it is true of the rejected TAPCO and El Paso II projects.

Another means used by the pipelines to minimize risk is to insist upon a full cost-of-service, or all-events, tariff. In the original El Paso coal gasification project, the company proposed full cost of service, and currently the Great Plains sponsors have filed an all-events tariff. The rejected TAPCO LNG proposal contained an all-events tariff. Another device used to shift the risk forward to the rate payer is reflected in the Great Plains proposal to set rates for full recovery of all debt, which is approximately 75 percent of all capital required for the project. In addition, the Great Plains sponsors propose a surcharge to cover the cost of interest on debt as well as a return on equity while the project is being constructed. Another device is the insistence on a minimum bill, which was included in the original El Paso coal gasification project and other LNG proposals.

Another common element is the use of commodity or market value to determine the price of gas. For example, the base price determined for Canadian gas is the commodity value in the end market. In each Canadian project, the supply and demand are competitive. Pacific Indonesia, TAPCO, and El Paso II all also provide for escalations above the base price predicated upon either No. 2 or No. 6 fuel oil. Finally, loan guarantees, as originally proposed by Transwestern, shift the burden to the taxpaying public.

The common elements in these various supplemental gas supply proposals indicate a shift of both financial and operating risks forward to the rate payer and away from the firm. This is accomplished either on the basis of a full cost-of-service or all-events tariff, in addition to which the use of a surcharge, or a minimum bill, will minimize risk to the firm. Another critical consideration is that resource cost is no longer the basis for pricing; commodity or market value becomes the premise for the price which is being sought. Finally, there appears to be an insensitivity to market structure because the pipeline sponsors of the various supplemental gas supply projects are either producers or distributors of gas, and therefore there is no arm's-length relationship relative to the supplier or the purchaser of gas. Pipeline suppliers
benefit from their monopoly position in trading with affiliates and enjoy the captive market in which there is little competitive pressure on the firm.

Public Policy Recommendations

In assessing various public policy alternatives, I think it is necessary to divide the choices into near term or immediate and longer term or basic institutional approaches to gas supply requirements. It is important that regulators look at the growth rate of energy in general and at gas requirements as part of the overall energy balance in making short-term decisions. It is necessary to realize that with the recent significant increase in the prices of natural gas, conservation will become a much more important factor in the future than it has been in the past. Regulators must examine the need for supplemental gas supplies after assessing the available conventional supplies of natural gas. It may be more judicious to look forward to the utilization of surplus Canadian gas from Alberta or the availability of one or two trillion cubic feet of gas from the intrastate market to meet immediate needs. The basic question that policy makers face is whether projects should be postponed in light of the potential reduction in natural gas requirements generally and in terms of the availability of conventional supplies. Special scrutiny should be given to costly capital-intensive projects, such as the Great Plains proposal, which will provide very small quantities of gas at exceptionally high prices. This is also true of the Alaskan gas pipeline, with capital cost estimated at $12 billion and a delivery cost in the $5.00 per Mcf range. In the near term it may be advisable for policy makers to ascertain the impact of the Natural Gas Policy Act's higher prices on supply options from wellhead sources and carefully examine supplemental supply projects to determine their need on the basis of economic costs and supply availability.

In the more basic area of developing a public policy approach to assure equity and resource efficiency, regulators must closely examine the supplemental gas proposals that shift risk forward to the rate payer and away from the firm. One of the ways the gas industry does this is to rely upon rolled-in pricing and eschew incremental pricing of supplemental supplies. This completely destroys a real market test for allocating resources as well as for determining the potential market for the gas based upon its true cost. Rolled-in pricing destroys the only meaningful economic test of the marketability of the gas on the basis of its true resource cost. If, in fact, the gas cannot be sold at the incremental price because that will not permit the financing of the project, then this is a prima facie case that the project is uneconomic. In order to have a meaningful resource efficiency test, as well as an equity test, policy makers must rely upon incremental pricing for these supplemental projects and must avoid the use of rolled-in pricing. Regulators must realize that the current practice of shifting risk forward to the rate payer through such devices as the all-events tariff, surcharges, and debt guarantees will remove incentives for the firm to operate efficiently and innovatively. It is important for the firm to bear the risks associated with the project. The regulated firm in this regard is not different from any other and must bear the risk if, in fact, the pitfalls of inefficient operation are to be avoided. It is a dangerous precedent to permit this shift of risk from the enterprise: it postulates the premise for socialization of the firm if rate payers bear the risks.

Finally, regulators should realize that market structure reform can be a significant aid to regulation, and the introduction of more competition in this industry would help regulators achieve more efficient performance and lower rates for customers. If pipelines were made common carriers and new entry were permitted, this would provide greater assurance that, in fact, supplemental supplies would come forward on the basis of lower cost alternatives. The current available capacity in interstate pipelines could be more fully utilized for providing benefits for other rate payers on these systems as well.

A fundamental institutional reform required is functionalization and separation in the natural gas industry. This would limit producers strictly to the production function, pipeline to pipeline, and distributors to distribution, instead of permitting the same firm to control service at all three levels. This would provide a greater degree of arm's-length bargaining and competitive pressure, thus bringing about lower cost supplies to customers. That the competitive relationships do not exist is illustrated by the use of commodity value or market value in the pricing of gas. In a workable competitive market, it is not the price of other fuels that determines the price that a firm can receive for gas, but the resource cost to the firm plus a fair rate of return on investment. This obviously goes counter to the type of pricing that takes place in a monopoly market, where actual resource costs are incidental to the price demanded and usually approved by the regulatory commission.

Notes

9. Ibid.

**Comments**

*Richard A. Solomon*

The three papers by John Holtzinger, John Curley, and David Schwartz all seek to accomplish that most difficult task, namely, evaluating the future of the natural gas industry in a time of rapid technological, economic, and legal change. Schwartz has combined an interesting overview of the developing trends toward enormously expensive projects for supplemental sources of natural gas with a healthy skepticism of their necessity or usefulness at the putative prices and a plea for better evaluation before existing consumers, who may not benefit from them, and future consumers, who will have to pay the price, are saddled with the costs. Curley, in contrast, has assumed that such projects are necessary, or at least inevitable, and has given the more or less conventional viewpoint of the financial community as to how they should be financed.

At first blush, Holtzinger’s excellent paper on the so-called incremental pricing mandated by the Natural Gas Policy Act of 1978 (NGPA) and the inevitable problems in implementing this largely political corollary to the act’s price increase provisions might appear to be out of context, but I think it is correctly placed. The very fact that such a scheme was enacted into law, over the opposition of the entire industry, as a counterweight to the NGPA ordered price increases for conventional gas supplied should, I believe, tell us something in evaluating the nonconventional projects discussed by Curley and Schwartz.
Let me first discuss Schwartz's paper. He is, of course, correct in insisting that potential supplementary gas supply projects must be evaluated from a national policy standpoint involving the entire energy mix, rather than from the necessarily parochial view of the gas pipelines and distributors, or the hardly unbiased view of the oil industry, which controls most gas production and, unfortunately, much of the raw material for any other gas supplies. In the past, one of the primary problems in dealing with supplemental gas projects sponsored by pipelines has been that the analysis of the place of the project in an optimum energy picture has rested on inadequate knowledge of the price elasticity of both supply and demand. In all too many cases, this has led to an assumption of an unlimited market for all supplies of gas at any price, particularly if relatively captive existing gas consumers can be made to subsidize the marginal ones. It may prove true that rising costs of alternative fuels will always leave a comfortable margin for gas sold at rolled-in prices, but it is hardly appropriate to formulate national policy based on such an assumption.

Schwartz has suggested that a solution to the problem might be to require high cost gas from supplemental projects to be sold on an incremental price basis, thus allowing the market to determine whether there is a need for the project. I am inclined to believe this is impractical in view of the long lead time involved. Customers will be extremely reluctant to sign up for gas three to five years or more before deliveries commence, especially when, as will normally be the case, the estimated initial price will be significantly in excess of both existing pipeline or distributor rolled-in gas costs and estimated costs for domestic natural gas under the NGPA. Yet, project financing would appear to be impossible without a guaranteed market in terms of purchaser contracts. As Curley has made clear, even firm pipeline contracts with distributors may be insufficient in the absence of state commission guarantees that consumers stand prepared to underwrite distributor commitments.

The probable impracticability of incremental pricing as a selling device does not, however, mean that it should be abandoned as an analytical tool. The commission was clearly correct, in my view, in evaluating the feasibility of the proposals to transport Alaskan gas to the United States on an incremental basis. The difficulty with any analysis of the supply or demand situation for natural gas five to ten years in the future supports rather than detracts from the need for an incremental approach.

This leads to another problem with Schwartz's paper. For purposes of evaluating supplemental projects, he is, I gather, prepared to accept relatively optimistic predictions of the available natural gas supply under the NGPA pricing scheme, and to match them with relatively optimistic predictions of the impact of higher gas and alternative fuel prices upon usage conservation at the burner-tip. While, as indicated, I believe that considerably more attention must be given to both of these possibilities than is normally reflected in the presentation of sponsors of major supplemental projects, I doubt whether Schwartz's assumptions can be carried to what I suspect is their intended logical conclusion, namely, rejection of virtually all supplemental projects. Some on-hand gas supply cushion seems necessary in the event the elasticity of supply at NGPA prices proves less than the industry projects in its more optimistic moments or in the event that conservation remains at a lower level than we all might wish.

None of this seems to worry Curley. As long as the bankers have no risk and the pipelines and other equity sponsors of supplemental projects are compensated by high returns for their relatively small risks, the need for the supplemental project and the costs to be borne by present and future gas consumers are not his concern. What does worry him is that the mechanism for ensuring that the ultimate risk is imposed upon the ultimate user may not yet be in place.

To the best of my knowledge, there has been no study of whether project financing of supplemental projects on an essentially risk-free basis for all but the ultimate consumer is the only way in which such developments can be financed. Curley is undoubtedly correct in suggesting that the insurance companies which traditionally have been the primary source of debt financing for the pipeline segment of the industry are a far from assured source of money for supplemental projects when technical feasibility and cost attractiveness are less certain as compared to normal pipeline capital investment. And there is an attraction for consumers in capitalizing expensive projects on the basis of the 75 percent debt—25 percent equity formula apparently used as a norm in project-financed endeavors, in view of the tax deductibility of the interest payments and the normally lower rate for debt than for equity. But it is very difficult, if not impossible, for me to accept the contrary answer that it appears a real one if it is only maintained by the type of unusually high rates of return for equity which Curley seems to support.

Where I have most difficulty with Curley is in his view that gas consumers must bear virtually all of the risks, not only of interruptions in established service but also of failure of the project to become operational due to technological or economic considerations. In the case of a project such as the Great Plains coal gasification proposal mentioned by Curley, the contrary answer seems clear. This project,
in its present form, has no economic justification except as a pilot plant for a new technique which may become a vital element of our energy mix in the 1990s or beyond (or conceivably, pursuant to a nationwide crash program, in the event of a cut-off of all imported oil resources). In my opinion, there is no conceivable basis for suggesting that the risk of failure, either before or after the plant becomes operational, should be placed upon the customers of the five pipeline sponsors who, in the event of success, will secure a minimum amount of gas at a price which will almost certainly be in excess of alternate fuel costs. If the project is necessary in the national interest, then the costs, including the risks of project failure, should be spread nationwide. This could be done by a surcharge on all gas consumers such as that used to finance the operations of the Gas Research Institute. Preferably, in my view, it should be accomplished by government loan guarantees.

The Alaskan transportation system case presents a similar problem on an even larger scale, complicated by the fact that many of the key determinations have been made prior to any firm gas sale contracts, that is, prior to any knowledge of which consumers would be the putative beneficiaries of Alaskan gas. Here, the more extreme claims of the project sponsors for all-events tariffs — placing the entire risk for noncompleteness upon existing gas customers — have so far been rejected, although it remains to be seen whether the commission and Congress will hold the line when the inevitable renewed claims for protection are made and the effective allowable rate of return on equity is determined. Here, again, it seems to me, we are dealing with a project for the assumed benefit of all our citizens who depend on energy, and not merely the existing users of natural gas upon which Curley wishes to impose virtually all risks of failure. In these circumstances, I do not see any validity to Curley’s theological arguments against governmental, that is, taxpayer, guarantees.

Finally, I come to Holtzinger’s definitive analysis of the “incremental price” provisions of Title II of the NGPA and his exposition of the implementation problems facing the FERC. As Holtzinger acknowledges and I would like to stress, incremental pricing was an integral part of the political compromise which was necessary to secure agreement to the provisions of the act for a substantial immediate increase in maximum rates, and eventual deregulation, of producer sales to pipelines. The intent was to load as great a percentage of the costs of the gas increases as possible upon industrial customers, to cushion the blow to residential and small commercial consumers, that is, voters.

Of course, there were complications in achieving this objective. It was soon recognized that placing too large a share of the expected NGPA cost increases on industrial consumers would lead to their switching to alternate fuels, and residential consumers would then have to absorb the pipeline and distribution fixed costs previously paid for by the industrial users. To cope with this difficulty, the higher costs incrementally allocated to industrial users were limited to alternative fuel levels. And, as can be expected, numerous exceptions were built into the scheme, some of which, in my opinion, have considerably more merit from a policy standpoint than others. Holtzinger has highlighted many technical problems in implementing the congressional mandate, and certainly the task before the FERC is not an easy one. I doubt, however, whether the commission’s experience with the egregious curtailment proceedings, to which it has devoted so much time in the last few years, forecloses all the difficulties in coping with its new responsibilities that Holtzinger fears. One would expect that the commission will have learned by its (and our) mistakes. A major trouble with the curtailment cases was that the need for immediate action caught the commission and the industry without warning, with the result that individual pipeline cases were set for hearing without any standards for evaluation, and the “guidelines,” when they eventually were issued, were extrapolations from an atypical situation which created as many procedural problems as they resolved. I would hope that the commission, given nine months to put the initial plan for boiler fuel users in place, will be able to move toward a reasonable resolution of the major issues. In this endeavor, the commission should be able to build upon the lessons, and more detailed data bank, which are the products of its experience with curtailment proceedings.

Holtzinger also questions whether the result will prove to be worth the effort and cost involved. To the extent that he speaks for the industry in expressing an unequivocal skepticism as to the outcome, he reflects the industry’s basic lack of knowledge of its existing and potential market at the retail level. I remember all too well the early curtailment days when the pipelines paraded with pride their lack of knowledge of the nature of the usage of their gas at the retail level, and their essential disinterest in the fights between distributors and end users for priority. I suspect, however, that even if the industry had been able to quantify the impact of the congressional scheme and to show that the potential savings to residential users were minimal, Congress would have acted in essentially the same way. For the basic question was political, and the fact that gas consumers may have to
Comments

pay indirectly the additional cost imposed on industrial users by any incremental costing scheme does not change the political equation.

My own uneducated guess is that the costs shifted from the residential consumers will prove to be more significant than Holtsinger suggests will be the case. But, since it will only serve to reduce the substantial increases which will, in any event, result from the producer rate increases authorized by the NGPA, the political impact of incremental pricing may well turn out to be less than its congressional sponsors hoped. To me, however, what is most significant about Title II of the NGPA is Congress's implicit recognition that, valid or not, residential consumers are not likely to accept with equanimity the higher costs for gas which are, we would all agree, the inevitable result of any comprehensive response to the contingency energy crisis.

The underlying problem is akin to that with which state regulators are all too familiar with respect to fuel clauses and CWIP. The natives are restless, and it is going to be increasingly more difficult to explain to them that increased costs are for their own good, or that they were subalidized by past regulatory policies and should now be willing to let the laws of supply and demand work their will. They will have to be convinced that these increases are necessary and that the responsible government officials have taken all available steps to mitigate the cost burden. This will include ensuring that private parties seeking a risk-based return are prepared, within the limits of their capacity rather than their desires, to assume their share of the risks. As indicated at the outset of these remarks, this question will turn out to be at least as applicable to the evaluation of proposed project-financed nonconventional gas endeavors as it was during the enactment of the Natural Gas Policy Act.

Richard A. Rosan

The three presentations by John Holtzinger, John Curley, and David Schwartz are interrelated since they consider separate aspects of several major gas industry problems that will affect the nation's ability to cope with its major energy needs.

William Melody has noted that the nation is confronted with a serious energy problem, and many of the activities being pursued — such as the rate design commandments of the National Energy Act of 1978 — are largely irrelevant in attacking the basic issue, which is meeting our nation's energy needs.

I view the incremental pricing commandment of Title II of the Natural Gas Policy Act as falling in that category. Apart from its irrelevance to helping our energy situation, I question the wisdom of legislation that incorporates directives of the kind embodied in Title II — ten pages of detailed legislative instructions. It was political gimmickery of the worst sort; it will cost the consumers of gas and the taxpayers vast sums. It is a further manifestation that our politicians do not understand that there are limits to what U.S. industry can bear and still remain competitive in domestic and world markets. But, even more important, the administrative difficulties, the continuing costs of com-
pliance (which consumers will pay in one way or another), will far outweigh the short-term benefits, if any, that the residential and small commercial user will realize.

Some background may be helpful in appraising the three papers. Commencing in the early 1950s, the nation’s energy consumption grew at an unprecedented rate. So did the standard of living of a large percentage of Americans. Equally important, the expectations of our people had no bounds. Our political leaders seemed convinced that there were no limits, in resources or money, to satisfying such expectations. In the case of basic energy services — gas and electric — a series of circumstances supported the view that everyone could have unlimited amounts of energy at bargain prices. The industries did nothing to counter this opinion. The need to grow was the focal point of most management thinking.

One measure of this situation is that natural gas production increased from 10 Tcf in 1955 to 21.9 Tcf in 1970, peaking at 22.6 Tcf in 1972. Gas, a premium fuel by almost any standard, was being made available to everyone at prices that in some cases were less than the cost of a coal-fired furnace.

It was not the projected decline in domestic gas production, however, that caused an examination of our energy posture; it was the oil embargo of 1973–1974 that set some people thinking about our nation’s vulnerability in the energy area.

There was a flurry of legislative interest. The Federal Energy Act was hastily drawn and passed in early 1974; FEA embarked on Richard Nixon’s Project Independence, and the resulting blueprint, hurriedly done by people with little, if any, energy background, proposed solving the nation’s energy problems with an all-electric economy. After a short time this blueprint was discredited. But even today, there are strong advocates in government of an all-electric economy and a phasing out of the gas industry. They, also, in my view, will ultimately be discredited.

Since then, we have stumbled from one program to another. The American people have been confused and frequently misled. The recently enacted National Energy Act only touches upon some of the areas needing consideration, but in large measure it rests on a conservation ethic. So five years after the oil embargo, we have failed to formulate sound relevant policies that will meet the energy needs of our nation. Thus, I am far from impressed with Schwartz’s appeal that we rely on the policy maker’s viewpoint as distinguished from the recommendations of industry. It is quite natural for one who has been closely associated with the policy makers to stress their approach. Based on the track record to date, we must be very skeptical of their thinking.

Let me return to the conservation solution to our energy problem. Clearly, the expectations of all Americans demand a dynamic society. The 30 percent of our people who are less fortunate expect to raise significantly their standards of living, to acquire better homes and all that goes with them, such as labor-saving devices. Our population is projected to grow from the present level of 220 million to 253 million in 2025. That is 35 million more people, or an increase of 700,000 per year. To maintain and create jobs for all requires economic expansion. That will require a large increase in our energy supplies. Even while propounding its conservation policy, the Carter administration projects in some of its studies a significant growth in energy use.

Thus, I suggest that we must promptly delineate and define our future energy needs and develop sound and realistic programs for satisfying them. This is the relevant course. There are some credible estimates of the problem. In 1978 the nation will use about 77–78 quads of energy; most forecasts project a range of need in the 1985–1987 period of 90–100 quads, an increase of 15 percent or more in only seven or eight years. Where will these added quads come from? Equally pertinent is how we propose to meet the need of 110–20 quads in the year 2000, only 22 years away.

Confining our consideration to natural gas or its equivalent, the subject of these papers, there is a strong belief that the current level of domestic production in the lower 48 states will slowly decline from the current level of 19–20 Tcf per year to 12–14 Tcf per year in 2000. Of course, this slow decline assumes adequate prices to allow us to find the new reserves; without adequate prices, the decline will be more severe. However, if gas is to play its rightful role in the nation’s energy posture, the industry must be able to augment its domestic production with a continuation of Canadian imports (8 to one Tcf per year), synthetic gas (5 Tcf per year), LNG imports (3 Tcf per year), Mexican imports (one Tcf per year), Alaskan gas (3.6 Tcf per year), coal gasification (3.3 Tcf per year), and new technologies (5 Tcf per year). Thus, by 2000, the gas industry should be able to supply 29–31 Tcf per year or maintain its present position of supplying about 25 percent of the nation’s total energy.

Within this context, how do the three papers relate to the issue of meeting our nation’s future energy needs?

Holtzinger has covered extremely well and in a comprehensive fashion the problems associated with the concept of “incremental pricing” as prescribed by the Natural Gas Policy Act of 1978. As he says, it is not incremental pricing in the sense economists usually use the term.
He has noted the severe administrative problems it poses. Obviously, establishing and maintaining an up-to-date data bank, developing a mechanism for a current determination of the competitive prices of a variety of alternate oil (and there are three types) for the various regions of the country, and setting up the administrative mechanism for passing through from the first purchaser to the last seller the cost of the higher priced gas will make incremental pricing a costly process. As indicated earlier, the concept is, in my view, wholly irrelevant to the central issue facing the nation. However, insofar as the concept may reduce gas industry earnings because of the delay in recovering higher gas purchase costs or by reducing industrial sales, it can be harmful to the nation's future energy posture. If, because of incremental pricing and its accompanying uncertainties to the buyer, the shift is to imported oil, the concept would aggravate our imported oil problems.

Curley's and Schwartz's papers are most relevant concerning the future role of natural gas in the nation's energy mix. Curley explains the very difficult financing problems confronting the industry on some of the major supplemental projects. Schwartz ignores this problem; either we do not need supplemental gas projects, or they can be financed in the manner of previous projects. Since I have worked very closely on some of the major supplemental gas projects that pose unique financing problems, I not only understand Curley's discussion but also am persuaded that the questions his paper poses must be faced. In money matters, we must recognize the real world, the concerns of those who put up the principal, the suppliers of capital are trustees of other people's money, such as the large insurance companies.

Assuming we need supplemental gas resources, and I believe very strongly that we do, they will not be developed in a timely manner under the approaches advanced by Schwartz.

Of course, my first major difference with Schwartz's paper concerns the underlying theme that because of conservation and other factors, we may not need supplemental gas supplies in the future. Instead, we should use short-term expedients (the bubble of gas in Alberta, Canada, and the one to two Tcf of intrastate gas) to meet the immediate needs and postpone the high cost, capital-intensive projects. This is, in my view, a most shortsighted and unrealistic approach. It would have us gamble with the future economic well-being and national security of this country. The approach ignores the lead times required to bring major new projects on line. It assumes that ventures such as the Great Plains coal gasification project can make the fuel, when it is needed desperately, immediately available. The Great Plains project is needed now as a tool to find the most economic and efficient way of providing gas from coal. The time for research and experimentation is now, not when the nation has been brought to its knees for lack of energy!

Schwartz cites the denial by the administrator of the Economic Regulatory Administration of the current El Paso LNG project on the basis of a lack of need for the gas. In my view, the ERA's decision is nothing less than incredible. It uses the demand for gas in 1985—six years from now—as the test of need. Our policy makers had better start thinking of our need in the year 2000. For all practical planning purposes, we are well past 1985. Aside from that, the ERA assumed that currently available domestic gas, or even less, would take care of our gas requirements in 1985.

The present negative policy of DOE toward LNG is difficult to understand when one realizes that there are vast amounts of natural gas in the world. It has been reported that the Arubians flare 6 Tcf per year, almost one-third of our production.

Many things must be done now in recognition of our future needs. They should not be postponed on the basis of some comparative current cost of a resource or the current availability of a resource, in most cases not priced at replacement cost. By urging the incremental pricing approach as a resource efficiency test, Schwartz would effectively stop all current supplemental gas projects. This would be, to my thinking, one of the worst decisions our policy makers could make. Many times, Curley explains the problems of financing new large cost projects such as the Alaskan Gas Pipeline or the coal gasification plants. Much of the problem stems from the rapid and continuing inflation in heavy construction costs. Thus, the natural gas industry, with total plant of about $55 billion at the end of 1977, net gas plant of $34.4 billion, and total capital stock equity of $23.8 billion, does not have the financial muscle to finance projected capital expenditures by 1985 of an estimated $100 billion without some new and innovative approaches, such as those outlined by Curley. This would not happen if Schwartz's advice is taken. If he cannot perceive the new conditions as to financing, then his conclusions are devoid of merit.

Schwartz urges that "regulators must realize that the current practice of shifting risk forward to the rate payer through such devices as the all-events tariff, surcharge, and debt guarantee will remove the incentive for the firm to operate efficiently and innovatively. It is important for the firm to bear the risks associated with the project."

This is wonderful theory, but it overlooks two very important factors. First, the current inflationary economy and the cost of new resources
have placed most major new projects beyond conventional financing, even if supported by a large number of industry members. The total gas industry does not have the resources commensurate with the undertakings, nor can it absorb the much higher risks associated with some of the projects. Thus, as Carley sums up the issue: “For the large projects of today, the problem of risk-bearing is like a double-edge sword. Lenders are concerned that the sponsoring credits are not strong enough to provide meaningful credit support at the same time that sponsors are finding it impossible to justify economically the direct and indirect guarantees required by the debtholders.” Either Schwartz has not been listening, or he does not believe what financial experts such as Carley have been saying to the sponsors of these major projects.

The second factor is the current “risk of regulation.” By this I mean the totality of regulation confronting any new project, including untoward delay, unreasonable requirements, and the uncertainty of future regulatory actions. Regulation has reached such levels that the suppliers of capital are increasingly wary.

As suggested earlier, Curley and Schwartz deal with topics that are most relevant to the future energy posture of the nation. Curley suggests what must be done in the financial areas by the regulators if our objectives are to be realized; Schwartz’s approach, which incidentally is shared by a few others who believe they are protecting the rate payer or the consumer, would move us, in my view, down the primrose path to national disaster.
Integrating Efficiency Standards into the Regulatory Process

Thomas K. Standish

The regulatory process must be guided by some set of principles if constitutional requirements are to be met. In general, in order to produce a legally defensible “end result,” regulators have adopted one or another of the traditional cost-of-service methods to set rates. In addition to this legal standard, reliance by regulators upon accepted cost-of-service methodologies for determining the level and structure of rates has generally been thought to conform with another standard — that of economic efficiency. In this sense, legal and economic standards have been regarded as coincident and mutually supportive.

Despite the availability of cost averaging, trending of costs, embedded costs categorization, and other elements of traditional normalizing techniques, existing cost-of-service methodologies to gauge fairness and reasonableness have produced what only can be described as a vague and imprecise “end result.” With regard to this latter observation, it is curious that this fact has been of little concern to regulators until recent times. This is especially odd when one thinks of the tremendous importance of regulated industries to our economy and society. One wonders why so little attention has been focused upon the development of a body of techniques to make the regulatory process more scientific.
The Traditional Basis for Producing a Fair and Reasonable "End Result" Is under Challenge

It might be argued that the poverty of sophisticated techniques used to assist in the regulatory process is to be expected because econometric and computer methodologies have been available only in very recent times. Comparison with other sectors of the economy, where these techniques are in widespread use, reduces this to a hollow argument. The answer must lie, not in an absence of available techniques, but in the failure of the administrative will to apply them. This in turn requires that another question be asked: Why this lack of regulatory will?

Passive Approach to Regulation Has Prevailed

Interestingly, the answer to this latter question may lie in the prevalence of a particular view, embedded in the fabric of our society, concerning the role of economic price. If there was ever an example of Lord Keynes's dictum to the effect that we are all unwitting proponents of some defunct economist's theory,\(^2\) this is a case in point. What is suggested is that the passive approach to regulation, which has dominated the field to this very day, is built upon the presumption that the operation of the competitive market will automatically create a socially desirable balance of conflicting interests making a claim upon economic resources and output. Convinced of the higher good which flows from leaving management to its own designs, regulators have willingly relegated themselves to the limited role of setting economic price at a level which reflects the so-called competitive standard. This conviction has been translated into acceptance by those entrusted with the regulatory function of the cost-of-service evidence offered by regulated utilities. In a word, passive regulation has resulted from the general acceptance of the view that economic efficiency is promoted by letting the utility businessman be captain of his ship, free to compete in the marketplace, unhampered by public intervention except as to the setting of economic price at a level which cuts monopoly profits.

Competitive Standard No Longer Sole Guide to Rate Making

The recent controversy over rate reform, of course, has changed all this. Not that the ideas of economists have disappeared from the current rate reform controversy; \textit{au contraire}, it is clear that more defunct economists are turning over in their graves than ever before. But what is equally clear is that the very existence of this controversy signals the end of the universal acceptance of the competitive standard as the basis for setting public utility rates. In contrast to their past history, theorists of the competitive standard must now compete for equal time with a growing group of theorists advocating alternative concepts for setting economic price.

By no means have the old ideas based upon the competitive standard disappeared from the the current struggle over setting economic price. These ideas are very much alive and, in their most refined and precise form, continue to dominate the current discussion, namely, in the form of marginal cost pricing of utility services. From this fact, some observers interpret the advent of widespread consideration of marginal cost ideas as proof positive that the competitive standard still reigns unchallenged as the guiding principle for meeting legal and economic needs in the ratemaking process. This view is mistaken.

Rather than marking the strength of the traditional view, today's marginal cost refinements are the last defensive act of the old view in response to challenge. Hence, what is apparent strength actually is a profound weakness. The appearance at this particular time of the extreme refinements of competitive standard theory signals the collapse of that theory, and with it, it signals both the end of universal acceptance of the competitive standard as the basis of ratemaking and the beginning of a new era of ratemaking.

It is not accidental that regulators are looking for alternatives at this time. Coupled with a secular trend toward increased risk in the economy, the advent of rapid change (particularly in the energy sector) has led to a profound rethinking of accepted ideas. In sum and substance, the lack of administrative will to apply new techniques in the past, referred to above, has been transformed into an active search for new answers to meet today's real world problems. As will be argued below, when reduced to its essential elements, this search is identical to a quest for the redefinition of the meaning of economic efficiency. The kind of regulatory "end result" flowing from this quest differs significantly from that with which we are now familiar, and, by definition, a new role has been cast for economic price.

Fair and Reasonable End Results and Economic Efficiency

Having opened the door to the abandonment of the competitive standard as the exclusive principle for setting regulatory price, regulators are now confronted with the direct responsibility of resolving two
this logic is that, under either the traditional or the new wave theory, a fair and reasonable regulatory price is identical to that which promotes economic efficiency. To evaluate the extent to which either of the competing views on regulation is correct, it is necessary to evaluate the model which underpins the two schools of thought.

Realism Determines Usefulness

In any theoretical investigation, the nature of the central question posed is what gives the inquiry structure; by definition, it constrains and circumscribes the kind of answers which can flow from the analysis. Framing the central question of any theoretical context, in turn, devolves into a whole set of secondary questions which are subordinate to, and which are circumscribed by, the nature of the primary question posed. Asking the wrong question can only produce wrong answers. In terms of appraising the great debate on rate reform, this problem is not without interest.

There is a severe penalty which must be paid for pursuing answers to wrong questions. Framing the central question improperly, in other words, has the unfortunate consequence of producing rather extensive and elaborate searches into blind alleys. The words of a French epistemologist (whose name escapes this author’s recall) are apposite: “We are indebted to Emmanuel Kant for the suspicion that problems which do not exist may give rise to massive theoretical efforts, and to the more or less vigorous production of solutions as fantastic as their object.”

As that French epistemologist would undoubtedly point out in this case, what is at stake is theory in search of the correct “object” of analysis; more properly, what is at stake is whether or not the theory will work when applied to real world problems. To recapitulate, an economic model is nothing more nor less than a central question and a set of subordinate questions; this being the case, the test of the usefulness of a model is the extent to which it represents reality—or, in our Frenchman’s terms, the extent to which the correct “object” of inquiry has been properly perceived by the theorist.

Based upon this criterion, an inspection of the models which underpin the traditional and new wave price proposals is, then, the acid test as to the usefulness of each of these proposals for setting public utility rates.

Curiously, Keynes was right about defunct economists, and each of the current positions discussed above is heir to the scrutinings of now-defunct economists. In fact, the models of economic price which pervade today’s regulatory discussions are derived from two schools of economic thought—the classical and the neoclassical—and, in par-
ticular, from the model of economic efficiency which forms the core of the thinking of each of these schools. For insights into the merits of current advocates' arguments, there is, then, value in picking over the thoughts of defunct economists.

The Historical Search for a Model of Economic Efficiency

From the writings of the French Physiocrats in the early 1760s, the single and overriding preoccupation which has dominated the theoretical content of virtually every major treatise in economics is that economists have been striving to understand the nature of increases in economic efficiency. It is not an accident that economic efficiency is the common thread tying together the history of economic thought, nor is it a coincidence that differences in central assumptions concerning the nature of economic activity define the key differences between models of economic activity. After all, increases in economic efficiency have been the source of improvements in material well-being to the household, and in mirror image of this same phenomenon, efficiency increase has been the principal source of profit for business. Naturally, economists have wanted to understand economic efficiency and have "posed questions" about it in their models.

In general, two basic models of economic efficiency can be discerned from the scruffing of defunct economists. Both of these pose a central question, and both can be tested against reality — their "object" of the theoretical inquiry. It is only this test which determines the degree to which each model may lead to a search of the unknowable due to having posed the wrong question. In addition, each model must be examined for consistency, stability, and the degree to which it is capable of including phenomena about which policy claims are made. To gain a perspective of the validity of each of these models for real world policy purposes, in turn, the classical and neoclassical models will be discussed, commencing with the latter because it still dominates regulatory thinking.

The Neoclassical Model of Economic Efficiency

Current marginal cost proposals are taken from theory which emerged in its complete logical form in the 1870s with the so-called marginalist revolution. This analysis is logically elegant and has been progressively refined by translation into geometric and mathematical representations of the basic theory.

The central question posed by this school is: What is the Pareto optimum price solution for the economy? This query, in turn, devolves into a matrix of secondary questions derived from and compelled by the nature of the central one. For example, the marginalist model requires asking such subordinate questions as: What is the short-run marginal cost of firms in the market sector of the economy? What is the short- and long-run marginal cost of monopoly firms in the regulated sector of the economy? What is the second best solution for a regulated firm when prices cannot be set directly at marginal cost? What are the marginal costs of firms producing substitutes and complements to public utility service? It asks these and other questions with which we on the regulatory scene are now all too familiar.

It might be argued that the central question posed by current proponents of marginal cost pricing does not necessarily involve the Pareto optimum, but it is clear that the model used in the EPRI study and by other well-known adherents to this version of economic efficiency does, in fact, reduce to this central question. To wit: "Economic efficiency is based on the concept of Pareto Optimality, an allocation of production such that no consumer can be made better off without another consumer's being made worse off."

From the perspective of the great rate debate, there is a more important reason why the Pareto optimum is the foundation upon which the marginalists build their conclusions. The form of economic efficiency which flows from this model of the economy is such that the operation of competition automatically provides optimizing answers to the questions of economic efficiency by what means and economic efficiency for what purposes and for the benefit of whom.

Without the validity of the Pareto optimum, marginal cost pricing collapses as an independent guide to economic price, leaving as the only alternative the active setting of regulated price according to criteria extrinsic to the operation of the market itself. In other words, if it can be shown that the neoclassical model is not valid, then the competitive standard cannot be relied upon for setting public utility prices. Hence, from what may appear to be dry and uninteresting analysis, there is a great deal at stake. In a word, the collapse of the marginal cost model of economic efficiency means the collapse of an era of regulation. Enter the concern for the central question posed by new wave theorists.

With this background in mind, a brief review of the neoclassical model follows.

FATAL FLAWS CONTAINED IN MARGINAL COST THEORY. An exhaustive analysis of the neoclassical model of marginal cost is not necessary to demonstrate that it is invalid for ratemaking purposes. What is sufficient is to point out flaws in the theory which are fatal to: (1) logical consistency, stability of model equilibrium, and theoretical completeness, or (2) the realism of the assumptions underpinning the
theory. While a fatal flaw on either count is sufficient to destroy the usefulness of the theory, marginal cost pricing principles are fatally vulnerable on both counts. An examination of the former, where the theory is least vulnerable, is offered first.

The marginal cost model lacks completeness, stability, and logical consistency, except in the trivial case.

In order for a theory to act as a valid guide for policy, no matter how accurately its assumptions reflect reality, it must be internally consistent in the most expanded and complete form from which policy claims are made. The second test is that, in its most complete form, a theory must encompass and represent the entirety of the phenomena for which claims are made about reality. Hence, the more sweeping the policy claims made about reality, the more difficult it is for a theory to meet these logical tests. In addition, even if these criteria are fulfilled, a model of economic activity must exhibit stability in its equilibrium condition in order for the conclusions drawn from the model to have policy meaning.

It is conceded at the outset that the restrictive conditions of general equilibrium spelled out in the model in all probability will result in a Pareto optimum and, hence, can be shown to promote economic efficiency as defined by the model. However, in order for traditional microeconomic theory to serve as a guide to policy for the overall economy, it must be shown that a stable, general equilibrium can be attained wherein all firms in all markets either are producing or will tend to produce at their optimum point of equilibrium. The condition underlies the normative conclusions derived from the Pareto optimum. Even the best case for the marginalist model cannot meet this stability condition due to circularities of causation in the model once out of equilibrium. The problem is that, even in the pure form of the theory, both demand and supply interact to determine price. Unless unique supply conditions exist in equilibrium, wherein the firm will offer slightly increased or decreased quantities at the same price, a disequilibrium of the general economy results from interaction between demand and supply in each market. It is important to note that, in the marginal cost model, the concept of equilibrium which is necessary to reach conclusions about economic efficiency is such that an equilibrium must be attainable at an instant in time. The Pareto optimum is not sequential.

Critical to the theory of the marginalists is the proposition that a stable general equilibrium exists. Yet, such an equilibrium does not exist because, once out of the unique equilibrium which the theorists postulate, the interaction between supply and demand in individual markets for goods and services will produce an infinite array of alternative income distributions. With each income distribution, a new set of demand and supply conditions will be generated, and so on into hopeless circularity of causation. There is no inherent reason that this spiral of interdependence, once in disequilibrium, should tend to produce a unique equilibrium necessary to Pareto optimum conclusions. In fact, it is well known that, except under the most restrictive of assumptions, the general equilibrium solution cannot be arrived at except by assuming this equilibrium in the first place.

Since stability conditions do not exist, even under assumptions most advantageous to the marginal cost model, the conclusion which follows inexorably is that the whole concept of economic efficiency spelled out in that model collapses upon itself. Clearly, once these highly restrictive assumptions are relaxed even further to include real world phenomena about which the marginalists make policy claims, the model becomes hopeless as a guide to economic efficiency.

To be complete, the marginal cost model should encompass the activities of the government sector; yet, the introduction of this sector into the model makes the Pareto optimum impossible.

Although proponents of marginal cost pricing make policy claims which recognize the existence of an active government sector in the U.S. economy, the theory from which the normative Pareto conclusions are drawn admits of no government sector. By this I do not mean that a mathematical model cannot be drawn up which sets under conditions a government sector which is totally divorced and independent from the private sector of the economy; what is meant is that a government sector which is an integral part of the economy destroys the ability of the marginal cost model to produce Pareto results.

The fundamental problem is that the Pareto optimum "works" only if the marginal costs of production can be shown to be important to changes in that sector, government or private, which produces any given type of output. If such a variation does exist, then the prices in the marketplace can no longer act as signals of economic efficiency.

Problems in the theory stem from two factors. First, the "pricing" of the public output cannot be made consistent in the model with prices of output produced in the private sector. Second, determination of the total social marginal cost of any output for any firm is conceptually inconsistent so long as there exists an interdependence and intermingling of private and government production and of consumption demand for public and private goods (and services).

Turning to the first of these issues — government supply of output — encompassing the government in the marginal cost model creates an
inherent inconsistency. This stems from the fact that the value of final output produced in the government sector differs from the value of the identical output produced in the private sector. This complication is compounded when the outcome of a branch or function of government becomes the input to another; hence, in the input-output chain of production starting from raw materials and ending with final product, the "marginal costs" in the government sector could contain significant undervaluation when compared with the result which would obtain had the public goods been produced in the private sector and priced at private marginal cost. This is a problem in the theory even if the government and private sectors are completely divorced from one another. We turn now to the second factor alluded to above—government divorced.

Not only is there an inconsistency which arises in the model due to the "mark-up" problem mentioned above, but also the very basis for valuation of government output is inconsistent with that of private output. Price values in the private sector are a direct function of the unequal distribution of income, whereas egalitarian participation determines the "pricing" of public output. Given this dramatic difference in the basis for valuation of output as reflected in economic prices of both the government and private sector of the economy, then a question arises: What is it that is argued by the marginalists when it is suggested that economic efficiency is promoted by setting economic price equal to marginal cost? The fact is that, given this inconsistent basis, setting price equal to marginal cost loses its operational and substantive meaning. From all of this there emerge unanswerable the two questions raised earlier: economic efficiency by what means and economic efficiency for what purposes and for the benefit of whom. Because the essence of this second source of inconsistency has its genesis in the demand side of determining price, the answer to these questions must be found in some means of specifying consumer demand in the ratemaking process.

In summary, if economic theory is to be useful to public utility ratemakers, it should be demonstrable how resources used in production (including labor) are best organized to promote the maximum possible output in response to the needs of society. This is to say that prices and levels of output should produce overall economic efficiency. As pointed out above, the normative question of economic efficiency for what purposes and for the benefit of whom is thought to be answered automatically by competition as demonstrated in the Pareto analysis. The introduction of the government sector into the theory presents, however, a contradiction in the very foundations of traditional theory. It is this: Marginal cost analysis only works if inequality in the distribution of income and wealth is accepted as a normative social norm. (The private marketplace then allocates resources and goods to conform with the social needs of individuals as expressed in dollar market demand.) However, whereas Pareto analysis demonstrates how to organize production to conform optimally with unequal money demands of individuals in the market, the same cannot be said of the ever-growing portion of total goods and services which are being produced in response to egalitarian demands originating from the government sector. What results, then, is one demand sector of the economy which responds to inequality and another which responds to equality.

Any number of logical devices can be contrived to reconcile this normative inconsistency. They all must fail, however, once the interaction between the government and the private sector is recognized in the model. This is to say that, once public output becomes an input into private production, the basis for reconciling public ends with private means (and vice versa) no longer exists.

For example, what good would it do to price private output equal to marginal cost when, in essence, portions of that output are actually "priced" according to nonmarket principles? How would adjustments be made in market prices so that firms in different industries, using a different mix of government inputs, could price according to the cost to society of producing the marginal unit of output? Exactly what kind of economic efficiency would result even if such adjustments could be made?

Both the inconsistencies resulting from incorporating the government sector into the theory are compounded and magnified by virtue of the interrelationship between the government sector and the rest of the economy. It is known from input-output analysis that there exists today a near complete interlock between sectors and industries of the U.S. economy. Even this type of analysis, however, does not depict the degree to which developments in technology, education, health care and medicine, defense, and so on, are causatively interrelated over time between and within both the public and private sectors. Nor does the analysis, among other things, show the interrelated aspects of government institutional factors (laws, programs, and so forth) with private sector determinations of product quality, price, and levels of output.

Although it is intuitively and analytically obvious that there is cross-sectional interdependence in production and consumption between the government and private sectors, this is further revealed
when an historical view is taken. Many of the functions presently
performed by the public sector were, at one time, private goods and
services sold in the marketplace. The symbiosis of public-private in-
teraction has produced an economic development path which is now
regarded as normal for the U.S. economy. Yet, at each moment during
this development, a cross-section of commodity and service output
would reveal differing segments of this output priced according to
profit and public benefit principles. Would the proponents of tra-
ditional theory claim that cross-sectional consistency exists, at any given
time, with the knowledge that dynamic shifts have placed an increasing
proportion of economic output in the public sector, where pricing
principles are at such variance as those in the private sector?
At least for these reasons, in the general equilibrium form neces-
sary for a Pareto optimum, marginal cost theory does not contain
a government sector. Attempts to integrate government into the general
equilibrium model lead to glaring inconsistencies in pricing principles.
The conclusion: Marginal cost theory either lacks completeness on
positive theoretical grounds (in the former case), or it is theoretically
inconsistent and unstable (in the latter).

The existence of noncompetitive resource and product markets makes impos-
sible the determination of Pareto optimum.

In the general equilibrium model used to generate a Pareto op-
timum, firms' operations are assumed to be of a size sufficiently small
such that no one firm can substantially affect market price. These
conditions of pure and perfect competition extend to both product and
resource markets. Noncompetitive industry and factor markets, by
definition, tend to produce output levels and prices which are not
defined as economically efficient.

The fact is, of course, that our economy is dominated by bigness,
not smallness, and as a consequence, the conditions necessary for a
Pareto optimum are not extant in the real world, that is, not without an
additional set of assumptions. To address this theoretical flaw, two
avenues of escape have been constructed by the marginalists. One is the
theory of "second best." The assumption is made that all other firms
but one monopoly firm are assumed to price according to marginal cost
principles induced by perfect competition; in addition, the monopoly
firm is forced to produce and price according to restrictive conditions.
Note that this second best solution works when the only exception to
perfect competition is single monopoly.

The second avenue of escape rests upon even more absurd analysis.
Having demonstrated that pure and perfect conditions of competition
are necessary for normative conclusions to flow from the theory, mod-
ern advocates of marginal cost pricing have resorted, curiously, to
calling a rose by another name. In the place of oligopoly and monopoly,
these writers have substituted the label "effective competition" to de-
scribe the U.S. economy. What this assumption reduces to is the asser-
tion that the giant firms in today's economy will act to produce exactly
the same pricing and output decisions as would be the case had
thousands of small firms produced the same goods and services.

In both cases, the resort by the marginalists to the use of theoretical
artifice shows two things. First, these theorists recognize the absolute
necessity for the effects of pure and perfect competition as a component
of their theory. Second, it shows the extent to which these theorists, like
Procrustes, will cut and stretch reality in order to preserve their idea fixe.

It should be noted that microeconomic theory has been developed to
analyze monopoly, oligopoly, and imperfect competition (considered as
partial equilibrium theory of the firm). These partial theories have not,
however, been integrated into a general equilibrium framework
because to do so would result in the problems of circularity
outlined above. (This is particularly the case if the industry structure
most prevalent in the U.S. economy — oligopoly — is admitted into
the general equilibrium framework.)

The outcome of the foregoing is that, on grounds of pure theory,
marginal cost analysis is either incomplete when considering only pure
and perfect competition, or inconsistent and unstable when expanded
to include the noncompetitive market forms of which most of the U.S.
economy is comprised.

The marginal cost model cannot account for social marginal costs.

Problems with completeness and consistency also arise when an
attempt is made to incorporate social costs into the marginal cost
model. This is particularly disturbing inasmuch as the raison d'être of
the model is to establish the most efficient method of utilizing economic
resources to meet social ends. After all, it is the marginalists who assure
us that they can provide optimal answers to such questions as these: How
are resources to be organized to produce what is socially beneficial,
what economic policies will bring about an optimal allocation of factors
of production, and how are social needs to be best expressed so as to
bring production into conformity with these needs?

In order to reach their conclusion that market price produced by
state-equilibrium competition will provide the best answer to these
questions, the marginalists must assume that market prices will reflect
all costs of production. This assumption is defective on two grounds.

The first defect in assuming that costs internal to the firm are, in
fact, true marginal costs of production stems from the existence of
significant costs of production and of consumption which are not internal to the private sector. Pollution, both physical and aesthetic, is an example of a cost of producing and of consuming output which is not generally reflected in private prices, even if private market prices are set at internal marginal costs. Litter, contaminated streams, flashing billboards, unhealthy air, and so forth, are all social costs which are external to prices set according to the marginal cost of the firm. Hence, absent some mechanism to adjust competitive prices to account for these externalities, marginal cost pricing cannot produce economic efficiency. Even if such adjustment calculations were possible, the Pareto optimum conditions could not reflect true social marginal costs of production in that they are the very basis of defining internal marginal costs of externalities. Private production and its associated marginal costs are a function of the existing income and wealth distribution. On the demand side of the market, social marginal costs of externalities resulting from private production are defined by their egalitarian impact on society as a whole. Hence, in the case of externalities, the basis for defining social marginal costs differs and is fundamentally inconsistent with the definition of social marginal costs reflected in the prices of private production.

Modern theorists mention the necessity to include social marginal cost pricing as an integral part of normative conclusions concerning economic efficiency; however, they have not developed an adequate mechanism to reconcile private costs associated with outputs determined by private market forces with social marginal costs, the amount of which must be determined according to nonmarket standards. As pointed out above in another context, this inconsistency renders the Pareto optimum invalid.

The second defect derived from the need to reconcile private costs (internal to the firm) with full marginal cost pricing relates to the fact that a new, and historically increasing part of the costs of producing goods and services is not reflected in market prices. This is because federal, state, and local governments subsidize private costs. Comparison of operating conditions for a typical firm today with those confronting a typical mid-nineteenth-century firm highlights the point. In the earlier period, a firm paid full market interest rates on new plant and equipment, absorbed all the cost of developing sites for new plant and equipment, could expect no substantial tax write-offs, or government subsidies for research and development, and so on. Today, by comparison, local high school curricula generally reflect the needs of local industry, state universities are heavily subsidized by state and federal governments, state and federal tax and interest subsidies stimu-

late new investment, state and local governments heavily subsidize industrial and commercial parks, labor training programs are government supported, and a myriad of additional government programs alters the private costs of producing goods and services in such a way that the true costs of producing output are not fully reflected in market prices.

Whereas it could be argued that the internal marginal costs of the nineteenth-century firm tended to reflect most of the monetized costs of producing output, the gradual "socializing" of private costs has shifted a large (and increasing) portion of total monetized costs of production into the public sector. It is exactly this phenomenon which occurs when one state or municipality competes against another for new business and industry. Each government competitor is attempting to socialize a greater portion of private costs in order to attract and retain investment. These influences create a sharp divergence between private internal marginal costs and full social marginal costs. This divergence is even more difficult to trace once the input-output paths of production are included in the analysis. To price output on a consistent basis and include full marginal costs of production in prices, one should include not only internal costs to private production, but also (throughout the input-output system of production) nonmonetized social marginal costs external to the firm and monetized costs of production paid by the public sector.

The two defects in the marginal cost model highlighted above exemplify the impossibility of using monetized costs internal to the firm as the basis for pricing. Even if the marginalists are correct in their conclusions that prices should be set equal to social marginal costs, the normative inconsistency between social marginal cost derived from the public and the private sectors points out that the quest for "true" social marginal cost is a quest for the unknowable. Hence, the marginal cost model is either inconsistent (when expanded to include these phenomena) or incomplete (when social marginal costs are excluded from the model.)

The marginal cost model is incomplete and may be inconsistent when technological change is added to the model.

The conclusions which flow from the marginalist thinking support a finding that economic efficiency is promoted by the existence of small firms. This reasoning flows from the conclusion in the marginalist model that bigness in markets thwarts competition and produces deviations of price from marginal cost. Hence, based on static equilibrium comparisons of alternative market structures, the marginalists conclude that more output to satisfy social needs can be obtained from a
The conclusion: Pricing according to marginal costs produced by pure and perfect competition may violate the very purpose of the static Pareto model.

DEFECTS IN ASSUMPTIONS FATAL TO THE MARGINAL COST MODEL. Having explored internal defects of the model, this section addresses the invalidity of the marginal cost model resulting from having asked the wrong central question. The policy conclusions which follow from traditional theory are dependent upon the capability of that theory to describe adequately the essential workings of the U.S. economy, that is, the normative worth of the theory is a direct function of the positive analysis. If it can be shown that assumptions concerning the workings of reality which are embedded in the traditional theory fail to describe the real world, then the theory would be rendered useless for policy purposes in rate design for public utilities. It is on such empirical grounds that marginal cost theory fails most dramatically.

Developed in the mid- and late nineteenth century, at a time when government played a very small role in the economy and when firms in most markets were competitive and small scale as compared with those of today, traditional microeconomic theory described the operations of the existing economy to an extent which many observers believed sufficiently accurate to justify the use of the theory for public policy purposes. This claim cannot be made today in light of several conditions. (Note that sections below indicating empirical flaws in traditional theory parallel sections above covering theoretical defects.)

Stability conditions in the theory are constrained by reality.

Although the market system apparently is inherently stable, this real world condition does not necessarily stem from the causes marginal cost theorists cite. To test whether or not the marginalist model explains why the market system is stable, an inspection of the real world is necessary to compare model assumptions with actual events. Suffice it to say that the conditions spelled out in the marginal cost model do not necessarily for a stable equilibrium are not prevalent in the U.S. economy. For example, even short-run costs of firms tend to be fluid and not concave; long-run costs tend to decline, reflecting secular improvements in productivity; and the wage rate does not go to zero when massive unemployment exists (such as in 1975). The plain fact is that the "reality" predicted and necessary for a Pareto optimum has little or no relation to the reality in which we live. In this light, to discuss stability conditions of the marginal cost model further would serve little purpose.

The government sector is inseparable from the real world economy, and yet it cannot be handled by the marginal cost model.

In order to be consistent on theoretical grounds, marginal cost theory has no government sector. If that sector were insignificant in the real world economy, the lack of a government sector in the pure theory would not be damaging to the use of marginal cost pricing for public utility rate determination. However, this is not the case; in 1976, federal, state, and local governments produced final output equal to
57.1 percent of all goods and services produced in the U.S. economy. As can be seen in Table 1, this proportion has expanded steadily during the twentieth century, and the trend can be expected to continue. Hence, traditional theory fails to describe the operation of an important and growing portion of the phenomena about which policy claims are made from the theory.

In addition to the huge and growing size of the government sector represented by goods and services devoted to public purpose, virtually all economic activity is significantly affected by state and federal regulation.

It is, therefore, inconceivable that market prices in the U.S. economy can result which replicate those which would be produced by pure and perfect competition. To gauge the extent to which government regulation is a factor in the marketplace, examine Figure 1, which lists some of the more important regulatory agencies at the federal and state level. A review of this figure leaves little doubt that policies aimed at promotion of Pareto optimality would, by definition, mean abolishing these agencies. In any event, the marginalists are on weak ground when they argue that competition of the sort envisioned in their model exists in the U.S. economy.

Not least in affecting the practical validity of the marginal cost model are government penalties and subsidies. Every time a government agency either subsidizes, penalizes, or controls the (1) level of output, (2) price of output, and/or (3) quality of output of an industry, the result is that Pareto optimum conditions have been violated. Federal subsidies in 1975 were $95.1 billion according to a conservative estimate of their magnitude. Since 1968, $55 billion has been expended by the federal government for manpower training programs. The Office of Policy and Management this year published a catalogue of 164 separate U.S. loan guarantees used to subsidize capital formation. Fossil fuel and other energy programs have been built upon huge and direct tax subsidies, and so on.

To assert, as the marginalists do, that the economy operates on prices which reflect social marginal costs and, then, to argue the uselessness of these prices as "signals" for determining a unique social optimum becomes increasingly unrealistic as one thinks of the reality in which government plays so large a role.

Monopoly, oligopoly, and labor unions are the reality of the U.S. economy, not the conditions necessary for the traditional theory to produce useful policy.

In order to be consistent on theoretical grounds, it is assumed in marginal cost analysis that all firms and all markets for resources in the U.S. economy are perfectly competitive. Translated into real world

<table>
<thead>
<tr>
<th>Year</th>
<th>Government Expenditures as a Percentage of Gross National Product</th>
<th>Subtotal</th>
<th>State and local</th>
<th>Federal</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1929</td>
<td>5.3</td>
<td>5.3</td>
<td>0.0</td>
<td>5.3</td>
<td>5.3</td>
</tr>
<tr>
<td>1930</td>
<td>6.1</td>
<td>6.1</td>
<td>0.0</td>
<td>6.1</td>
<td>6.1</td>
</tr>
<tr>
<td>1931</td>
<td>6.9</td>
<td>6.9</td>
<td>0.0</td>
<td>6.9</td>
<td>6.9</td>
</tr>
<tr>
<td>1932</td>
<td>7.4</td>
<td>7.4</td>
<td>0.0</td>
<td>7.4</td>
<td>7.4</td>
</tr>
<tr>
<td>1933</td>
<td>7.9</td>
<td>7.9</td>
<td>0.0</td>
<td>7.9</td>
<td>7.9</td>
</tr>
<tr>
<td>1934</td>
<td>8.4</td>
<td>8.4</td>
<td>0.0</td>
<td>8.4</td>
<td>8.4</td>
</tr>
<tr>
<td>1935</td>
<td>8.9</td>
<td>8.9</td>
<td>0.0</td>
<td>8.9</td>
<td>8.9</td>
</tr>
<tr>
<td>1936</td>
<td>9.4</td>
<td>9.4</td>
<td>0.0</td>
<td>9.4</td>
<td>9.4</td>
</tr>
<tr>
<td>1937</td>
<td>9.9</td>
<td>9.9</td>
<td>0.0</td>
<td>9.9</td>
<td>9.9</td>
</tr>
<tr>
<td>1938</td>
<td>10.4</td>
<td>10.4</td>
<td>0.0</td>
<td>10.4</td>
<td>10.4</td>
</tr>
<tr>
<td>1939</td>
<td>10.9</td>
<td>10.9</td>
<td>0.0</td>
<td>10.9</td>
<td>10.9</td>
</tr>
<tr>
<td>1940</td>
<td>11.4</td>
<td>11.4</td>
<td>0.0</td>
<td>11.4</td>
<td>11.4</td>
</tr>
<tr>
<td>1941</td>
<td>11.9</td>
<td>11.9</td>
<td>0.0</td>
<td>11.9</td>
<td>11.9</td>
</tr>
<tr>
<td>1942</td>
<td>12.4</td>
<td>12.4</td>
<td>0.0</td>
<td>12.4</td>
<td>12.4</td>
</tr>
<tr>
<td>1943</td>
<td>12.9</td>
<td>12.9</td>
<td>0.0</td>
<td>12.9</td>
<td>12.9</td>
</tr>
<tr>
<td>1944</td>
<td>13.4</td>
<td>13.4</td>
<td>0.0</td>
<td>13.4</td>
<td>13.4</td>
</tr>
<tr>
<td>1945</td>
<td>13.9</td>
<td>13.9</td>
<td>0.0</td>
<td>13.9</td>
<td>13.9</td>
</tr>
<tr>
<td>1946</td>
<td>14.4</td>
<td>14.4</td>
<td>0.0</td>
<td>14.4</td>
<td>14.4</td>
</tr>
<tr>
<td>1947</td>
<td>14.9</td>
<td>14.9</td>
<td>0.0</td>
<td>14.9</td>
<td>14.9</td>
</tr>
<tr>
<td>1948</td>
<td>15.4</td>
<td>15.4</td>
<td>0.0</td>
<td>15.4</td>
<td>15.4</td>
</tr>
<tr>
<td>1949</td>
<td>15.9</td>
<td>15.9</td>
<td>0.0</td>
<td>15.9</td>
<td>15.9</td>
</tr>
</tbody>
</table>

terms, this means that there are no unions, no large firms which dominate markets, and no serious imperfections in resource markets. In essence, this assumption requires that the very fabric of the U.S. economy be ignored when using the marginal cost standard for utility rate design. The fact is that, by 1954, 500 firms accounted for over 50

<table>
<thead>
<tr>
<th>Figure 1. Agencies in Federal and State Government Which Have Regulatory Responsibilities Affecting the Private Market</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FEDERAL AGENCIES</strong></td>
</tr>
<tr>
<td>Power — 14</td>
</tr>
<tr>
<td>Alaska Power Administration</td>
</tr>
<tr>
<td>Bonneville Power Administration</td>
</tr>
<tr>
<td>Defense Nuclear Power Administration</td>
</tr>
<tr>
<td>Energy Research and Development Administration</td>
</tr>
<tr>
<td>Tennessee Valley Authority</td>
</tr>
<tr>
<td>Southeastern Power Administration</td>
</tr>
<tr>
<td>Southwestern Power Administration</td>
</tr>
<tr>
<td>Federal Energy Administration</td>
</tr>
<tr>
<td>Energy and Technical Development — 9</td>
</tr>
<tr>
<td>Bureau of Economic Development</td>
</tr>
<tr>
<td>Federal Advisory Council on Regional Economic Development</td>
</tr>
<tr>
<td>Federal Council for Science and Technology</td>
</tr>
<tr>
<td>National Science Foundation</td>
</tr>
<tr>
<td>Institute for Applied Technology</td>
</tr>
<tr>
<td>Office of Technology Assessment</td>
</tr>
<tr>
<td>Office of Experimental Technology</td>
</tr>
<tr>
<td>Office of Water Resources and Technology</td>
</tr>
<tr>
<td>Economic Development Administration</td>
</tr>
<tr>
<td>Commerce — 9</td>
</tr>
<tr>
<td>Bureau of Domestic Commerce</td>
</tr>
<tr>
<td>Bureau of Economic and Business Affairs</td>
</tr>
<tr>
<td>Office of Industrial Economics</td>
</tr>
<tr>
<td>Securities and Exchange Commission</td>
</tr>
<tr>
<td>Interstate Commerce Commission</td>
</tr>
<tr>
<td>Bureau of Resources and Trade Assistance</td>
</tr>
<tr>
<td>Small Business Administration</td>
</tr>
<tr>
<td>Federal Trade Commission</td>
</tr>
<tr>
<td>Transportation — 5</td>
</tr>
<tr>
<td>Civil Aeronautics Board</td>
</tr>
<tr>
<td>Federal Aviation Administration</td>
</tr>
<tr>
<td>Federal Highway Administration</td>
</tr>
<tr>
<td>Federal Railroad Administration</td>
</tr>
<tr>
<td>Urban Mass Transit Administration</td>
</tr>
<tr>
<td>Consumer — 5</td>
</tr>
<tr>
<td>Office of Consumer Affairs</td>
</tr>
<tr>
<td>Consumer Product Safety Commission</td>
</tr>
<tr>
<td>Office of Product Standards</td>
</tr>
<tr>
<td>Bureau of Quality Assurance</td>
</tr>
<tr>
<td>Federal Food and Drug Administration</td>
</tr>
<tr>
<td>Conservation — 6</td>
</tr>
<tr>
<td>Environmental Protection Agency</td>
</tr>
<tr>
<td>Office of Land Use and Planning</td>
</tr>
<tr>
<td>National Commission on Water Quality</td>
</tr>
<tr>
<td>Ocean Mining Administration</td>
</tr>
<tr>
<td>Bureau of Mines</td>
</tr>
<tr>
<td>Agricultural Stabilization and Conservation Service</td>
</tr>
<tr>
<td>Other — 7</td>
</tr>
<tr>
<td>N.A.S.A.</td>
</tr>
<tr>
<td>Federal Communications Commission</td>
</tr>
<tr>
<td>Equal Employment Opportunity Commission</td>
</tr>
<tr>
<td>Federal Maritime Commission</td>
</tr>
<tr>
<td>Health Resources Administration</td>
</tr>
<tr>
<td>Health Services Administration</td>
</tr>
<tr>
<td>Federal Housing and Urban Development Administration</td>
</tr>
</tbody>
</table>

| Figure 2 — continued |

<table>
<thead>
<tr>
<th>STATE AGENCIES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of states (including District of Columbia)</td>
</tr>
<tr>
<td>Agency</td>
</tr>
<tr>
<td>Agriculture</td>
</tr>
<tr>
<td>Aeronautics</td>
</tr>
<tr>
<td>Alaska Power Administration</td>
</tr>
<tr>
<td>Alaska Railroad Commission</td>
</tr>
<tr>
<td>American Arbitration Association</td>
</tr>
<tr>
<td>American Bar Association</td>
</tr>
<tr>
<td>American Medical Association</td>
</tr>
<tr>
<td>American Public Relations Association</td>
</tr>
<tr>
<td>American Telephone and Telegraph Company</td>
</tr>
<tr>
<td>Anheuser-Busch, Inc.</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Newark, New Jersey)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (St. Louis, Missouri)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Pittsburgh, Pennsylvania)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (San Francisco, California)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Washington, D.C.)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Cleveland, Ohio)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Boston, Massachusetts)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Indianapolis, Indiana)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Chicago, Illinois)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Minneapolis, Minnesota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Atlanta, Georgia)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (St. Louis, Missouri)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Kansas City, Missouri)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Columbus, Ohio)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Cincinnati, Ohio)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Hilton Head, South Carolina)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Charlotte, North Carolina)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Raleigh, North Carolina)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Greensboro, North Carolina)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Winston-Salem, North Carolina)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Durham, North Carolina)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Richmond, Virginia)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Roanoke, Virginia)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Charlottesville, Virginia)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Alexandria, Virginia)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Washington, D.C.)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Huntsville, Alabama)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Birmingham, Alabama)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Memphis, Tennessee)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Nashville, Tennessee)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Knoxville, Tennessee)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Jacksonville, Florida)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Tampa, Florida)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Miami, Florida)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Jackson, Mississippi)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Sportsman’s Museum, New Orleans)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Caddo Parish, Louisiana)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Beaumont, Texas)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Houston, Texas)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Dallas, Texas)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (San Antonio, Texas)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (El Paso, Texas)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Brownsville, Texas)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Laredo, Texas)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (McAllen, Texas)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Lubbock, Texas)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Abilene, Texas)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (San Angelo, Texas)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Midland, Texas)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Odessa, Texas)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Carlsbad, New Mexico)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Roswell, New Mexico)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Albuquerque, New Mexico)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Las Cruces, New Mexico)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Durango, Colorado)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Alamosa, Colorado)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Colorado Springs, Colorado)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Denver, Colorado)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Greeley, Colorado)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Boulder, Colorado)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Loveland, Colorado)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Pueblo, Colorado)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Rapid City, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Sioux Falls, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Yankton, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Aberdeen, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Deadwood, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Fort Pierre, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Mitchell, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Sioux City, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Sioux Falls, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Yankton, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Aberdeen, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Deadwood, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Fort Pierre, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Mitchell, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Sioux City, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Sioux Falls, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Yankton, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Aberdeen, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Deadwood, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Fort Pierre, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Mitchell, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Sioux City, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Sioux Falls, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Yankton, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Aberdeen, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Deadwood, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Fort Pierre, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Mitchell, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Sioux City, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Sioux Falls, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Yankton, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Aberdeen, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Deadwood, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Fort Pierre, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Mitchell, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Sioux City, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Sioux Falls, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Yankton, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Aberdeen, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Deadwood, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Fort Pierre, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Mitchell, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Sioux City, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Sioux Falls, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Yankton, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Aberdeen, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Deadwood, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Fort Pierre, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Mitchell, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Sioux City, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Sioux Falls, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Yankton, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Aberdeen, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Deadwood, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Fort Pierre, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Mitchell, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Sioux City, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Sioux Falls, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Yankton, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Aberdeen, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Deadwood, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Fort Pierre, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Mitchell, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Sioux City, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Sioux Falls, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Yankton, South Dakota)</td>
</tr>
<tr>
<td>Anderson, Clayton Corporation (Aberdeen, South Dakota)</td>
</tr>
</tbody>
</table>
cost for even one of the many firms operating in the U.S. economy; therefore, this theory is silent as to whether the existing price of any firm is above, below, or equal to the optimal price level. Hence, unlike imperfections in Newtonian physics, where theory can give us useful information as to order of magnitude and direction of change, marginal cost theory is virtually useless as a guide to solve real world problems precisely because even the direction of desired change is an unknown.

It is interesting that, although they entered the great debate on rate reform from an apparent position of strength, the advocates of marginal cost pricing are now in a defensive stance. As the practical usefulness of marginal cost pricing has come to be seen at best as "a fifth wheel attached onto peak load pricing," and as more fundamental criticisms have shown the invalidity of the theory itself, advocates have adopted the defense that "second best" considerations are still relevant. Furthermore, they claim that it is incumbent upon those who criticize marginal cost theory to show that firms in the U.S. economy are not pricing at levels which represent a Pareto optimum condition.14

Although it may not be apparent to these theorists, this line of reasoning is tantamount to a complete abandonment of the marginalist position because it throws all theories of pricing onto the same ground. What these theorists are admitting is what they do not know — whether prices are set at levels which constitute a socially desirable optimum — nor do their critics know, they say. This is similar to medieval scholars asserting that an infinite number of angels can dance on the head of a pin and, when asked to prove it, saying that it is up to the critics to prove that there are less than an infinite number of angels dancing on pinheads. The problem which is apparent is that the central question posed by the marginalists in the 1870s has led to more than a century of searching for the unknowable. As do their utopian socialist counterparts, who would have us organize society into communes, these utopian market theorists would have us organize society based upon their own particular idea fixe — small competitive firms controlled by market forces so as to price at marginal cost.

Today's marginal cost theorists, by defensively asking this new question of their critics (that is, prove that marginal costs are not the basis for prices in the U.S. economy) have done a great service to us all. We are indebted to these theorists because, by asking a new question, they have shifted the ground from pin dancing to reality. Once on that ground, the pragmatic weighing of the marginalist position rises or falls alongside all alternative proposals, dependent upon whether or not it works.

Where this leaves us is not as bad as it seems. After all, society did not rise or fall because monks and the theorists of the Dark Ages were posing questions about angels dancing on the heads of pins. Instead, society found pragmatic, workable solutions to real world problems exactly because the important questions presented themselves directly from the nature of those problems. This brings us to the second school of thought, the pragmatic approach to setting public utility prices born of the classical tradition.

The Classical Model of Economic Efficiency and the New Wave Theory

The classical model of economic efficiency, unlike its historical successor, is more humble in its design. Whereas the neoclassicists make sweeping claims as to the precise level at which all firms must price output at any given time, classical thinkers simply state that the forces of competition will produce increased economic efficiency over time. The former model is based upon the assumption that there is a knowable, instantaneous, general equilibrium price solution (leading to a Pareto optimum). The latter model depends upon sequences of market interaction to produce a material summum bonum for society. This important difference is highlighted by the methodologies employed by these schools of thought.

The neoclassical model isolates social and political factors from consideration and asserts that the economy is, and should be, separate from other influences in society. In contrast, the classical model is systems oriented and views social, political, and economic interaction as the basis for sequences of economic development. Whereas the neoclassical model is built upon a box-of-tools methodology and is dominated by what is known as static equilibrium analysis, the classical model is oriented toward outlining the successive phases of development of the market system;14 hence, it employs dynamic equilibrium analysis as its principal technique.

What is of importance in the discussion here is that the classical view of economic efficiency depends upon a concept of forces working toward an end; it depicts a "circular cumulative causation" of interrelated decision makers.17 They, in their economic roles, weave an imprecise path of development which cannot be specified except as to its direction and as to the general order of magnitude of pieces which make up the economic pie.

It is important to see that the classical tradition does not make grandiose claims about an ideal or optimum pricing solution; rather, it is premised on the muddling through, trial-and-error view of move-
ment toward increasing efficiency of production. Economic pricing thus results from competition, whereby the market price is driven down to a level reflecting average efficiency. Businessmen in the competitive sector are viewed as being constrained in their behavior such that they must innovate and adopt the most efficient methods of production in order to survive. Stating the same thing in a different manner, the market system provides an opportunity for the most efficient competitor to make excessive profits in the short run. Eventually, other competitors emulate the most efficient producer, and the resulting price decline wipes out short-run profits. This process is iterative and sequential; it is based on pragmatic “rounds” of adopting best-practice technology and is aimed at producing more output for less cost. It is blind to, and unconcerned with, any concept of optimality. It depends upon a pragmatic search among available alternatives for workable solutions to output and pricing decisions. This lack of precision has its virtues.

The classical model does not ask “precise” questions which lead down a path of logical elegance toward the unknowable. This pitfall is avoided not only because the reality which the model seeks to represent defies precise definition, but also because much of the intent of the classical school is to question the general direction in which society is going. Hence, we find no questions as to the “precise” level of marginal cost as produced by competition, exactly at one or another level. Rather, the classical model incorporates the general expectation that competition will operate so as to promote technological advance. Nor do we find the precise definition of optimum consumption; instead, in the classical model there is merely the presumption that “more is better” and that high quality is a desideratum. The key is the pragmatic, means-ends approach of classical economic thought; for it is here, coupled with the interactive systems approach, that the new wave theorists are in concert with, and derive inspiration from, their historically distant antecedents.

The New Wave and Efficiency Standards

The traditional theory supporting “competitive standard” regulation is devoid of real content. It is defunct. With the demise of the fiction that Pareto optimum pricing can be relied upon to provide for the best use of society’s resources, so too dies passive regulation as a rationale for setting economic price.

Interventionist Regulation

In the place of passive regulation, there is no choice but to accept active, interventionist regulation as the only means to promote economic efficiency. The logic of this position goes as follows. (1) The use of the competitive standard for regulatory price has not been, and cannot be, proven valid as a method for promoting economic efficiency; (2) regulators are, de facto, interventionists since they do set economic price and, in so doing, have a profound impact upon the path of economic and social development; (3) the effects of setting regulatory price can be measured to some extent, and alternative costs and benefits can be estimated; therefore (4) explicit efficiency standards by which to gauge the costs and benefits of alternative regulatory actions are required to produce a fair and reasonable “end result.”

The point is that, with the death knell of the neoclassical model, there exists no other choice but to regulate according to explicit standards of economic efficiency. Whether it is recognized or not, regulators are already interventionists, and by establishing economic price in the regulatory process, they select the shape of the future society. The plain fact is that regulatory choice made to define the form of public utility price is, at one and the same time, a choice among alternative future technologies. For example, whether or not electric utility companies are encouraged to promote conservation is, in fact, a choice whether or not to promote “soft” energy development. In this connection, whether or not electric rate orders contain rates to promote cogenration, waste recycling, and off-peak energy use constitutes policy choice which will constrain future opportunities available to our society. Certainly, we are not unaware of the fact that a decision which supports the expansion of “hard” electric technology (which by definition is centralized in its production) is also a decision supporting less centralized social organizations. At least in this area, we know from experience that it is “hard” electric technology which makes possible a geographically dispersed population in the suburbs.

The sum and substance of this line of reasoning is that, in a social system inseparable from the interrelated technologies upon which it depends, not only are regulators making a constitutional choice to balance the short-run interests of consumers, investors, and the public, but also they have a determining impact on the future balance of those interests by promoting or restraining alternative technological forms. Given that, de facto, regulators significantly shape and form both the future and present development of society, and given that the forces of competition cannot be presumed to produce a socially optimum distribution of resources and technology, then the only conclusion which follows is that goals for regulation must be addressed which make explicit regulation of the effects of ratemaking for electric utilities. The question of whether or not to employ such goals in electric rate design is
Integrating Efficiency Standards

no longer relevant; rather, the appropriate question to be resolved in the regulatory process is which particular balancing of goals will produce a result which is fair and reasonable in both the short and the long run.

Future Controversies

While the great debate on rate reform marks the end of the era of passive regulation, it is the beginning of active regulation based on explicit efficiency standards. The ground has now shifted to a new search for workable models to assist in setting such standards. This also creates the context for an extension of the great debate far into the future. The shape of this debate will, this author suggests, take the form which presently exists — two general camps, retooled neo-classicists and pragmatists using the classical approach. The shift to regulation by efficiency standards has been long in the making, yet, it has received considerable stimulus over the last five years due to the influence of the energy crisis, eroding household income, and the general push for better government performance. The signposts of this change are already there, if we care to heed them, and, as observed above, they break into the two camps. On the side of the neo-classicists is the recent work of the National Bureau of Standards, the Federal Energy Administration, and certain consultants using productivity models based upon marginalist precepts of economic efficiency. (Of course, the marginalist price theorists, including those at EPRI, are to be counted as a major force within this school.)

In the new wave are the various “movement” theorists — ELCON, AIRCO, GM, the conservationalists, environmentalists, consumer advocates, and so on — all of whom base their arguments for economic efficiency on some concept of what will work best from among the competing alternatives which are known and which can be implemented. Notably, all of these proponents of the new wave view the economic and society as an interrelated system; they all look to promoting dynamic sequences of events which are subject to a calculation of results according to simple means-ends tests.

Finally, there is already institutional support for the new wave. Legislatures and chief executives at the federal and state level have established policies which are aimed at regulation by efficiency standards. For example, Connecticut was one of the first states to mandate efficiency as an explicit criterion for setting rates. Many more states require management audits as part of the regulatory process. Governor Richard A. Snelling, in a speech before the New England Conference of NARUC and in Public Utilities Fortnightly, has advocated rewarding and penalizing utility management according to efficiency performance. Not least, at President Carter’s insistence, the Federal Energy Regulatory Commission has instituted efficiency based regulation of the Alaska gas system. Hearings are presently under way to determine a procedure to reward the gas pipeline builders with an above-normal rate of return on equity if the project is completed at less than budget cost, and to penalize the builders with a lower return if they have cost overruns.

Summary and Conclusion

It is true that, in the private competitive market, pressures which enforce efficiency standards are automatic and implicit. Variations in the rate of profit are the tangible result of variations in operating efficiency achieved by management. This conforms with the classical view of things and not that of marginal cost theorists.

In the regulated sector of the economy, the very forces which promote efficiency in the market sector are kept at bay by legislative design. The economic cost of state-enfranchised monopoly may be significant inefficiencies — both in operations and in the construction choices of public utility management. In addition to the potential for inefficiencies internal to the operations of utility companies, due to the interrelatedness of the modern industrial economy, decisions concerning economic price (and therefore future investment) can have the effect of creating inefficiencies in the economy as a whole.

The conclusion which follows from the analysis herein is that regulation must be fashioned to penetrate traditional management prerogatives, and in the place of management autonomy, pressures of regulation must be structured so as to reward economic efficiency and penalize inefficiency. Because there does not exist a “competitive standard” intrinsic to the market system itself, and because regulators do have a strong impact upon the shape of our future economic and social development, there is no alternative but to construct such explicit efficiency standards — based on a pragmatic, mean-ends calculus — and to use them as the basis for establishing regulated economic price. This is the new wave which will shape the future of regulation in the United States.

Notes

1. Note, for example, the fact that the research arm of NARUC, the National Regulatory Research Institute, was formed only within the last two years. True, the Institute of Public Utilities was formed in 1965 at Michigan State University — to the very great credit of one man, Harry Trebing. The
point is, however, that ongoing work on the development of regulatory
techniques has been restricted to a few isolated areas until recently.

2. The ideas of economists and political philosophers, both when they are
right and when they are wrong, are more powerful than is commonly
understood. Indeed, the world is ruled by little else. Practical men, who
believe themselves to be quite exempt from any intellectual influence, are,
usually the slaves of some defunct economist. Madmen in authority, who
hear voices in the air, are distilling their frenzy from some academic
scribbler of a few years back. I am sure that the power of vested interests is
vastly exaggerated compared with the gradual encroachment of ideas. J.
M. Keynes, General Theory of Employment, Interest, and Money (New York:
Macmillan, 1936).

3. EPRI, Rate Design and Load Control (Palo Alto, Calif.: the Institute,

discussion of the distinction drawn between cost of service based upon the
Higher Federal Natural Gas case (which best represents the competitive
and "end result") and that of Permian Basin (representing the legal basis for
the new wave), see Testimony of the author before the Colorado Public

5. We are reminded, for example, of the well-known "object" of the investiga-
tions of the scholastics. Having engaged in a search for the meaning of
infinite grace, these theorists were reduced to asking such important
subordinate questions as how many angels can dance on the head of a pin,
and so on, in an endless search for the unknowable.

6. EPRI, Rate Design and Load Control, p. 20. The fact that the particular
cost concept of economic efficiency referred to in the EPRI study depends
upon the Pareto optimum is clear from the quotation. That economic
efficiency and the Pareto optimum are inextricably linked explains why
Charles Cichetti, Irwin Steltzer, Alfred Kahn, and others must support
their analysis with assumptions about "all other relevant goods and
services in the economy are priced at marginal cost" (Steltzer), and "if all
the goods and services provided by a society are produced and supplied
according to marginal cost...then social welfare and economic efficiency
will be maximized" (Cichetti). Quotations are from Irwin M. Steltzer,
Testimony in Generic Hearings Concerning the Rate Structure of Electric Utilities,
Case No. 3693, State of Colorado, 5 August 1977, p. 12; and Charles J.
Cichetti and W. J. Gillen, "Time of Electricity Pricing: Correcting Some Con-
tinuing Confusion," paper given before the Annual Convention of NARUC,

7. The restrictive assumptions necessary to maintain a general equilibrium
include, among others, that (1) the production function be such that
Euler's adding up theories produce a balance between factor payments and
profits, on the one hand, and product market demand, on the other;
(2) the production functions remain the same for each factor owner regardless of the product "mix" (necessary to
maintain the same level of demand for each type of output in equilibrium);
(3) Government demand (i.e., Arrow), to name a few points of reality ignored;
(4) all firms are sufficiently small so that no one firm can set market price
levels; (5) there is no unemployment and no unions; and so forth, and so
on. The full extent of these restrictive assumptions can be gauged by the
fact that they take up 15 pages of single-spaced explanation in an inter-
mediate economics textbook. See George Malanos, Intermediate Economics

The import of these restrictive conditions is that the very existence of a
Pareto optimum is dependent upon all firms producing at an output level
at which long-run marginal cost (LRMC) equals long-run marginal reve-
ne (LRMR) equals short-run marginal revenue (SRMR), which is to say
that all firms produce under technological conditions wherein either
long-run costs are always at the same level over all possible levels of output
or that long-run costs first decline and then increase as output expands.
Under the former technological assumption (constant long-run costs)
there can be no productivity increase in the economy over time; under the
latter assumption concerning technology, there can be no growth in the
economy except by departing from the Pareto equilibrium into an unstable
circularity of causation.

8. For example, Malanos points out that the insights of the Keynesian revolu-
tion lead to a conclusion that the price of labor should go to zero if the
marginalists are right about how the market operates. See Malanos, Inter-

9. It is interesting that Kenneth Arrow's impossibility theorem, long the bête
noire of the marginalists, focuses upon one aspect of this source of inconsis-
tency of the neoclassical model. As it is well known, no theorist has been able to
challenge Arrow's finding successfully; hence, the marginalists are
without a theory of demand to establish economic price.

10. It should be noted that the empirical evidence on the question strongly
suggests that small competitive firms which cannot substantially affect
their market price are the least innovative. For the most recent review of
the literature on this subject, see O. E. Williamson, Markets and Hierarchies:

As pointed out by Joan Robinson, in "What Are the Questions?" Journal of
Economic Literature 16 (Winter 1978): 1519-20, it is probably impossible
ever to test the U-shaped long-run cost curve due to the confounding of
technological change with pure scale effects.


12. Ibid.


15. For example, see Kahn's statement on page 152 of the EPRI study and also
the statement on page 8 of Cichetti and Gillen, "Time of Day Electricity
Pricing." In both instances, the retreat is complete. Cichetti and Gillen are
more forthright about this than Kahn. They state: "If marginal cost is not
the basis of pricing something else must take its place. We do not know of
any obvious approach." (p. 28). Cichetti and Gillen even suggest that a
pragmatic approach to time-of-day pricing is compatible with their view of
marginal cost pricing.
16. Note, for example, that virtually all the major classical theorists have a concept of natural phases of economic development. Adam Smith's vent-for-surplus phase, the gloomy prospects depicted by both Thomas Malthus and Karl Marx, and the ultimate equilibrium phase feared by David Ricardo when economic surplus would be captured by the landed rentier class.

17. This is Gunnar Myrdal's well-known phrase.


---

**Incentive Regulation**

*Daniel J. Demlow*

Beginning in the late 1960s, increasing inflation and decreasing utility productivity, among other factors, caused the cost of utility services to begin to rise steadily as contrasted with previous trends of diminishing costs per unit of output. One result has been that customers have begun to question the propriety and necessity for the various costs incurred by utilities. Intervention in formal rate proceedings has multiplied, and even larger industrial customers, well aware of the significant cost pressures, have begun to question the wisdom of varying utility cost decisions.

The traditional ratemaking process has been scrutinized by both the public and regulators. In many instances, it has been found lacking. Briefly and incompletely summarized, regulatory decisions were made on the basis of historical data, which resulted in many decisions that were out of date when made. The decision-making formula was a type of cost-plus system: A utility spent whatever it wanted and submitted the bill to its customers. Regulators were put in the position of examining management decisions several years after the fact and trying to decide whether they were “reasonable and prudent” at the time they were made. These problems were greatly compounded in states such as
Michigan, where regulated utilities were burdened with multibillion-dollar construction programs.

Regulators are presented with voluminous evidence in a formal hearing and are held publicly accountable, yet they have no effective means of influencing utility decision making in a timely manner. This system of regulation, with its untimely review of costs, introduces an unacceptable and unnecessary risk into the decision-making process which troubles Wall Street, ill serves the public, and greatly limits the realistic options open to the regulator. In response to these shortcomings, Michigan has embarked upon a new regulatory approach which emphasizes incentives in public utility regulation.

Incentive regulation, as practiced in Michigan, is intended to ensure that utility policies reflect the public’s interest in minimizing the costs of operating a utility. Realistic targets for performance established in public hearings are the basic vehicle for implementing this approach. In addition to addressing the public’s concern for cost minimization, incentive regulation can provide wide scope for managerial initiative by allowing management even more freedom to manage. In a system of incentive regulation, the utility commission sets general performance targets, and management is then free to meet these targets and to receive the credit for success or accept the responsibility for failure. No interest is served by restricting the flexibility and operating latitude of utility managers.

The most important task for incentive regulation is to determine those measures of performance associated with and affecting the cost of doing business. Numerical targets must be set, and tolerable degrees of deviation from them must be established.

The system generally consists of more than one incentive provision. Each principal measure of performance, such as equipment availability or fuel expense, may be addressed through separate provisions. To avoid interval inconsistency, it is essential that the various provisions mesh so as to achieve rational resource use. Adopting two provisions, for example, one of which rewards management for reducing investment cost of new generating plant and another which penalizes management for building generating plant which is susceptible to failure, would make very little sense. Given a greater reward for good effort in the first instance than penalty for poor effort in the second instance, management would literally be encouraged to build plants which did not work.

The same need to assure rational results is also applicable to the structuring of individual incentive provisions. Absent suitable numerical yardsticks and tolerances, an incentive provision could have serious counter-productive consequences.

The determination of appropriate goals is difficult, but without quantification of targets, utility management lacks the explicit guidance necessary to enable it to achieve the objectives of incentive regulation.

General Criteria for Incentive Provisions

What are some of the specific considerations entering into the development of incentive provisions? Normally, they will contain six features or terms that guide future regulatory actions: (1) the evaluation time frame; (2) the target or neutral zone in which there will be no incentive or disincentive; (3) the action trigger points; (4) the incentive-disincentive bandwidth; (5) the incentive-disincentive amount; and (6) the implementation mechanism.

The first two are concerned with the expected accomplishment to be realized with a given performance dimension. For all practical purposes, the evaluation period will be either one calendar month or twelve continuous calendar months. Other time frames, although unlikely, could be used. Specific numbers for the target or neutral zone in which there will be no incentive or disincentive are somewhat less tangible. These, of course, deal directly with the normal or expected level of managerial effort. To effect, the neutral zone is that range of values in which the reward-penalty factors are not operative. Clearly, the range must be such as not to penalize management unjustly for performance in the lower range. Likewise, careful attention must be given to what level of performance justifies activating the reward mechanism at the upper end of the performance scale. The guiding motivation should always be reasonableness.

By varying the width of the neutral zone from a single to a widening range of values, it is possible to address the difficulties of establishing equitable numerical yardsticks. If there is little precision associated with a measure of performance in terms of expected routine management effort, the neutral zone must be widened. It should be noted that, even under traditional regulatory approaches, a neutral zone is being used. When incurred cost is the yardstick, the neutral zone covers the gamut of possible operating results.

Once the neutral zone has been established, the trigger points for incentive and disincentive action and the bandwidth of incentive-disincentive measures must receive consideration. It is necessary to develop policies that will continue to be effective over a long period. The difficulties resulting from poorly conceived incentive programs illustrate this need. At one extreme, if goals are easily and quickly attained, a stifling of initiative and technical stagnation could result. A management which easily maintains the targeted goals would be able to
sit back and relax. At the other extreme, if targets are successively increased, the utility is effectively penalized for success. Ultimately, regardless of the effort exerted by management, only frustration, failure, and penalty are experienced.

Obviously, a balance must be achieved. The overfulfillment problem can be handled by an incentive provision incorporating a neutral zone based on the assumption that normally expected technical or financial performance would result in neither a reward nor a penalty. Targets or incentive bandwidths beyond the neutral zone can then be set which progressively require greater initiative and innovation on the part of management to reach and maintain. Use of a norm in this fashion could also handle the "penalizing success" problem. Diskincentive bandwidths below the neutral zone would penalize unacceptable results progressively and thus avoid financial disruption.

Another means for achieving balance would be the use of appropriate targets or yardsticks, such as the nationwide increase in productivity or cost of living, external to the operations of the utility. To be appropriate, the external yardstick and the utility performance measure must be allowed to fluctuate as the result of events reasonably common to both. The performance of business entities can be compared on several bases: when the businesses market exactly the same product, serve exactly the same territory, or are affected in the same way by phenomena such as inflation, productivity, or other widespread circumstances.

Performance measures, then, are of two types: those arising from technical considerations, which might be viewed as internal to the utility’s operations, and those arising from economic or financial considerations which might be deemed external. An expected electric generating plant availability of 70–80 percent is illustrative of a yardstick determined through technical considerations. An acceptable annual increase in “Other Operation and Maintenance Expense” (total operation and maintenance expense less fuel expense, purchased power expense, and production maintenance expense) equal in percent to that of the national all-items consumer price index exemplifies an externally determined economic factor.

There is one exception to the use of targets based on either technical or economic considerations: When utility performance is affected by many different cost factors. Fuel expense and purchased power expense are instances. At first glance, it might appear that many diverse performance factors, such as investment in generating plant, should be used to determine the efficiency of an electric utility in terms of fuel use or the purchase of power. But when a reasonable or equitable target for this area of utility performance is not available or practical, it may be necessary to fall back on regulatory delay as the means to encourage cost minimization. In the case of fuel burned expense, this might involve relatively instant recovery of only a portion of increases from a base level. Full recovery would have to be justified in a time-consuming formal rule proceeding.

The next feature of an incentive provision to be discussed is the amount of economic incentive or disincentive to be associated with various levels of performance accomplishment. In theory, this factor should be neither higher nor lower than that necessary to induce cost minimization. Since no empirical evidence presently exists to support the application of theory, the specific levels have to be, initially at least, set on the basis of judgment.

“Common sense” standards can be used. For example, if the reward is to be based upon attainment of a specific operating target, then the amount of reward must be less than the cost reduction achieved; if the reward is to be based upon management’s ability to live within an externally determined budget for selected expenses, then the amount of such expenses chargeable to the customers in cost of service must reflect a lower percentage increase than the company’s historical experience. Use of practical standards such as these will benefit customers and stockholders.

The process is similar for disincentives. Once the amount of reward has been determined, simply equate the amount of penalty to it. Assuming that the neutral zone has not been skewed or biased toward higher accomplishment than could be routinely expected, given normal operating circumstances, this does not appear inequitable. Reduced costs resulting from accomplishment beyond the neutral zone would very likely not differ to any significant degree from increased costs resulting from accomplishment below the neutral zone. This assumption would be invalid if changing costs and changing accomplishments did not occur on a roughly parallel basis. Due to this possibility, some care must be exercised in equating the numerical values used.

As a final step in determining incentive and disincentive factors, it is necessary to approximate the savings to be realized by utility customers when performance levels are reached above the neutral zone. If the dimension has been carefully selected, such savings should be susceptible to relatively straightforward calculation.

Implementation

The implementation mechanism pinpoints the exact procedure for moving from the determination that a reward or penalty is in order to effectuate rate schedule changes affecting customer billings. Alternan-
Incentive Regulation

tively, it provides for a change in revenue requirement level as a routine matter due to an allowable change in a yardstick or neutral zone level. The mechanism indicates how the changed revenue requirement level is to be calculated. It specifies the effective date for the related rate changes, states the period during which the rate changes are to apply, and spells out the future regulatory treatment of revenue effects stemming from the incentive provision.

The development of workable provisions is crucial to the broad acceptance of the concept of incentive regulation. Nothing could be less in the public interest than a program viewed with suspicion by the financial community, or one which handicapped able management in doing its job. Nor could the public accept what it perceived to be a giveaway program. Thus, incentive provisions must be developed which recognize routinely expected performance and then reward or penalize deviation from it.

The Michigan System

In 1976, the first step was taken toward instituting incentive regulation in Michigan. A new comprehensive fuel and purchased power cost adjustment clause was implemented. Under Michigan law, the fuel cost portion of this clause can operate without public hearings. It is, therefore, automatically applied every month. The purchased power cost portion is operable only after monthly public hearings.

Two critical aspects of these items were considered. The first deals with their magnitude. Fuel and purchased power expense represents about 40 percent of annual operating revenues for the average utility. The impact of increases in this expense and the current high probability of increases strongly argue not only for adjustment clauses, but also for all-inclusive ones, that is, to cover energy and demand components. It does not make sense to ignore the realities of utility operations and to assume away any of these expense components.

Providing an umbrella for these potentially volatile expenses encourages management to use the lowest cost power supply available at any time. Short-run marginal cost is the overriding economic consideration for day-to-day operations. Taking into consideration the magnitude of expense of items not only strengthens a utility’s financial viability, but also encourages the economic interchangeability of self-generated power supply and purchased power supply. Customer billings today tend to be increased by the former and decreased by the latter.

The second critical aspect addressed involves management decisions to procure coal supplies and purchase power. It has become quite clear that second-guessing management decisions several years after the fact is not an effective regulatory tool. Experience also has taught us that costs are not going to be reduced much, if at all, by increased auditing activity. A totally different means of influencing managerial behavior is necessary, and it must be operative when this decision is made.

Given the interest in increasing management’s competitiveness, it was thought desirable to spur cost minimization whenever a fuel or power contract was negotiated. A self-pricing incentive-disincentive was introduced by arbitrarily allowing a pass-through of only 90 percent of changes from a base level for these expenses. After a rate case, these incurred costs do show up in rate levels. This use of regulatory lag as the incentive mechanism, although far from ideal, appears to be the best means available. Setting equitable targets or neutral zones for this very significant expense category is impossible. Comparisons are virtually meaningless, given the many cost variables. Considering the influence of global economics and governmental policies, as well as actual, perceived, or managed scarcity, targets would tend to be whimsical at best.

One certainly can criticize the arbitrariness of the 90 percent figure, but it does provide a realistic and workable incentive to management. Since the purchase of fuel and power always has the potential to affect earnings per share, management faces financial expectation competition which acts to reduce costs. Furthermore, absorbing a 10 percent of expense increases is arguably not punitive nor, under an all-inclusive cost provision such as this, is a real threat in itself to financial viability. In Detroit Edison’s case, even assuming a $100 million increase, which is possible but one would hope not likely, swallowing $10 million reminds management that costs have gone away and are unacceptable, yet earnings per share are not affected in a disruptive manner. The magnitude of impact is about 10 cents per share. For these reasons, the fuel and purchased power cost adjustment clauses are a positive first step for incentive regulation.

The second step, taken in 1977, put in place an “availability” incentive clause. Availability is a quantitative measure of the ability of a number of electric generating machines to serve customer needs. The best measure is megawatts per hour. Between 1973 and 1976, Detroit Edison was unable to keep the machines running. System performance progressively deteriorated during that period and then dramatically rebounded in 1977. This happened to be the year the availability incentive provision was implemented and the first year for reward-penal tation review of performance in this area.

The provision was intended to motivate management to evaluate its
performance continuously in this costly area and to provide a very real impetus for improvement. Costs are much higher than necessary when generating machines do not work. An improvement of 5 percentage points in availability can reduce fuel and purchased power expenses for Detroit Edison by a minimum of $12 million.

In the past, management has spent a great deal of money on availability problems. Expenditures by Detroit Edison for production maintenance increased from not quite $50 million in 1973, or about 0.9 mill per kwh sold, to almost $65 million in 1977, or 1.8 mills per kwh sold. Even so, availability has remained at roughly the same unacceptably low level. Apparently, some crucial element of the problem has yet to be identified. The situation is similar for Consumers Power Company. From a regulatory perspective, this low availability seriously affects costs passed through to the consumer under the fuel and purchased power clause. It is also clear that increased maintenance expenditures have not solved the problem.

The relationships among production maintenance expenditures, availability levels, and fuel and purchased power expenses deserve discussion. As noted above, increased availability significantly reduces fuel and purchased power expenses. Production maintenance expenditures alone are very significant, amounting to about 4 percent of the utility's revenue dollar. These expenditures have a paradoxical quality. Inadequate expenditures can impair availability by increasing the possibility of forced outages. Excessive expenditures can impair availability by increasing the downtime needed for scheduled maintenance. Plainly, given a relationship between scheduled maintenance and forced outage rates, an optimization potential exists.

The availability incentive provision, by offering a bonus of 25 or 50 basis points to common equity rate of return, depending upon the availability level in a calendar year, should motivate management. It should ameliorate the availability problem by encouraging continued higher levels of availability and by reinforcing cost minimization.

A neutral zone availability standard for expected managerial accomplishment ranging between 70 and 80 percent was established. This zone was based on technical considerations set forth in a rate proceeding. The neutral zone reflected the regulatory desire not to be disruptive, and the rather wide spread was intentional. The East Central Area Reliability Council calculation procedure was adopted. The trigger points were set at the upper end to provide two accomplishment levels or incentive bandwidths, one ranging from 80.1 percent availability to 85 percent, and the other for availability exceeding 85.1 percent. A trigger point of 70 percent was set at the lower end of the availability disincentive bandwidth. The reward system involves a bonus of 25 basis points allowable for common equity rate of return at the first incentive level and an additional 25 basis points for attainment of the second incentive level. The penalty system is a one-step loss of 25 basis points. In terms of increased revenue requirement, 25 basis points amounts to approximately $5.5 million for Detroit Edison. This is well within the range of expected annual savings of at least $12 million. Stockholders benefit through an increase in earnings per share, and customers receive a lower billing. The provision effectively means that what is profitable to the public is also profitable to the stockholder, a paralleling of interests often thought to be mutually exclusive.

The third step, approved in 1978, instituted an incentive provision which ties the allowable level of "Other Operation and Maintenance" expense to the base level determined for calendar year 1978, adjusted for inflation as represented by annual changes in the national all-items consumer price index. It also acts to increase rate levels promptly in response to the impact of inflation on such expenses.

The other operation and maintenance expense item accounts for about one-sixth of a utility's revenue dollar. With this 1978 provision, in combination with the first two, the Michigan program now addresses almost two-thirds of a utility's cost of doing business.

The 1978 provision was implemented to induce competitive behavior on the part of management. By tying the standard or target for accomplishment to the performance of numerous other businesses in the economy, as represented by annual changes in the consumer price index, it was felt that the manager would be effectively subjected to outside cost pressures.

One obvious expense incurred by both utilities and other businesses is salaries and wages, which accounts for about 60 percent of other operation and maintenance expenses for utilities. Employee pensions and fringe benefits, and materials and supplies account for another 30 percent. It would be reasonable to assume that these proportions would be approximately the same among all businesses, although we do not have a specific breakdown.

A few figures will indicate the effort required of management to live within an operating budget for other operation and maintenance expenses when incentive regulation prevails, that is, when inflation increases are tied to increases in the consumer price index. I have compared the actual other operation and maintenance expenses per kwh of sales from 1973 to 1977 with the figures under an incentive system. The per unit expense in 1973 was used as the base for subsequent increases related to the consumer price index. The maximum difference in any year was 0.35 mills for Detroit Edison and 0.38 mills...
case and should be treated as mechanisms for making future adjustments to the specific rate award resulting from the rate case.

The incentive clauses are designed to change a utility's level of revenues upon the occurrence of certain specific events in the future. The adjustments need to relate to an instance when a utility's rates were determined by a regulatory body. Equally important, the regulatory body must decide that the adjustments in utility revenues produced by the foreseeable operation of the clauses, taken in conjunction with the revenues produced by the tariffs approved in the overall rate case, will produce a "reasonable rate of return" for the utility.

Failure to link the initial approval of the clause to an overall examination of the utility's need for revenues could result in the need for a full rate case when the first rate adjustment is made as a result of the operation of a clause.

In Michigan (General Telephone Co. v. Public Service Commission, 341 Mich 620; 67 NW2d 882 [1954]) and throughout the United States (McCardle v. Indianapolis Water Co 272 US 400 [1926]), it is accepted that a regulatory commission should consider factors not only relevant for the present but also for a reasonable period into the future. In essence, a commission must determine that, although a specific rate level is just and reasonable at the moment, events which are foreseeable (such as inflation or changed levels of efficiency) will necessitate specific future adjustments.

Common sense, as well as legal requirements, would dictate the need for initial consideration of an indexing formula, cost adjustment formula, or any other adjustment mechanism in the context of a full rate case. The need to meet this requirement is reinforced by a Massachusetts supreme court case, Consumers Org for FEE, Inc v. Department of Public Utilities, 335 NE2d 341 [1975]. It confirmed the nationwide pattern of cases which initially require full consideration of any cost adjustment formula within the context of a rate case, but which also hold that subsequent adjustments resulting from application of such formulas need not be the subject of full rate case hearings.

In regard to New Jersey's "comprehensive clause" (an adjustment clause for telephone rates), the state supreme court affirmed the necessity for an adjustment clause to relate back to an overall determination of the adequacy of a utility's rate level. In the Matter of the Board's Investigation of Telephone Companies, 66 NJ 476; 333 A2d 4 [1957], the same court struggled mightily to find a connection between the hearings which applied the New Jersey comprehensive clause and the preceding rate case hearing which had adopted the clause itself. What is important about this case is the obvious discomfort which the court felt in having to uphold the New Jersey Public Utility Commission's

Legal Prerequisites for the Adjustment Clauses

It is essential to the functioning of the Michigan regulatory system that each incentive proposal be applied through mechanisms which are timely in terms of providing rate relief for the utility and make economical use of the resources of the commission and its staff. If implementation took as much time as a full rate case, an opportunity would be lost to remedy one of the most significant defects of the traditional process, that of regulatory delay.

The Michigan commission has decided to use adjustment "clauses" to implement incentive regulation in a timely and efficient manner. If four basic legal prerequisites are met, the likelihood is greatly increased that the clauses will comply with constitutional and statutory law.

Implementation in a Rate Case

The clauses initially should be implemented as part of a full rate
comprehensive clause as a complete and separate ratemaking process rather than as a method by which rates could be adjusted subsequent to a more conventional rate case or as interim relief prior to a final conventional rate case decision.

The Need for Public Hearings

Unless the automatic clauses are authorized by statute, public hearings should be held after notice to interested parties and prior to implementation of rate increases resulting from changes in an "index" level or from application of a "system availability" incentive.

It is not clear that statutory authorization is a prerequisite to any clause which could increase rates without a hearing. As will be shown below, it seems clear that when a clause is not authorized by statute, notice and public hearing before implementation of increased rates resulting from operation of a clause will greatly enhance the legality of such rate increases.

One leading case which is specifically applicable is Petition of Allied Power and Light Company, 321 A2d 7, 132 VT 354 (1974). The Vermont Public Service Board attempted to implement an "automatic" power and fuel adjustment clause by means of an administrative ruling. This clause provided for the periodic adjustment of utility rates on the basis of changes in the cost of fuel and purchased power to a utility. The adjustments were to be accomplished without notice or hearing. The state supreme court held that a Vermont Statute, 30 VSA 225, 226, which required notice to interested parties prior to a change in rates, was applicable to changes in rates attributable to the operation of fuel and purchase power clause.

The supreme court of Massachusetts, in Consumers Org for FEE Inc. v. Department of Public Utilities, 335 NE2d 341 (1975), sides with the majority of states in reaching a different conclusion than that set forth by the Vermont court in Allied. Despite footnote 11 at page 346 urging the contrary, the Massachusetts court seems to depart from Allied by approving a fuel and purchase power clause which was not authorized by specific statute and which provided no notice or hearing for rate increases caused by the clause. This case can be distinguished from Allied because Massachusetts does have a statute, GL c 164 sec 98; St 1970 c 615 sec 4, which permits any customers to file a written complaint concerning the price of electricity. The Department of Public Utilities is then required to give notice, hold a hearing, and order any suitable changes in price. This distinction fails to explain the fact that numerous other jurisdictions have approved adjustment clauses, even when the law does not contain the consumer protections of the Massachusetts law.

As will be seen below, the more practical view of the law would seem to be that there is considerable, although not universal, legal precedent for judicial approval of fuel and purchase power clauses, even in the absence of specific statutory authority. The rationale for these decisions, as summarized in Consumers Org for FEE, is that fluctuations in fuel, gas, and purchase power costs are either regulated by other entities or are largely beyond the control of utilities. This specific factual rationale does not always apply to the subject matter of expanded clauses which are not specifically authorized by statute.

Many state legal officers and courts, therefore, have been reluctant to dispense with notice requirements when regulatory commissions have attempted to expand the scope of existing adjustment clauses which are authorized by statute or by case law.

In Michigan, statutory law 300 PA 1972; MCLA 460.6a(2) specifically authorizes "fuel or purchase gas adjustment clauses" which do not require public hearings, but the Attorney General, in Opinion No. 4844, 1974, opined that such a provision could not be expanded to include "automatic" compensation for purchased power expenses.

In Wisconsin, fuel and purchase power clauses had been permitted, although such clauses were not specifically authorized by statute. Wisconsin Electric Power Company (WEPCO) applied to the Wisconsin Public Service Commission for both rate relief and expansion of its fuel and purchase power clause to include such items as labor, supplies, steam, electrical expenses, and supervisory expenses. Although WEPCO's application for expansion of the clause was heard after notice to the public as a part of the rate case, WEPCO's application contemplated that subsequent adjustments to the clause would be accomplished without notice or hearing. The Wisconsin Public Service Commission approved this request in 1974.

In 1978, the supreme court of Wisconsin held, in Wisconsin's Environmental Decade Inc. v. Public Service Commission of Wisconsin, 81 Wis 2d 344, 260NWd 712 (1978), that the expanded WEPCO clause was in violation of the laws of 1931, ch 183, Sec. 2, sec. 196.20 (2), states, which require public hearings prior to a change in rate schedules which constitutes a change in rates. The Wisconsin supreme court was careful to note that its ruling did not affect the validity of Wisconsin's current fuel and purchase power adjustment clauses. These operate to increase rates without public hearing and are not based on a specific statutory exemption for elimination of public hearing.

This Wisconsin case may not decide the issue of the need for public hearing before increasing rates through clauses which include items other than fuel and purchased power, but the case does raise important questions in this area.
In summary, the legality of rate increases through adjustment clauses not specifically authorized by law is greatly enhanced by notice and public hearing prior to implementation of such increases.

The Scope of Hearings

The proceedings authorizing changes in rate levels through adjustment clauses may be restricted to consideration of only those rate changes directly attributable to the clauses.

In order to avoid the delays and expenses of full rate proceedings, the scope of adjustment clause hearings must be restricted to only those factors covered by the clause. Adjustment clauses are beneficial, in theory, because they allow the timely matching of revenue and expenses. The regulatory lag inherent in full rate proceedings would destroy this benefit. There is substantial legal precedent for restricting the scope of public hearings which implement specific adjustments.

A case in point is *In Petition of Allied Power and Light Company*, 132 VT 554; 321 A2d 7 (1974). The Vermont supreme court stated (page 13) that, where frequent adjustments for purchased power and fuel cost had to be made and where such adjustments had to be made in the context of a public hearing: “to require on every occasion complete reexamination of the rate structure of a utility would be wasteful and redundant. Absent a demonstration of a significant change, it would seem that the previously filed evidence on other aspects [of a full rate case], as evaluated by previous Board decisions, should stand.”

Consumers Power Company v. Public Service Commission, 85 Mich App 73; 237 NW2d 189 (1975), upheld a decision of the Michigan Public Service Commission which, in the context of a rate case decision, provided for subsequent, automatic reduction of plaintiff’s revenues upon reduction in the federal tax surcharge. The plaintiff argued that such a reduction could not take place without consideration of offsetting increases in plaintiff’s expenses of operation which had occurred subsequent to the commission’s rate decision.

The Michigan Court of Appeals tied its finding to the concepts set forth in *General Telephone Company v. Public Service Commission*, 341 Mich 620; 67 NW 2d 882 (1954), which specifically holds that commission rate increases are prima facie lawful and reasonable and hence that such rates would be a legitimate basis for an adjustment due to foreseeable events.

It should be noted that the court’s approval, in *Consumers Power*, of an “automatic” adjustment without public hearing can be distinguished from other states’ court decisions requiring hearings by a reading of the Michigan Law, 300 PA 1972, 6(a)(1); MCLA 460.6(a)(1), which permits reductions in rates without notice or hearing.

A recent Michigan decision affirms the concept that clauses which were adopted to adjust rate cases but not approved by statute could be implemented in hearings which are restricted in scope to exclusive consideration of the adjustment resulting from the clause. In *Attorney General et al v. Michigan Public Service Commission*, Circuit Court of Ingham, Docket No. 76 18795 AA (1977), the Attorney General challenged a purchased power clause approved by the commission. It was not authorized by statute but was authorized in a rate case and was to be implemented by a series of monthly public hearings. These were to be publicly noticed and were to be restricted to consideration of purchased power costs only. The Michigan circuit court rejected the Attorney General’s motion for a partial summary judgment. In so doing, it found the commission’s adjustment mechanism “not . . . unreasonable.” No final disposition of this case has occurred as of this writing.

Nature of Proceedings

The hearing concerning the adjustment should be conducted in accordance with all the procedural and due process requirements of a full case with the exception of the restriction of the scope of the proceeding to one subject.

As noted above, full notice and hearing in adjustment proceedings should occur if such adjustments are not statutorily exempt from notice or hearing requirements. The hearing itself must be conducted as a rate case, with witnesses, cross-examination, and a commission order.

Both the Allied Power and Light Company and the Wisconsin Environmental Decade cases require that adjustments be made in the context of a “hearing.” With the exception of restrictions in scope, as discussed above, it must be assumed that a “hearing” includes the full procedural requirements which would apply to any other rate proceeding before a commission.

Reactions to Incentive Provisions

Two major themes emerge from the numerous comments reaching the Michigan Public Service Commission. First, the concept of tying rate levels to some measure of performance is both understood and approved by the public and news media. Second, the use of the CPI as a yardstick for allowable rate increases has been popularized in Michigan by its widespread use in the cost of living allowance clauses in automotive wage contracts. Hence, this use of the CPI makes sense to many
citizens who otherwise would never understand rate regulation and are violently opposed to any "cost plus" contracts.

This element of improved citizen understanding cannot be over-emphasized. In the recent past, utility regulation has become highly visible. Unless the public can understand the process, it will have only one basis for judging the result: Are the rates higher or lower? If we are to avoid or reduce citizen unhappiness with regulation, we must make the regulatory process relevant to other economic concepts which citizens do understand and accept, if not approve. The necessity for cost increases due to inflation or rewards for efficiency are concepts which are understood and widely accepted.

Conclusion

Obviously, no system or principle of regulation can hope to solve all problems in all jurisdictions. The system of incentive regulation described here is a response to a set of conditions and a regulatory environment which may be unique to Michigan. At the very least, they may be unique to regulatory jurisdictions in which the traditional ratemaking process is performing unsatisfactorily.

In Michigan, incentive regulation appears to have solved or mitigated some of the most serious problems facing the commission. Most important, it has gained widespread public support because it avoids the negative image associated with any cost plus scheme of pricing. Furthermore, the incentives and disincentives apparently convince many citizens and opinion makers that utility rate increases must be earned on the same basis as the majority of workers earn their salary: by acceptable performance, rather than as a matter of right. An additional benefit to consumers is the concept that any incentive payment to a utility is usually exceeded by the benefits to consumers which are produced by superior utility performance.

For utility management and regulators, incentive regulation is an approach which can minimize many of the problems traditionally associated with the ratemaking process. The use of clauses administered in restricted scope hearings can achieve decisions which not only are more timely for the utility than a full rate case, but also require less commission manpower. Furthermore, setting goals for efficient operation or fuel procurement, which is inherent to incentive regulation, harnesses management talents while allowing management almost unlimited latitude to determine the best way to meet those goals.

Perhaps most important, incentive regulation ties rate relief to a utility's ability to meet performance standards which were specified before the utility attempted to meet them. This concept is not as prevalent in the utility ratemaking process as one might assume. Indeed, the most common method of determining utility compliance with the so-called reasonable and prudent standard is to specify performance criteria after a plant has been built or fuel has been purchased. Such a system of performance evaluation has disadvantages for both regulators and utility management. Regulators are often frustrated because an after-the-fact penalty can jeopardize a utility's financial health, whereas no penalty will signal regulatory indifference to inefficient management decisions. Utility managers often find a perfectly reasonable management decision being subjected to hindsight judgment, which takes into account factors which no prudent manager could have foreseen. The system of specific goals contained in incentive regulation allows both utility management and regulators to know the rules of the game before the game starts.

As noted earlier, no one system or concept of regulation can solve all regulatory problems or suit all regulatory jurisdictions, but in Michigan we believe that incentive regulation has been, and promises to be, a very positive force in our environment.
Our topic, “Assessing New Regulatory Concepts and Tools,” first raises the question of what is intended to be achieved by the use of new regulatory tools and methods. One obvious answer would seem to be to provide a system of regulatory rules that function more effectively to secure adequate service, now and in the future, at the lowest cost, while maintaining the financial integrity of the utility and providing reasonable compensation to the investors. Still another answer might be that, in addition to rules that achieve the condition of adequate service rendered by financially healthy utilities, a desirable goal is to restore and maintain public confidence in the regulatory system as being a cost-sensitive mechanism in an era of rising costs and higher utility rates. Personally, I think that public perception of the system as being cost conscious and effective in holding costs down is tremendously important. It is essential that the public generally, and particularly legislative branches of government which are most responsive to public attitudes, be made aware of the fact that the utility industry and its regulators are constantly looking for ways to improve and to keep costs from rising any more than absolutely necessary.

It should be emphasized and more effectively explained to the public that the traditional methods of rate regulation in the utility industry have promoted a number of incentives to reduce costs. All costs incurred immediately affect present earnings adversely; whether increases in costs are subsequently allowed to be recovered through higher rates to the public will depend usually upon whether they are normal and recurring and whether they become a part of a test period for ratemaking. Even assuming that increases in costs will be covered in future rates, there is always a period during which they are not recovered and therefore lost forever. In recent years the impact of regulatory lag upon most electric utilities has been so great that a few regulatory commissions have turned to forward test periods, to estimate for some costs such as fuel, and to allow for attrition in setting the rate of return.

Today, the public protests against rising electric rates have been such that utility management feels even more keenly than in the past the need to operate as efficiently as possible. Even when higher costs are finally permitted to be recovered through higher rates, there is always the prospect of emotional and hostile public reaction, including attacks through the news media, that may be translated into restrictive legislation which has the effect of preventing rates from rising to cover costs. It should go without saying that the faster costs increase, and rates rise, the more difficult and unpleasant it is for utility management.

The utility regulatory commissions with which I am familiar have strong and effective staffs. Should this not be the case everywhere, then staffs should be strengthened. Commission staffs, through investigative audit functions, exert considerable pressure upon the utility, which lives in a “goldfish bowl,” to be as diligent as possible in controlling costs. This fact should be made better known to the public, perhaps by communications directly from the commissions themselves.

Recently inaugurated management performance audits by qualified outside consulting firms also can provide both ideas and incentives for more efficient operation.

Incentive regulation and efficiency standards seem to offer promise if used realistically. Certainly they are preferable to regulation which usurps managerial decision making. The latter would mean, in effect, two managements and two full planning staffs. The role of the regulator as one who determines whether or not the proposals of management are “just and reasonable,” rather than a new role of trying to manage decisions from among available choices, was wisely developed and continues to be the proper function of regulation.

The development of new concepts and tools, a current emphasis which I favor, should include retaining those traditional components which have been tested over many years and continue to have value. As
new techniques are developed, it is often best to try them slowly on an experimental basis in order to obtain reasonable indications of results, thus avoiding drastic changes which could either severely cripple the utility by providing revenues that are not adequate to cover costs, or impose financial or other hardship on the consumer. Incentives: a refinement of regulation ought to be encouraged, certainly on an exploratory basis; again, care must be taken to see that they are equitable and are not disincentives rather than incentives. To supply the electricity needs of our country, we must be fair to several groups—employees, suppliers, rate payers, and investors. Without proper treatment of all four groups, we cannot expect quality service at the best price. The rate payers are entitled to efficient management of the utility; the investor must have a sound regulatory climate if he is to be persuaded to place his capital at risk. Specific incentives to efficient operation must be designed equally to encourage the risk of capital and protect the rate payer. The true incentive is a climate which makes sensible action by management a realistic possibility.

Any incentive plan, such as the one used in Michigan, must be readily reviewable by the regulator. This facet is particularly applicable to the Michigan expense adjustment for “Other O&M Expenses.” If the initial base has been set too low, or if substantial lag begins to develop, then there must be a quick means of making adjustments. In that specific regard, I note the general statements in Daniel J. Demlow’s presentation about the relationship between the Consumer Price Index and the components of the “Other O&M Expenses” of utilities. I do not fully understand, however, whether an incentive study was made to see if the Consumer Price Index components truly represent the makeup of the utility “Other O&M” price increases.

The 10 percent underrecovery in the Michigan automatic fuel clause raises the general problem of whether the deliberate use of regulatory lag as an incentive is a sound procedure. I question whether an arbitrary disallowance of recovering 10 percent of a cost increase is an incentive, or whether it is an arbitrary penalty, particularly when we have every reason to think that the prices of coal and oil are going up. Is there a valid incentive if there is no way for the utility to win, but only a way to minimize losses? It seems to me that the 90 percent arbitrary provision, when coupled with the outlook for continuing increases in mining costs due to wage increases, safety requirements, and probably limited productivity improvement, as well as continued price increases for oil, may produce frustration and a tendency to cut corners in other ways which may actually impair the overall quality of service over time.

The generating plant availability clause established in Michigan is a commendable effort at innovation, but it may also be subject to a critical comment. It would appear from the statement that increased in pro-

duction maintenance expense have not provided increased reliability, and that some factor is standing in the way. Therefore, the adjustment appears to have been adopted, it seems to me, in the absence of a full understanding of cause and effect. Under those circumstances, could not the classical or other hardship be an incentive to operate at the expense of safety, or necessary maintenance, or to build a more expensive plant than normal engineering practice calls for, in order to maximize reliability? If so, the result would not be effective in the effort to hold down the rising costs of attracting new capital.

There is a fundamental concern with limiting “Other O&M” expense increases to increased percentages in the Consumer Price Index. As Demlow now noted, the assumed increase in the CPI is 7 percent annually, whereas the increase per kWh of sales will be 3.9 percent in two future years. As he notes, the increase to the customer is roughly one-half of the annual percentage increase in the CPI. He states that productivity increase should account for a part of this difference, but there is little in the offering today to indicate that such an increase in productivity is achievable. In fact, the seemingly continuous moving target of expenditures for environmental and safety requirements raises questions as to whether productivity can be increased at all. As with the fuel clause, regulatory lag appears to be the ultimate incentive, if I understand correctly the application of this provision. Again, either there will be frustration of a management which may be doing the very best possible, or else there must be a reduction in the only place left—the manpower for providing quality service.

Otherwise, however, this provision requires a ready willingness to adjust as circumstances develop, and I am happy to understand that the commission recognized such a need in its order establishing this regulatory tool.

If it is accepted that a goal of regulation is to achieve the same results in a monopoly industry that competition would achieve, then deliberate underrecovery of expenses has no place in the regulatory scheme. Such seems to be the situation in the Michigan fuel and purchased power adjustment clause, and, realistically, this probably is the situation in the “ceiling” recovery of “Other O&M Expenses.”

On balance, however, I commend the efforts of the Michigan Public Service Commission to try innovative ideas through incentives to improve efficiency. My purpose here is simply to note, or to emphasize, problems that may exist and that urge that a key to adopting incentives such as these is the willingness to recognize that results may not be what were anticipated and to make corrections without delay. Technology and other circumstances can change quickly. Will the commissions and their staffs be able to change the standards quickly? Rapid response to inequitable results of any incentive formula is essential to adopting the
plant availability and the "Other O&M" provisions. If the 90 percent limitation is to be retained in the fuel and purchased power clause, then the other 10 percent must be rolled in much more expeditiously than is normally done in general rate cases. While Demlow feels that the potential impact on earnings is not disruptive, from my viewpoint an impact of about 10 cents per share could be quite disruptive.

Now, as to Thomas K. Standish's thesis, I cannot agree that regulators have not and cannot effectively regulate. The existing system must not be all bad by any standard of measurement. Generally, good results have been produced, such that in this country we have probably the best utility services, and at the most reasonable prices, in the world. Just because we have changing circumstances does not call for thorough condemnation. While new approaches based on incentives may merit consideration and in many cases implementation, they should be recognized as refinements, not replacements, of the present system. Refinements are well in order, but the traditional regulatory approach remains sound for the basic problems. The correctness of this viewpoint seems to me to be recognized by the Michigan commission in the approach that it has taken.

Standish's extended criticism of marginal cost pricing seems to me to document very well many of the basic flaws in the theory of "marginalism." From some of his other statements, however, I assume that he would advocate time-of-day rates. He notes that calculations have not been made that net cost benefits would result, but he seems to feel that such calculations would show net benefits. Perhaps it should be noted here that we still lack a fully developed methodology for determining comprehensively whether adoption of particular time-of-day rates under specific circumstances would result in system improvement.

The apparent dismissal of auditing and monitoring of the purchasing and operating practices of utilities as a thoroughly workable incentive is disturbing. If the monitoring of fuel and purchased power practices, plant construction and operation expenditures, and the amounts of other operation and maintenance expenses cannot establish the efficiency or inefficient of performance in these areas, how can studies produce the basic figures for adjustment clauses or expenditure ceilings to any better degree, or regulators judge whether the resulting consequences are in line and just and reasonable? Performance audits and monitoring have almost become a way of life in American business, and this is because they have been tested and proven to be effective techniques in controlling costs and promoting efficiency.

One of the problems I perceive from the practical application of Standish's doctrine is the prospect of never-ending rate proceedings, as each public interest or consumer group seeks to make its position dominant. In recent years we have seen that when all parties seek to take advantage of "due process," the regulatory system can become manipulated, especially through delay. There is great danger, too, in the concept that "regulation must be fashioned to penetrate through traditional management prerogatives and in the place of management autonomy." Such a practice would surely lead to purely defensive management by utilities; and if the regulators become the managers, who then will regulate the regulators?

Although I have pointed out some problems with specific types of incentive-disincentive mechanisms, let me state my firm belief that the system ought to encourage efficiency. To me, it already does in a number of ways. I am convinced that, for the great majority of utility people, the personal incentive for success in doing a good job for customers and investors is very strong. With continuing audits, internally, by outside consultants, and by commission staffs, the pressure to manage well is intense. The existing system goes much farther. It already contains a built-in financial incentive in the combination of inflation and regulatory lag. At present rates of inflation, the goal should not be to increase the lag or even tolerate its present level, but to reduce it. Clearly, the inability to reflect promptly more cost-increase items than the ones singled out in this discussion is a great problem.

The real incentive should be to remove obstacles to efficient operations. Utility managers today are sometimes forced to operate in this situation: Until their companies are in dire financial distress, they will not be allowed rate relief. Regulators should work to reduce regulatory lag. Also, an aid to more efficient operation would be a more realistic appraisal by regulators of the needed return on common equity, and rates that will produce it, which means that an attrition allowance should be provided.

Some years ago we instituted a regular system of making comparisons of key cost statistics of our company with other companies in the southeast. While some may not favor this type of comparison, and there are limits to its use, we now find our company being approached for information about our method for conducting the study and appraising the results. Improperly done, such comparative studies can be misleading; properly done, however, these can be a great incentive to efficient operations, when one remembers to look carefully behind the bare figures into the details of the operations, which usually can be made available.

I have said that it is a disincentive to efficient operation if companies are put into a "no-win" situation. An excellent illustration is the case when a number of companies place their stockholders' capital in mining ventures in order to ensure fuel supply. The system should encour-
age such prudence rather than discourage it. In hindsight the undertaking turns out not to have been the most advantageous, but to have been made in good faith with reasonable judgment under the circumstances at the time the decision was made, then the utility should not be penalized. Conversely, if future developments establish that the decision was very fortuitous, then the stockholders should not be held to a bare-bones return for having risked their capital.

One very interesting innovation in utility ratemaking is the “cost-of-service index,” as developed and applied in New Mexico beginning in August 1975. The index originally provided for quarterly reviews of electric rates and permitted an upward adjustment in rates when the return on common equity for the past twelve months was below 13.5 percent; a downward adjustment was required when the return was above 14.5 percent. After several years, the New Mexico Public Service Commission conducted an extensive investigation into the operation of the index and determined that the system had apparently succeeded in attaining two basic objectives: reducing capital costs and improving the company’s ability to attract capital. In commendation of this indexing approach, the commission concluded that traditional forms of ratemaking were not adequate to enable a very rapidly growing electric utility to raise large amounts of new, outside capital. The commission, however, has now moved from a quarterly to an annual review, is requiring additional reporting of information by the utility, is permitting some lag, and is providing for more regulatory oversight of the utility’s operation. It will be interesting to follow the application of this indexing system, although it seems now to be very evident that the utility’s cost of capital has been substantially reduced, from what it otherwise would be, because of the improved likelihood of the recovery of costs in a way that will sustain an adequate rate of return. Also, this element of improved likelihood of earnings realization should enable a utility to improve management performance through better long-range planning.

Let me return to the management audit. A recent survey of 28 southeastern utilities shows that 16 have undergone such audits, either a proprietary audit by an outside consulting firm sponsored by the utility, or an “adversary” audit directed by the regulatory commission. One may argue the relative merits of each; the proprietary audit is likely to be of more lasting benefit to the utility, and the commission-directed audit is likely to have the greater public acceptability. In either case, I believe that such audits can be very valuable in improving the efficiency of utility operations.

In 1976, Carolina Power & Light Company underwent a thorough audit by a leading consulting firm selected by the North Carolina
Comments

Robert J. Rohr

There can be little dispute that rate-of-return regulation of utilities is not working very satisfactorily. Utility executives are unhappy. They argue that the combination of regulatory lag and basing rates on a historical test year result in the inability of firms to earn their allowed rate of return (as meager as it may be viewed). Commissioners are frustrated. Since, in many instances, large portions of a requested rate increase have been mandated by past investment decisions, there are often few options but to approve such requests. Frequently, there is little regulators can do to encourage efficient utility performance in supplying the regulated good. Economists who have studied regulated utilities are concerned about the economic losses that result from failure of utilities to operate efficiently. Given this general level of dissatisfaction, the proposals of Commissioners Daniel Demlow and Thomas Standish to introduce efficiency incentives into the regulatory process should be welcomed. While the incentive properties of these proposals may not prove acceptable in practice, their focus on efficiency is certainly a movement in the right direction.

There are essentially two identifiable schools of thought on how to achieve the common goal of more efficient utility operations. The first, with which Demlow has aligned himself, bases its approach on economists' respect for market solutions. This school would create incentives based on market substitutes, with utility executives free to maximize their own best interests in any manner they choose. The second approach, propounded by Standish, is based on a belief that markets fail to produce goods and services in efficiently preferred quantities. For this school, called the "New Wave" by Standish, the only means to achieve efficient utility operation is the direct involvement of commissions in the management of utilities (called "interventionist regulation" by Standish).

The Standish Approach

The justification for Standish's interventionist regulation is that markets fail in essentially two respects: They fail to produce goods efficiently and to distribute the economy's output in a way that meets some unspecified standard of a "fair" income distribution. Standish's conclusions concerning failures of the first type are not surprising. Most of the particular cases that he considered have been identified in the literature by welfare economists. The standard against which market failures traditionally have been compared is the abstract criterion of Pareto optimality, a standard to which everyone should subscribe. As Standish points out, to the extent that the benefits or costs of any given decision are not fully internalized by those making choices, the results of the market will not meet the Pareto standard. The classic example is the divergence between the marginal private and marginal social cost for a polluting firm that imposes external costs on others in society. Since these external costs are not taken into account in the firm's decision processes, the Pareto criterion will not be achieved by a market solution. This type of economic analysis provides the justification for government intervention.

The failure of the market to distribute society's product in a "just and reasonable fashion" has not been as rigorously developed in the economics literature as have failures of the first kind. Under a market solution, income will tend to be distributed to those possessing large initial resource endowments (talents and skills, as well as capital). This inevitable result offers a plausible basis for government intervention in the market to achieve a more "equitable" distribution of income in society. Although the "injustices" of a market income distribution may be universally recognized, without criteria for comparison of various distributions, it is impossible to achieve general agreement on the appropriate method of correction. Lacking a criterion, coalitions are typically formed that attempt to shift income and wealth toward their best interests and away from others. The end result may well be an
income distribution that does very little for those groups originally considered deserving.

After comparing an idealized model of competitive behavior with the less than perfect world in which we operate, Standish concludes that "competitive standard" regulation has not and cannot work and must be abandoned in favor of his interventionalist form. Standish states, without documentation, that traditionally "regulators have willingly relegated themselves to the limited role of setting economic price at a level which reflects the so-called competitive standard." The available evidence indicates that public utility commissions have not behaved this way until perhaps very recently. In the recent past, commissions were primarily concerned with the total dollars that a utility should be allowed to raise and considered price structures only in terms of revenue realization and customer claims of discrimination. Consequently, a public policy concerning utility price structures and the effect they would have on the scarce resources of society was not considered. This is not to suggest, however, that utilities and commissions were not cognizant of costs when designing rates. Cost relationships have always been a consideration in designing a value of service rate structure.

The decreasing block rate structure offered residential electric customers provides a good example. Because the system requires a minimal distribution grid that depends on customer location and because customers require meters, meter reading, and billing regardless of the number of kilowatt-hours they consume, it seemed clear that average cost per kilowatt-hour would fall as average consumption per customer increased. Recognizing this, believing that larger users have better load factors, and considering that the value of service falls as a customer extends his use of electricity, ratemakers quite naturally proposed and commissions accepted declining block structures.

In a world in which continually expanding operations led to lower unit costs and in which technological advances were reducing unit costs at every level of operation, it made perfect sense to price electricity in a manner designed to promote its use. Expanding use meant capital-intensive plant expansion, with the associated reduction in costs, while revenues increased. New investment opportunities justified profit growth and resulted in a more than adequate rate of return.

The clearest evidence that this was a happy circumstance for utility stockholders was the level of stock price. In the 1960s, electric utility stocks generally sold at more than twice book value.

Even with stockholders enjoying exceptional returns, rate payers found no obvious reason to protest. With the entire price structure falling absolutely and even more dramatically relative to the price of other goods and services, and with expanded consumption in the lowest priced final block, customers viewed electricity as a bargain. In its time, the declining block price structure was not a source of customer dissatisfaction.

Because the rate structure satisfied both utility companies and their customers, it also appealed to regulators. There was none of the unpleasant friction that develops when one or another of the parties to regulation is unhappy. Rate cases were infrequent, making leisure an additional reward for service as a commissioner. Appointments to a commission were riskless rewards a governor could provide without fear of subsequent embarrassment. Careful scrutiny by legislators was unlikely. Expanding electric utility operations meant greater prestige and power for regulators and increased the opportunity, as well as the likelihood, of employment in the industry when the term as commissioner ended.

With all parties satisfied, prices declining, quality improving, and new uses of electricity appearing, there was no need for extensive class rate-cost-of-service studies. There was no need for regulators to concern themselves with setting economic prices.

Had regulators been concerned with economic price, stocks would not have sold at twice book. There would not exist electric utility rate structures with the tail block set below the marginal cost of fuel to provide an additional kwh. The tail block for natural gas rates would be pegged to the price of oil, an action that was only recently taken by the New York Public Service Commission. To conclude that "competitive standard" regulation has been applied in the past and failed is just not consistent with the facts. Regulation of utilities in the past may have been inadequate precisely because commissions did not attempt to set price in accordance with the competitive standard. That standard has only recently (if at all) been adapted by commissions in setting the price of utility services.

By proposing to "regulate according to explicit standards of economic efficiency," Standish seems to advocate that regulatory decisions be judged by the effect these decisions will have on the use of society's resources. This criterion is also advocated by the "competitive standard" school in its support of cost-benefit analysis. Where the two schools may differ is in the meaning of a fair and reasonable end result. Standish's proposal seems to be a plea for "taxation by regulation," with primary emphasis given to the distribution of society's product and the implementation of different technologies. As Richard Posner points out, when commissions attempt to satisfy costly social or political goals, they are forced to practice taxation by regulation, and the tax has not always been collectible.
The Demlow Approach

The Michigan approach to encouraging increases in utility operating efficiency, as presented by Demlow, is the creation of incentive schemes under which both the utility and its rate payers would benefit from improved utility performance. In this system the commission would establish (it is hoped) realistic parameters for utility performance. Given these parameters, the utility is then left to its own devices to meet or exceed them. It is a plan whereby a utility can gain in terms of allowed rate of return if it exceeds an upper target parameter, will not lose if it ends up in the neutral zone, and will be penalized if it falls below a lower parameter target. As Demlow properly observes, the key to this incentive program is the development of the upper and lower parameters.

The specific proposals implemented in Michigan are basically adjustment clauses: (1) a fuel adjustment clause that allows a 90 percent pass through of fuel charges over a rate case determined base; (2) a plant availability clause that permits the allowed rate of return to increase by 25 basis points and 50 basis points when plant availability exceeds 80.1 and 85 percent, respectively; and (3) an operation and maintenance clause that allows increases in such expenses to be passed through only to the extent that they can be explained by increases in the Consumer Price Index.

Because these incentive programs have only been in effect since 1976, their value in increasing the efficiency of utility operations is yet to be established. There is no question that these programs must be dynamic, and they must change as conditions warrant. The institution of these clauses is important not because they will solve existing problems (as they certainly will not), but because they were introduced at all.

Although I feel that the Michigan proposals are a movement in the right direction and that their introduction was a proper departure from past practice, they are not without potential pitfalls.

A fuel adjustment clause that allows a utility to pass through something close to the full fuel increase is not new or original. The Michigan proposal to allow 90 percent recovery through the clause appears to be a compromise that provides utilities with some degree of financial stability while being acceptable to rate payers. As Demlow properly acknowledges, the Michigan clause is not an incentive mechanism; rather, it is a penalty mechanism driven by regulatory lag. Since bygones are bygones in this business, a firm that is extremely aggressive in locating and purchasing the lowest cost fuel available would still be forced to swallow 10 percent of any increase over the established base.

A utility regulated under this scheme could never gain from efficient management when fuel prices are above base level but could only attempt to minimize its losses. The plan's only positive incentive is to encourage utilities to establish a sufficiently high base for the fuel clause to provide a 90 percent credit for rate payers and a 10 percent bonus for the firm. An attempt to follow this strategy would most certainly be viewed by the commission as an attempt to abuse the clause.

The plant availability clause is certainly innovative and does provide a positive incentive for firms to improve availability. It is not without potential problems, however. The first such difficulty stems from a definition of availability that includes the entire mix of a utility's generating plant. Since potential fuel savings of significant magnitude must come from increased availability of efficient (base load) units, it may be necessary to refine the clause in a manner that reflects the availability of those plants where large fuel savings can be realized. There appears to be very little fuel savings potential from 100 percent availability of gas turbines.

A second area of concern is the clause's bias toward the use of proven technology. It is reasonable to assume that, recognizing the impact of the clause, utility executives, when faced with investment decisions, will opt for those generation technologies that are certain to provide a high level of availability. This may be proper in the short run, but it may impede the introduction of new technology in the future.

The clause may also bias decisions concerning the commercial declaration date of new plants. Rather than declare a generation plant commercial, utility executives may opt to continue the testing period and defer the accumulation of AEPUC for a longer period than would otherwise be the case.

Since no discussion of regulatory policy would be complete without a word on the Averch-Johnson hypothesis, it seems appropriate to examine whether, if any, the proposed clauses may have on utility capital-labor ratios. According to Harvey Averch and Leland Johnson, rate of return regulation encourages utilities to invest in excessive amounts of capital. A widespread interpretation of this theory holds that utilities would expand their invested capital and thus their return by excessive capital use or "gold plating." Both the plant availability and operation and maintenance clauses may tend to bias utility investment decisions toward the employment of more capital, or to induce additional Averch-Johnson type effects. The generally held view is that additional capital expenditures on plant increase reliability. If this is true, then the plant availability clause will result in more capital expenditures on plant than otherwise would have been the case. To the extent that capital costs do not fluctuate as much or are more predicta-
ble than labor costs (operations and maintenance), utilities will tend to use more capital under the operation and maintenance clause than would be employed without it. Thus, the effect of both clauses would be a tendency for utilities to increase capital-labor ratios over what they otherwise would have been.

Conclusion

These papers have addressed the problems associated with the development and introduction of incentive schemes into the regulatory process. They have emphasized the importance of creating incentives that discourage the economic waste inherent in the existing regulatory process. Opposition to changing the rules is understandable. One source of opposition is simply administrative inertia. A second is that the proposals are not free of potential defects. However, the only reason for not moving ahead in this area would be that the theory is so underdeveloped and the method of implementation so imprecise that the recognized economic distortions of the present system should be continued. The point is that incentive schemes can be developed that reduce the economic losses associated with regulation as currently practiced.

Notes