NRRI 94-30

MEASURING THE IMPACT OF ALTERNATIVE REGULATORY PRICING REFORMS IN TELECOMMUNICATIONS

Raymond W. Lawton, Ph.D.
Associate Director

Edwin A. Rosenberg, Ph.D.
Senior Research Specialist

Mary Marvel
Associate Professor
The Ohio State University

Nancy Zearfoss
Graduate Research Associate

THE NATIONAL REGULATORY RESEARCH INSTITUTE
The Ohio State University
1080 Carmack Road
Columbus, Ohio 43210-1002
(614) 292-9404

December 1994

This report was prepared by The National Regulatory Research Institute (NRRI) with funding provided by participating member commissions of the National Association of Regulatory Utility Commissioners (NARUC). The views and opinions of the authors do not necessarily state or reflect the views, opinions, or policies of the NRRI, the NARUC, of NARUC member commissions.

EXECUTIVE SUMMARY

At the same time that state regulators are examining and implementing innovative regulatory pricing strategies, telecommunications markets are changing in ways that may challenge the assumptions underlying the pricing reforms. State regulators have designed pricing policies that seek to simultaneously advance universal service and competition. This has often resulted in sets of workable pricing policies built on the uneasy and dynamic tension that exists between universal service and competition. Awareness of these dynamic tensions and the rapidly changing environment has resulted in the expressed need for assistance in evaluating and analyzing the costs, results, and impact of alternative pricing policies. Regulators are more interested in a pragmatic evaluation than in dogmatic ideological defense of a particular pricing reform.

Status of State Regulatory Pricing Reforms

In chapter 2 the results of an updated survey of state pricing reforms by The National Regulatory Research Institute are presented. The number of states with incentive regulation, price caps, flexible pricing, and ratebase/rate-of-return pricing are identified. Other features of the regulatory pricing reforms covered include quality-of service, length of plan, and infrastructure commitments. The survey notes that a third of states with pricing reforms also have identifiable infrastructure agreements. Twenty-four states reported some form of profit or revenue sharing and five states had implemented price caps. Approximately $348 million was shared in 1993 as refunds, rate reductions, and infrastructure improvements. Some thirty-two states reported using some combination of deregulation, pricing flexibility, or detariffing. The survey also found that twenty-one states had explicitly included a quality-of-service feature in their regulatory reform.

While the number and identity of states using different pricing reforms is important, a need also exists for a neutral discussion of the features of these pricing reforms. In a

regulatory proceeding the merit of alternative approaches is often hard to discern when the different parties advance competing views. The survey and analysis in chapter 2 covers each reform component and presents a snapshot of the status of regulatory pricing reform efforts. The program evaluation, economic, and qualitative approaches identified and analyzed in this report can be directly used by state commissions to determine the impact and results of each pricing reform. Armed with accurate evaluative information, state commissions can make better decisions about the need to further refine or modify existing pricing techniques.

Program Evaluation

This report examines various approaches and introduces one relatively new approach, program evaluation, that offers much promise in identifying the success of pricing reforms. The report does not evaluate whether a particular reform has been successful but instead shows how program evaluation techniques can assess the results of pricing reforms.

Program evaluation has been widely used in health care, education, social services, agriculture, aerospace, and defense programs to see if rhetoric meets reality. Program evaluation helps answer the "So what?", "Did it make a difference?", "Did the pricing program do what it was supposed to do?", and, the "What was the impact?" questions that commissioners care about. It does this by systematically comparing results and impacts to the goals of the pricing program or reform. It gains its analytical rigor through the use of methodologies derived from classic experimental design techniques, including the use (where feasible) of control groups. Its policy relevance comes from its explicit linkage of goals and outcomes to a pricing reform. Program evaluations do not try to uncover everything of interest and instead are designed to produce specific goal achievement information needed by decision makers.

Pricing reforms are generally contained in an order or other authoritative document that specifies the timing, affected organizations, the pricing program components, and intent of the pricing reform. Armed with this information it is possible to design a program evaluation that can answer many of the impact and goal achievement concerns of state regulators. Because of the explicit nature of commission orders and the compliance of

jurisdictional utilities, the program evaluation approach is well-suited for evaluating these pricing experiments.

Commissions, of course, are inundated with data and studies analyzing projected and/or actual results of pricing reforms. For instance, a cost-benefit study may be submitted. Without going into the merits of a specific cost-benefit study or the overall usefulness of cost-benefit analysis, it is important to realize that a cost-benefit study does not answer impact and goal attainment questions. It does, however, answer another important question: namely whether the benefits of the pricing reform outweigh the costs. No study method or data set can answer all questions. It is the thesis of this report that the unique set of regulatory concerns about impact and goal attainment can be better answered by using the program evaluation concepts and techniques identified in chapter 3, than by many other analytical approaches available to regulators. Further, as regulators typically see themselves in a transitional period they are especially sensitive to the need to adjust and change pricing policies to meet needs in a changing environment. Program evaluation can provide important data that regulators can use to retain or modify pricing policies.

Because of the difficulties that can arise in using control groups, a sample regulatory pricing evaluation indicator system is presented in chapter 7, which commissions can use to conduct a pricing evaluation with or without control groups. This indicator system would allow a commission to monitor systematically the impacts of a pricing reform. The indicator system is a method for identifying which impacts are to be monitored and the linkage to regulatory goals. It is argued that a workshop-type approach that includes all parties is the optimal way to design and implement a pricing reform indicator system.

Qualitative Analysis

The report discusses the strengths and weaknesses of qualitative analyses in the context of evaluating telecommunications pricing reforms. Several specific techniques are identified that can produce reliable qualitative information and that can be used to critique qualitative studies given to a commission.

The primary advantage of a qualitative approach lies in its ability to let the analyst understand or have an improved insight about a process, event, trend, or a relationship. For example, a testimonial from small startup firms about their need for advanced telecommunications services could properly be dismissed from a quantitative perspective because the sample would not be representative or reliable enough to be used to make generalizations concerning the telecommunication needs of all firms. A qualitative analysis would not necessarily challenge this conclusion, but might look at the testifying firms as a prototype or forecast of the future needs of new telecommunications firms.

Several difficulties arise in using qualitative information to analyze pricing reforms. First, most regulatory analysts are trained in quantitative methods. Yet, a significant amount of information received by commissions is qualitative, consisting of testimonials, scenarios, and propositions. Second, commissioners, senior commission staff, and other relevant policymakers are often more comfortable with the qualitative analyses and information presented to the commission than are commission staff. This can cause communications problems. Last, many regulatory analysts rely on quantitative techniques, just when the environment is changing in ways that may render the rigorous assumptions underlying a quantitative analysis inappropriate. A qualitative approach that stresses insight and understanding may prove more valid in some instances than a more elaborate quantitative method.

Chapter 4 provides tools that regulators can use to test qualitative information. Quantitative studies will not be replaced by qualitative studies. Rather qualitative studies of pricing reforms allow a more holistic and intensive investigation of new, emerging, and rare events and trends. Quantitative studies are significantly superior in terms of replication, the explicitness of the research methods, and the ability to make precise findings.

Economic Perspective

Chapters 5 and 6 appraise pricing reforms from an economic perspective, with special attention to issues surrounding competition. In particular, chapter 5 presents basic decision rules that regulatory analyst can follow in assessing the merits of pricing studies submitted to

the commission. It does not attempt to duplicate the massive and relevant economic literature. It focuses on key features that can potentially overturn or invalidate a study. Often times in the rush of events the absence of standard and basic parts of an economic analysis may be overlooked. This chapter provides robust analytical benchmarks that can be used in assessing an economic study on pricing reforms.

Competition is covered in chapter six and is a key underlying feature of regulatory pricing reforms. although the exact categories differ among the states, states generally characterize telecommunications services provided in their state as being competitive, partially competitive, and noncompetitive. The design and adoption of pricing reforms is often accompanied by a debate over whether the existence of competition must be proven before a service can be classified competitive or whether, based upon trends, a service can be classified as competitive before empirical evidence on the extent of competition is available. This difficult issue is, however, a judgement that regulators must make in the context of their state regulatory goals and economic conditions.

Regardless of the service classification approach chosen in a regulatory reform, regulators have a continuing interest in how to distinguish between partially competitive and competitive services. Commissions are concerned about having information that allows them to walk the fine line between allowing competitive entry without unduly restricting the local exchange carriers' (LECs') ability to respond to competition.

Criteria are proposed in tables E-1 and E-2 for telecommunications services sold in partially and fully competitive markets. Because most segments of the telecommunications business have traditionally been monopolized, proposed criteria are relatively conservative in the sense that markets classified as competitive in telecommunications would be considered highly concentrated in other industries. Even for highly competitive services, other classification schemes, such as the Department of Justice merger criteria would indicate that the market was "concentrated" rather than competitive. These criteria are sensitive to the peculiar history and concentration levels found in the telecommunications markets. They should be useful to state regulators as they provide a robust and realistic set of indicators of the competitive nature of given telecommunications markets.

TABLE E-1

PROPOSED CRITERIA FOR PARTIALLY COMPETITIVE CLASSIFICATION

Partially Competitive Services		
Number of Competitors	*Market Share Criteria*	*Other Criteria*
There must be at least one (but preferably more) viable competitor, unaffiliated with the LEC, which (1) is capable of providing an adequate alternative service and (2) is actively soliciting business throughout the relevant geographic area. The entire geographic area need not be served by the same viable competitor, but each subpart of the area must have at least one such viable competitor.	No service for which the LEC has as much as 70 percent of the market can be considered to be partially competitive. If the HHI can be calculated, it should be no greater than 5,200 for the market to be considered partially competitive.[a]	There must be sufficient evidence, based on acceptance of one or more competitors by the market, that existing customers are choosing the competitors' services over the LEC's offerings and that new customers consider the alternative providers to be effective competitors. Such evidence can be a reduction in the LEC's market share, a reduction in the number of (or percentage of) customers served by the LEC or a reduction in usage of the LEC's service offerings.

Source: Author's construct.

[a] The Herfindahl-Hirschman Index (HHI) used in the table is a measure that is used to determine the size structure of a market by looking at the market shares of all firms.

TABLE E-2

PROPOSED CRITERIA FOR FULLY COMPETITIVE CLASSIFICATION

Fully Competitive Services		
Number of Competitors	*Market Share Criteria*	*Other Criteria*
There must be at least three viable competitors, unaffiliated with the LEC, each of which (1) is capable of providing an adequate alternative service and (2) is actively soliciting business *throughout* the relevant geographic area. The same three competitors need not serve the area uniformly, but the entire area must have at least three active competitors.	No service should be classified as highly competitive so long as the LEC has over 45 percent of the market for that service. Moreover, at least two identifiable competitors must, individually, have at least 10 percent of the market, or one identifiable competitor must have at least 25 percent of the market. If the HHI can be calculated, it should be no greater than 3,000 for the market to be considered highly competitive.	There must be market confirmation that existing customers are choosing the competitors' service over the LEC's offering and that new customers consider the alternative providers to be effective sources of service. Confirmation includes evidence of a reduction in the LEC's market share, number of customers served, or usage of the LEC's services.

Source: Author's construct.

A service that is not either competitive or partially competitive will have a score that is less than the minimum criteria identified in table E-1. A partially competitive service will meet the criteria in table E-1 and competitive services will meet or exceed the criteria in table E-2. This classification scheme is workable and should offer protection to nascent competition.

In many ways regulators are like the captain of a ship with a universal service engine propelling the ship in one direction, while a competition engine pushes the ship in another direction. The captain hopes that by judicious use of the ship's rudder the ship can be steered. The captain knows from experience that a lot of the information available about regulatory shoals, competitive tides, and access to ports and shipping channels comes from unreliable sources having different world maps. It is hoped that the program evaluation, economic, and qualitative techniques examined in this report can help provide the captain with the information and methods needed to steer the ship in a desired direction.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF TABLES

FOREWORD

In the past decade the pace of telecommunications pricing reforms has increased at state and federal regulatory commissions. A wide variety of pricing plans have been considered and implemented. This report analyzes the status of pricing reforms implemented at state commissions and analyzes the different approaches available to assist commissions in evaluating the impact of the pricing reforms on universal service and competition.

Douglas N. Jones
Director
November 1994

ACKNOWLEDGEMENTS

The authors wish to thank Linda Schmidt and Jacquie Shepherd for their excellent typing and layout assistance, Catherine Reed for research assistance, Wendy Windle for preparation of graphics, and Dr. Francine Sevel for editing. The report was measurable improved by the thoughtful review comments received from Dr. Douglas Jones, Kenneth Costello, Charlotte TerKeurst, and Debra Kriete.

CHAPTER 1

INTRODUCTION

Introduction

State and federal regulatory commissions have considered, designed, and implemented a significant number of regulatory pricing reforms since the mid-1980s. These reforms were intended to correct pricing problems derived from commission reliance on traditional ratebase, rate-of-return regulation (RBROR) to set prices for services sold by telecommunications utilities. New pricing policies were adopted in order to further particular state and federal goals. These goals include ensuring affordable pricing of universal service, promoting economic development, maintaining equity, increasing competition, eliminating anticompetitive practices, improving access to advanced information services, and a number of telecommunications-based health, education, and public safety goals.

It would be fair to say that these new regulatory pricing policies were adopted for pragmatic reasons, namely to help a state achieve its regulatory goals. As befits the states' traditional role of being "laboratories of democracy," a wide range of pricing reforms were examined and implemented. As with any good experimenter, states have a strong and practical interest in evaluating the results of their experiments. States want an answer to a "Did it work ?" question and are far less interested in a dogmatic or ideological defense of the merits of a particular pricing reform. For instance, a state is likely to be more interested in determining whether the actual results of its price-caps program show it to be superior to the form of regulation it replaced, than in any academic debate over alternative pricing schemes.

Now that a number of years have passed since some states have implemented new pricing policies, it is appropriate to ask which pricing policies were most effective. A number of states are in policy cycles where they will soon be evaluating existing and proposed reforms. Unfortunately, a large amount of rhetoric, abstract theoretical arguments, and political posturing by affected stakeholders has accompanied evaluation efforts. This makes

an objective evaluation difficult to undertake for a state commission, as each affected party may view any change in the policy as hurting its interests and disproportionately benefiting others.

Before it can be determined which pricing policy is the most effective in achieving state telecommunications goals, it is first necessary to examine the approaches used and available to objectively appraise the results of these pricing policies. This report focuses on how to evaluate, rather than on evaluating a particular reform in a specific state. This distinction is important because if invalid measures and inappropriate procedures are used, or if the evaluation is premature, then less confidence can be had in the results of the evaluation.

Organization of Report

This report is organized in four parts. In chapter 2, a brief overview of the key features of telecommunications pricing reforms is presented. The features of the reforms are examined in a nonevaluative manner, so that the full range of options available to a commission can be seen in a nonadversarial context. In chapter 3 the concept of program evaluation is presented and its use as a regulatory tool is explained. In other areas the program evaluation approach has been used to help governmental agencies determine the impact and success of particular programs. Pricing reform is an ideal candidate for the program evaluation approach because it has a clear start date, a specified geographical area, and its features are explicitly described in an order, a rule, a written agreement, or a piece of legislation. This approach has not been widely used by state commissions in exercising their responsibilities for oversight of telecommunications utilities. Its use should significantly improve the ability of state commissions to get an accurate answer to the "Did it work?" question each state typically asks of its pricing reforms. Use of data produced through this approach could improve, complement, and even partially substitute for some of the data submitted in evidentiary-type hearings and used by state commissions in their reform appraisals.

In chapter 4 the use of qualitative information to evaluate a pricing reform is scrutinized. This examination is important because of the comfort with and reliance on qualitative information, such as testimonials, scenarios, and pilot studies that occurs at many commissions, agencies, legislatures, and governors offices. While qualitative data may seem to be a distant and less reputable cousin to the quantitative data preferred by most analysts, policymakers of all stripes are routinely presented with and do make good use of qualitative data. This chapter reviews issues associated with the use of qualitative data and presents a perspective on its best use in the context of assessing telecommunications pricing reforms.

In chapters 5 and 6 concepts and measures used to evaluate pricing reforms are assessed from an economic perspective. It is a normal part of nearly all regulatory reform proceedings to have econometric, regression, competition, and pricing data and studies submitted. Due to their daunting nature, and because they are cloaked in an aura of scientific precision, these studies often carry great weight. Used and interpreted properly, quantitative studies are extremely useful in identifying impacts and pricing outcomes. These chapters focus, in particular, on use and interpretation; and do so without attempting to duplicate the massive methodological literature that exists. Special attention is paid to the use of various measures of competition because of the linkage of pricing freedom to the degree of competition.

In chapter 7 an illustrative regulatory evaluation performance indicator system is presented, which state commissions can use as a prototype for their own evaluation efforts. Use of this material could save time and money, which a commission could redirect to the achievement of other commission goals.

CHAPTER 2

STATE REGULATORY PRICING REFORMS FOR TELECOMMUNICATIONS UTILITIES[1]

Introduction

Starting in the late 1980s state regulatory commissions have studied, ordered, and implemented a number of important regulatory reforms for jurisdictional telecommunications utilities. These reforms were occasioned by the divestiture and breakup of the old Bell system and the emergence of new and improved telecommunications technologies. During this period, state and federal regulatory commissions (and legislatures) acted to change the existing market structure in ways that promoted the emergence of competition, the use of new technologies and services, and ensured that universal service goals were advanced.

No one state necessarily used all parts of all the possible reforms, yet it was clear that waves of similar approaches were seen occurring at the same time at a number of state commissions. Partial proof of the pervasiveness of these reforms is the familiarity regulators now have with reforms whose names were virtually unknown prior to the late 1980s. These include price caps, exogenous adjustments, profit and revenue sharing, incentive regulation, pricing flexibility, and infrastructure commitments.

The good news is that states in their role as the laboratories of democracy have designed many regulatory reforms; the bad news is that the resulting variety often makes it difficult to compare states and to gain an overview. The differences observed across states are due to the economic circumstances of the states, the evolutionary nature of the reforms, the varying perceptions of regulators, the policies and operations of regulated

---

[1] This chapter is based, in part, on a report prepared under contract to the Idaho Public Utilities Commission. See Raymond W. Lawton, Nancy Zearfoss, and Catherine Reed, *Alternative Forms of Regulation: A Status Report* (Columbus, OH: The National Regulatory Research Institute, May 1994).

telecommunications utilities, and the competitive nature of telecommunications markets in each state.

Unlike a federal or nationwide approach, which often may have all march to the same drummer (and conformity and implementation of rules are the paramount features), state commissions are primarily in a searching, learning, and evaluative mode. States are interested in determining what works, rather than in dogmatically defending a particular feature of a regulatory reform. States use a pragmatic perspective and typically ask the following questions regarding each proposed or implemented reform:

1. Who pays?
2. How is universal service affected?
3. Does it promote competition?
4. Is it equitable?
5. Does it promote innovation?
6. What is the impact on the economy?
7. Does it encourage economic efficiency?
8. How are the disabled affected?
9. Will all parts of the state and sectors of the state's economy benefit?

Although states are in various stages of considering alternative forms of regulation (AFOR), a number of states have had an AFOR in effect long enough to evaluate its effectiveness in meeting state regulatory goals. Often the need to do this is because the pilot or test period, or agreed-to length of the AFOR expired, or is about to. States in a learning and in an evaluative mode are interested in analyzing the features and impacts of their AFOR in order to get information that will assist in designing the AFOR features needed for the next few years.

A number of regulatory reform surveys were conducted over the past few years. Unfortunately, some of them have been designed to promote a particular viewpoint. Additionally, each survey used its own categories such that very few surveys are directly comparable. However, in 1989 Missouri surveyed the states in a format that the NARUC

Communications Committee found useful. This survey was repeated by Maine in 1991 and by The National Regulatory Research Institute (NRRI) in 1992.[2] All three surveys were widely circulated and used by state commissions, utilities, consultants, and other interested parties.

These surveys were updated in April 1994 when the Idaho Public Utilities Commission asked the NRRI if it could quickly update the NRRI 1992 survey and add a few items.[3] In the update about half of the states were called and the 1992 information (along with the new information) was updated. States initially called were those thought to have AFOR changes in 1993 or 1994. Also a literature survey was conducted.

The data in this chapter should be regarded as accurate in terms of painting the big picture and should be sufficient to identify where a state's pricing reforms fit in relation to other states. Given the sensitivity of these and other kinds of responses a careful follow-up is required. It is not expected that the overall picture will shift markedly, rather the improvement will be that the details for each state will have been clarified and confirmed.[4] Because of the complexity of the data and the different ways each state may use a similar term or concept, a separate second phase of the survey is being conducted. The survey will pinpoint the exact features of each state's AFOR. This NRRI report is expected to be available late Fall 1994.

Many state commissions and the District of Columbia commission have made an assortment of telecommunications regulatory pricing reforms. Often these reforms or features of the reforms are known by different names in different states. A freeze in one state may be considered as a price cap in another state. Further the surveys that were used to classify the states had different purposes and used different classification schemes. In this chapter the

---

[2] Vivian Witkind Davis and Nancy N. Zearfoss, *Update to the Maine and Missouri Reports on Alternative Regulation Plans in Telecommunications*, ed. National Regulatory Research Institute (National Association of Regulatory Utility Commissioners, June 1993).

[3] Lawton et al., *Alternative Forms of Regulation: A Status Report.*

[4] For example, it is not clear if the infrastructure investment amounts reported by states are actually incremental expenditures or if they simply represent "business as usual."

most salient features of the pricing reforms are presented, along with a sense of the range of policy options often associated with each reform. The intent of the description is to avoid being prescriptive or favoring any particular reform or feature.

An important problem often occurs when the reforms are examined in isolation in a survey; as an example, the states often have complex pricing policies that include traditional and reform features. A state may have a ratebase system with revenue sharing and flexible pricing. Such a state could be counted one, two, or three times depending upon the sophistication and reporting format of the survey. In this report states are tabulated according to whether or not the feature is included in their pricing policy. This approach could count all three attributes of the state used in the above example.

In surveys conducted to date and from a reading of the regulatory reform literature the most salient pricing reform features are (1) pricing, (2) quality of service, (3) pricing flexibility, (4) price caps, (5) freezes, (6) efficiency incentives, (7) presence or absence of sharing mechanisms, (8) infrastructure commitments, (9) length of plan, (10) review features, (11) standards for classifying competitive services, and (12) exogenous adjustments.

Pricing

The primary focus of AFOR reforms across the country has been in the area of pricing. Two types of pricing reforms were undertaken. The first is pricing for basic or regulated services. The second is pricing flexibility for unregulated and competitive services. At some level the two types of pricing are frequently joined through a bargaining process. Generally, telephone utilities prefer less regulation, but commissions and state legislatures determined that complete deregulation of all services would not be in the public interest due to deleterious effects on competition and universal service. Instead many states have delineated some services as basic and/or noncompetitive and instituted various pricing reforms to set prices for these services. Other telecommunications services are seen as facing enough competition that the utility can be allowed various degrees of pricing flexibility, up to and including complete price deregulation. Each state defines its set of pricing flexibility services differently, but utilities seem to have accepted the bargain and have dealt with the complexity

associated with providing services under different sets of pricing rules. This section of the report deals with pricing schemes for basic services and a later section looks at pricing flexibility.

In tables 2-1 and 2-2, states are identified by holding company and the basic type of pricing reform. Results indicate the following:

- five states have price caps,
- twenty-four states report revenue or profit sharing,
- thirty-two states have pricing flexibility, and
- three states use a social contract approach.

A state may be in more than one pricing category so the totals do not equal fifty, plus the District of Columbia. For instance, in U S West states, one state, (North Dakota) has price caps, six states (Colorado, Idaho, Minnesota, New Mexico, Oregon, Washington) have sharing, thirteen states (Arizona, Colorado, Idaho, Minnesota, Montana, Nebraska, New Mexico, North Dakota, Oregon, South Dakota, Utah, Washington, Wyoming) have pricing flexibility, and one state (Nebraska) uses a social contract.

Price caps are the most recent pricing reform with North Dakota, New Jersey, Rhode Island, Delaware, and California having recently instituted price-cap plans. Price caps were proposed and are under consideration in Ohio, Illinois, Wisconsin, and Indiana.

As discussed in a later section on efficiency incentives, price caps are intended to promote efficiency by removing the "cost plus" incentive of ratebase, rate-of-return regulation (RBROR), and replacing it with a price-caps system such that (ideally) the only way a utility can keep or increase its current profits is by being more efficient. Detailed rules are used to set price floors and ceilings for all covered services.

TABLE 2-1

STATES WITH PRICE-CAP OR SHARING PLANS

	PRICE CAPS		REVENUE OR PROFIT SHARING	
	Jurisdictions	Total	Jurisdictions	Total
Ameritech[a]	- -	0	- -	0
Bell Atlantic	DE, NJ	2	DC, MD, NJ, VA	4
BellSouth	- -	0	AL, FL, GA, KY, LA, MS, TN	7
NYNEX	RI	1	CT, NY, RI	3
Pacific Telesis	CA	1	NV, CA	2
Southwestern Bell	- -	0	MO, TX	2
U S West	ND	1	CO, ID, MN, NM, OR, WA	6
	TOTAL	5		24

Source: NRRI 1993 survey.

Note: [a] Price-cap proposals are under consideration in several states served by Ameritech.

TABLE 2-2

STATES WITH PRICING FLEXIBILITY OR SOCIAL CONTRACT APPROACH

	DEREGULATION, PRICING FLEXIBILITY, DETARIFFING[a]		SOCIAL CONTRACT	
	Jurisdictions	Total	Jurisdictions	Total
Ameritech	IN, MI, OH	3	- -	0
Bell Atlantic	DE, DC, MD, NJ, PA, VA, WV	7	- -	0
BellSouth	TN	1	- -	0
NYNEX	ME, RI, VT	3	ME, VT	2
Pacific Telesis	CA, NV	2	- -	0
Southwestern Bell	KS, MO, TX	3	- -	0
U S West	AZ, CO, ID, MN, MT, NE, NM, ND, OR, SD, UT, WA, WY	13	NE	1
	TOTAL	32		3

Source: NRRI 1993 survey.

Note: [a] Includes all states that reported some deregulation, not only those that clearly have an alternative form of regulation.

As depicted in table 2-3 the primary incentives contained under a pure price-caps plan are sufficient to have a utility engage in any one of five behaviors or strategies.[5] State commissions have designed additional features for price caps to avoid or ameliorate some of the problems that may occur under four of the five basic price-cap strategies available to a price-cap regulated telecommunications utility. Each state's price-caps plan, accordingly, can look quite different.

In Strategy One the utility engages in the same kind of cost control activities typical of many unregulated firms and cuts its costs by layoffs, lowered service quality, and elimination of unprofitable services and facilities. It controls its costs and increases its profits but does not increase its efficiency-increasing investments. In order to ensure an acceptable outcome, a possible regulatory response to this behavior would be to establish or strengthen quality-of-service (QOS) standards. This would have the twin advantages of ensuring that consumers have universal access to reliable and ubiquitous service and encouraging the utility to focus on increasing efficiency.

Strategy Two is for the utility to make exactly the same level of investment in efficiency-increasing investments as it made before. The utility has "done nothing wrong" and charges the same prices as previously and enjoys the same level of profit. As long as inflation is mild and the utility's natural level of productivity meets or exceeds the target set by the commission, the utility and its customers are as well off as under RBROR. A regulatory response here could be to use a "consumer dividend" or to set higher than historically observed productivity improvement rates in order to encourage a utility to increase its efficiency improvement efforts. Without these corrective actions, it would be unclear why price caps would be superior to RBROR.[6]

---

[5] Other incentives may exist that could produce different outcomes.

[6] A subtle but pervasive concern of regulators and public policymakers is that the productivity rates used in price caps are those developed under RBROR regulation. It is a reasonable presumption that if price caps provide a profit-driven incentive to increase productivity rates by using new technologies, the productivity rates used under price caps should be greater than those observed under cost-plus regulation.

TABLE 2-3

FIVE UTILITY STRATEGIC RESPONSES TO PRICE-CAP INVESTMENT INCENTIVES

Strategies	Utility Efficiency Increasing Investment Actions	Result	Original Cost	New Cost	Price	Original Profit	New Profit	Regulatory Tool
Strategy One	Make less investment than normal	Lower cost and increased profit	10	9	11	1	2	Infrastructure plan/quality of service
Strategy Two	Same level of investment as ratebase, rate-of-return regulation	Same cost/price relationship (inflation)	10	10	11	1	1	Consumer dividend
Strategy Three	Increased investment	Lower costs and increased profits	10	7	11	1	4	Profit sharing
Strategy Four	Increased investments	Lower costs, lower prices, and increased profits	10	7	10	1	2	No new tools needed
Strategy Five	Extreme increase in investments	Uncertain: No profit if demand is weak	10	13	11	1	-2	Risk shifting rules

Source: Authors' construct.

Strategy Three is to increase investments and efficiency. The utility could sell its services at the same old price and have increased profits. The undesirable outcome here is that profits may be too high. A well-established regulatory practice is to set limits to ensure that excessive or monopoly profits are not earned. Here the rationale is not to capture profits in a redistributive mode but rather to make sure that the increased profits are due to the core efficiency-increasing incentives of price caps. If increased profits are due to artificial entry barriers rather than increased efficiency, sharing of these nonefficiency-driven profits is thought to be in the public interest. Further, there exists a widespread but somewhat fuzzy belief that there is a politically defined upper limit for utility profits. Profit sharing defuses this issue and directly benefits ratepayers.

In Strategy Four everything works out according to plan. The utility makes increased investments in efficiency, lowers its costs, lowers its prices, and has a higher level of profits. These outcomes occur when the utility's investment preferences, and pricing practices as derived from the incentive structure, are sufficiently close to those of society such that the desired outcomes occur. This congruence is unlikely enough that existing price-cap plans contain incentives beyond the initial profit incentive to encourage pricing and service outcomes beneficial to society. Benefits here also include the need to make sure that anticompetitive conditions do not inadvertently arise that would delay the introduction of widespread competition in telecommunications markets.

In Strategy Five a utility makes an extreme increase in investments such that its installed infrastructure is not financially viable. In the example shown the utility loses money because when using the old price as a de facto cap, it cannot increase its prices to the higher level needed. The regulatory response to this is problematic and would likely focus on risk shifting within the boundaries established by case law.

States concerned that the core price-caps efficiency incentive is not strong enough will tend to have more regulatory tools in order to promote state goals. To date, in order to ensure that desired outcomes actually occur, nearly all price-cap plans used some of the additional tools identified in table 2-3. Further, as commissions begin to better understand the dynamics of the price-cap incentives, more regulatory tools were developed to protect consumers and to advance the larger public interest.

Infrastructure Commitments and Achievements

In 1993, thirteen of the thirty-six states with AFOR plans appeared to have clearly identifiable infrastructure upgrade agreements. In table 2-4 the states with these plans are listed along with the timing and dollars involved. Inspection of the table (supplemented by conversations with staff of the surveyed commissions) shows a number of interesting features of the plans, as follows:

- Amount. The amounts listed differ quite a bit both absolutely and whether or not any amount is specified. The most frequently chosen option is one where the amount is "not specified."

- Timing. Some plans envision a set amount of money being spent each year, while others focus on a net sum that will be spent over the period. Some are front-loaded with a significant amount of the modernization investment being made in the early part of the plan. In New Jersey and Tennessee plans have specific timetables.

- Length. The plans show some variation in the length of the time period. Generally the length of the plan appears to be longer than the regulatory reform. That is, a reform may have a three-year life, but the infrastructure plan is five or more years in length due to deployment and revenue constraints. In New Jersey, for example, integrated services digital networks (ISDN) will be made available everywhere in eight years and broadband will be available in eighteen years. The average plan appears to be about five years in length.

TABLE 2-4

STATES WITH INFRASTRUCTURE UPGRADES

State	Amount of Upgrade
Kansas	$160 million over 5 years
Nevada	$0.465 million over 5 years
New Jersey	$1.5 billion over 7 years
West Virginia	$450 million over 5 years
Vermont	$284 million over 5 years
Tennessee	$400 million
Missouri	Investment level uncertain
Rhode Island	Investment level uncertain
Texas	Investment level uncertain
Washington	Investment level uncertain
Wisconsin	Investment level uncertain

Source: NRRI 1993 survey

- Incremental. Most plans are unclear as to whether or not the amounts shown are incremental increases to originally planned efforts or whether they simply represent amounts that would have been spent under a "business as usual" perspective. Some suggest that the modernization plans predominantly reflect amounts that would have been spent absent any regulatory reform. On the other hand, in New Jersey the infrastructure plan represents a 20 percent increase over "business as usual."

- Detail. A number of plans are lacking in detail about exactly which equipment or facilities will be installed at specific locations. Others, such as the Tennessee Plan, have specific details and timetables for deployment. This is partially due to the

need to protect proprietary deployment data from competitors. It also reflects the interest of the local exchange company (LEC) in having freedom to respond to demand and changing economic situations, as opposed to being locked in to an official and rigid deployment schedule. It may also be based on a commission's desire to give the utility the freedom it needs to respond to various competitive forces. A commission may also choose to avoid specifying details because of the monitoring and oversight costs implied. As shown in table 2-5 most states with explicit infrastructure agreements rely on annual filings or reports to monitor activity. Monthly reports are required in Washington and Texas.

- Strength of Infrastructure Investment Commitment. This is perhaps the most difficult and elusive part of any infrastructure agreement. Some plans contain commitments that the company will invest a specified amount over a given time period. Some only specify an amount. Other plans are silent on amounts or timing. Plans with commitments are often thought to be better because the company has explicitly committed funds to be invested in infrastructure modernization. In Ohio, Ameritech said that it would make specific commitments to make infrastructure investments. In Kansas and Missouri, commitments may be required for $65 million and $200 million respectively. In Pennsylvania, legislation says companies seeking alternative regulation are to commit to a 100 percent broadband network by the end of 2015.[7]

[7] *Telecommunications Reports*, "Missouri Advances Bills Requiring SFAS-105 Acceptance; House Substitute Bill Promotes Price Regulation Plans," *Telecommunications Reports* 60, no. 14 (April 4, 1994): 3-4; "Kansas Legislature Extends SW Bell Regulation Plan," *Telecommunications Reports* 60, no. 14 (April 4, 1994): 30; and "Bell of Pa. Submits Competition and Modernization Plan," *Telecommunications Reports* 59, no. 41 (October 11, 1993): 15.

TABLE 2-5

REPORTING ON STATUS OF INFRASTRUCTURE UPGRADES

Utility	State	Type of Reporting
1. Ameritech	Wisconsin	Annual filings with commission
2. Bell Atlantic	New Jersey	Annual reports
3. Bell Atlantic	West Virginia	Not clear
4. BellSouth	Tennessee	Annual report to commission
5. NYNEX	Rhode Island	Semi-annual reports and meetings with Division of Public Utilities and Carriers
6. NYNEX	Vermont	Monthly financial reports
7. Pacific Telesis	California	Compliance filings
8. Pacific Telesis	Nevada	Reports to commission
9. Southwestern Bell	Kansas	Not clear
10. Southwestern Bell	Missouri	Quarterly progress report
11. Southwestern Bell	Texas	Filing of monthly reports
12. U S West	Washington	Monthly service indicator reports

Source: NRRI 1993 survey.

It is not certain, however, exactly what a commitment means. Rhetoric surrounding commitments suggests that these commitments are binding agreements. In practice they may not be very binding as it would be unlikely that a commission would want to let a company become overextended—that is, let its modernization investments exceed its ability to recover revenues—simply to conform to what may prove to be an uneconomic timetable. Alternatively, a commission would be unlikely to want to slow down an even faster deployment if economic conditions warranted a faster deployment.

If part of the regulatory bargaining accompanying a regulatory reform is predicated upon the inviolate nature of infrastructure commitments, this feature may need to be reexamined. Commission compliance monitoring and reporting can ease some of this problem but may require more oversight.

- Urban/suburban/rural. Intrinsic to achieving their universal service goals, state commissions have been concerned with modernization plans that accelerate the deployment of digital switching, fiber, and the various advanced telecommunication services. The underlying tension contained in each plan is the timing differences that occur between when urban/suburban/rural areas receive the benefits of an advanced infrastructure. The Tennessee Plan is widely regarded as one of the most sophisticated plans and recognizes the urban/suburban/rural dilemma. In Tennessee, broadband capability is planned to start in urban counties in 1995, suburban counties in 1997, and rural counties in 1999. The Plan envisions broadband penetration of 10 percent in urban areas, 5 percent in suburban, and 2 percent in rural by the year 2000.

  Underlying economics result in the deployment pattern shown in table 2-6. The issue before a state commission is how fast the deployment occurs and not the basic urban-first pattern. To the extent that deployment of a modernized infrastructure is significantly accelerated beyond a reasonable payback period, states may take offsetting actions that increase utility revenue streams.

  In table 2-6, the core public policy problem is that although revenues and payback periods for the "first recipient" are more than adequate to obtain advanced services and facilities, this is not the case for the second and third recipients.

  The payback period is longer and the revenues stream is weaker for the "second recipient" and so deployment is delayed because of the utility's need to economically justify its investments. It is the "third recipient" that causes the most serious and complex problem. This is because using standard computerized "prove-in" models, utilities deploy here last, whereas this customer classification may contain the largest number of customers and may cover the largest geographical

TABLE 2-6

GENERIC DEPLOYMENT PATTERN FOR NEW SERVICES OR TECHNOLOGIES

	Order New Services or Technology Received		
Customer characteristics	First Recipient	Second Recipient	Third Recipient
Location	Urban	Suburban	Rural
Customer class	Business	Institutional	Residential
Customer type	High income	Medium income	Low income

Source: Author's construct.

part of a utility's service area. Speeding up the deployment may lead to pricing and costing decisions that use "first recipient" revenues to assist accelerating deployment to the second and third recipients. These actions may interfere with the development of competition at the local exchange level. Bill 1289 in California illustrates this problem and authorizes the Commission to permit noncost-effective investment to ensure service to noneconomic and underserved areas.[8]

Some of these issues are expressed in the debate over whether deployment should be supply or demand driven. In a supply deployment, a utility constructs facilities significantly ahead of demand. In a demand deployment facilities are installed only to meet the demand foreseen in a fairly short planning period. This approach is slower, incremental, and only achieves ubiquity when matched by ubiquitous demand. It is the type of deployment followed by unregulated firms.

[8] *Telecommunications Reports*, "Two Infrastructure Bills Clear California Legislature," *Telecommunications Reports* 59, no. 38 (September 20, 1993): 8.

In practice, utilities have actually followed a supply deployment, modified by a "prove-in justification" approach. This means that utilities do not uniformly, ubiquitously, and instantly saturate their entire service territory with the most modern and advanced facilities available. Rather, a computerized analysis is made on a local or central office basis where equipment is installed that will pay for itself within a specified payback period. QOS standards may effectively require longer payback periods or other adjustments in order for a ubiquitous supply of basic telecommunications service to be provided. Then, the practical difference is the length of the payback period and the different deployment pattern that results.

- Cost. It may be difficult to compare the amounts to be invested with the past investments of the same utility. This is because the cost of digital switching, glass fiber, and supporting components has been rapidly dropping. Simply stated, a dollar buys more switching and fiber than it did five years ago. The same amount of investment may yield more infrastructure and, possibly, more advanced features.

  This makes comparison somewhat more difficult as it is necessary to know whether the amounts are being expressed in terms of the purchasing power of 1994 dollars or in absolute amounts. Ameritech, for instance, developed a multi-year purchasing plan. In the plan, vendors make "bids" that set prices for major equipment purchases. This approach used the combined purchasing power of all of the Ameritech states to presumably get lower prices than might otherwise be available under single year bids or efforts by single states. Comparison, accordingly, between years, or before and after the multi-year purchasing plan, becomes somewhat more difficult.

- One wire. Infrastructure plans are built around a utility's vision of a one-wire world. If multiple providers are willing and able to deploy glass fiber so that every consumer has unfettered access to multiple wires, commission interest in the infrastructure plans of one (albeit large and incumbent) provider could decrease

accordingly. Because of universal service goals and the urban-first deployment pattern, commissions and telecommunication utilities have overlapping (but not identical) goals in ensuring that at least one high-capability wire reaches all consumers.

Each commission with an infrastructure plan has implicitly addressed in the plan the trade offs it made among various competing goals, such as the unleashing of market forces (by having multiple wires), ensuring adequate revenues to achieve any stated deployment goals, advancing universal service, and preventing anticompetitive behavior.

- Type of regulation. It is not clear that the form of regulation is the definitive factor affecting a utility's infrastructure deployment. In table 2-7, Ohio, a RBROR state, is compared to the four states that had price-caps plans in effect. As can be seen, Ohio does at least as well in this regard as the price-cap states measured. The same conclusion can not be drawn from data provided in table 2-8, where Ohio is compared to thirteen states with various AFORs in effect. Together these tables suggest an uncertain relationship necessarily exists between the type of regulation and infrastructure deployment.

  Wall street analysts and other informed observers have said that the goals of the holding company, terrain, and the economic strength of demand for telecommunications services are also definitive factors influencing deployment. It can also be argued that no telecommunications utility can afford to stand by and allow others to deploy first or to provide superior service offerings simply because of a disagreement over the type of regulation.

TABLE 2-7
1993 STUDY COMPARING THE FOUR PRICE-CAP STATES

	Rank of Ohio, a Ratebase, Rate-of-Return State Compared to Four Price-Cap States
Digital switching	1
Fiber channels	2
ISDN switches	3
Fiber links	3
Signal system seven (394)	3
ISDN lines	4

Source: Vivian Witkind Davis, Raymond W. Lawton, Edwin A. Rosenberg, Nancy N. Zearfoss, *An Analysis of Selected Aspects of Ohio Bell Telephone's Application for Alternate Regulation,* (Columbus, OH: NRRI, February 1994), 195-196.

Note: [a] Delaware was not included in the analysis.

TABLE 2-8
1993 STUDY COMPARING THIRTEEN STATES WITH ALTERNATIVE REGULATION PLANS IN EFFECT

	Rank of Ohio, a Ratebase, Rate-of-Return State Compared to Thirteen States With Alternative Regulation
Digital switching	8
Fiber channels	5
ISDN switches	11
Fiber links	10
Signal system seven (394)	11
ISDN lines	11

Source: Vivian Witkind Davis, Raymond W. Lawton, Edwin A. Rosenberg, Nancy N. Zearfoss, *An Analysis of Selected Aspects of Ohio Bell Telephone's Application for Alternate Regulation,* (Columbus, OH: NRRI, February 1994), 195-196.

Profit- and Revenue-Sharing Mechanisms

The AFORs developed in twenty-two states (see table 2-9) have some form of profit or revenue sharing. Part of the logic of including sharing is based upon the principle that a part of any additional profits earned or revenues gained by a telecommunications utility is due to the efficiency-increasing investments made by the utility and another part is due to the consequences of the commission-awarded utility franchise. Equity considerations require that the gains be shared accordingly. In the case of price caps, the first price-cap plan for British Telecom did not include any profit sharing. The plan did, however, have a price adjustment factor that started at 3 percent in the early 1980s and is now at 7.5 percent. It is widely understood that the increase in the adjustment factor was partially to decrease British Telecom's profit level.

The AT&T price-cap plan at the Federal Communications Commission (FCC) does not include profit sharing, but does include a "consumer dividend" that in effect raises the hurdle that AT&T must pass before it is fully in a profit-making mode. The FCC offered the LECs two sharing levels: (1) a 3.3 productivity offset has a 12.5 percent sharing level and (2) a 4.5 percent offset has a 13.5 percent level.

As of February 1994, three of the four states with price caps had profit sharing. In North Dakota no provision for profit sharing is contained in the legislation. In California sharing occurs based on a market-based return that is higher than previously authorized rates of return. The California Commission established the higher level to match the greater risks faced by the utilities. Essentially a 13 percent benchmark was created where 50/50 sharing would occur. One hundred percent of all earnings over 16.5 percent would be returned to ratepayers as credits on future bills. In Rhode Island 50/50 sharing occurs at 12.25 percent return on equity (ROE) and 100 percent at 15.7 percent. New Jersey, the fourth price-caps state, has 50/50 sharing above a 13.7 percent ROE.

Profit sharing occurs for the price-cap LECs and three of four price-cap states. Additionally, a de facto profit sharing occurs for British Telecom, and to some extent for AT&T via the consumer dividend.

TABLE 2-9

STATES WITH REVENUE OR PROFIT SHARING

Company	States	Total
Ameritech	None	0
Bell Atlantic	DC, MD, NJ, VA	4
BellSouth	AL, FL, GA, KY, LA, MS, TN	7
NYNEX	NY, RI[a]	2
Pacific Telesis	NV, CA	2
Southwestern Bell[b]	TX	1
U S West	ID, CO, MN, NM, OR, WA	6

Notes: [a] Connecticut had profit sharing from 1991-1993. It was eliminated in 1993 and had produced no sharing.
[b] Missouri had profit sharing, but it expired December 31, 1993.

Source: 1993 NRRI Survey.

Sharing also is included in incentive regulation, freezes, and flexible pricing states. While the pricing plans in these states are not as elegant as the price-caps approach, the sharing notion is predicated upon the same rationales. The very valuable rights-of-ways that a regulated utility has allows it to be more efficient than would be the case if it had to act as a private party and seek all of the rights-of-ways or access to rights-of-ways that it would need in order to achieve ubiquitous coverage of a service territory. Other valuable "assets" provided by the public that increase a utility's ability to operate efficiently include protection from antitrust rules, entry barriers for competitors, commission approved QOS standards that protect the utility from service complaints, a court-sanctioned protection to earn a fair rate of return, and the economic stability provided by having a long-term state-awarded franchise. To gain further appreciation of how valuable these assets are, consider how genuinely difficult (and expensive) it is for cable television companies to negotiate and renegotiate franchises

with hundreds of local communities (when each may have somewhat different terms and conditions).

As noted above, explicit formulas were developed that specify the amount of sharing that will occur based upon specified levels of profits or revenues. In developing the sharing levels commissions implicitly tried to provide a financial incentive (such as higher returns) that encourages a utility to make more efficiency-increasing infrastructure investments so that a wide array of telecommunications services can be ubiquitously provided.

A distinction is made between profit sharing and revenue sharing. Profit sharing was the first sharing concept and is based on return on equity, and identifies sharing levels as described below in table 2-10.

Based upon the above example, no sharing occurs for a utility that does not "cross the 10 percent level." At each level above 10 percent a specified amount of sharing occurs. In Connecticut, for example, the sharing threshold was never reached and no sharing occurred. GTE of California shared some years and had a rate increase one year. The ratepayer portion may be received in the form of credits to future bills, reduced rates, or in the form of commission-approved and specified infrastructure or service upgrades.

A revenue-sharing state follows a slightly different logic and recognizes that profit sharing may be a second-best way to provide a balanced and equitable set of incentives. Revenue sharing accepts the principle that utility profits can be increased (and be shared), but the profit increases may not necessarily be exclusively due to efficiency incentives. Just as a utility could control expenses and increase its profits without any growth in revenues in a profit-sharing system, revenue-sharing states have effectively said that an increase in demand as measured by revenues is a more appropriate sharing mechanism. These states want the utility to not only have an incentive to be more efficient but to have pricing, marketing, and deployment policies that increase the availability and use of information age telecommunications services. Revenue-based incentives are thought by some states to be the best way to encourage the utility to provide affordable and advanced telecommunications services. The actual mechanics of the sharing calculation may look quite similar to the profit sharing example presented above. Both profit and revenue sharing have explicit methods to determine the profits or revenues available to be shared.

TABLE 2-10

ILLUSTRATIVE PROFIT-SHARING LEVELS

Profit Level	Sharing That Is To Occur
10% or below	100% shareholder 0% ratepayer
10-11%	75% shareholder 25% ratepayer
11-12%	50% shareholder 50% ratepayer
12-13%	25% shareholder 75% ratepayer
13% and above	0% shareholder 100% ratepayer

Source: Authors' construct.

In a number of states, revenues and profits have been shared since the inception of the sharing plans. As would be expected these amounts received considerable attention in the media, specialized telecommunications newsletters, and from the members of the regulatory and telecommunications communities. Attention has largely been on the amounts shared, and this seems to generate an interest in comparing amounts shared in one state with those in another state. Table 2-11 displays some recent sharing reported in telecommunications newsletters and shows that twelve of the twenty-four states with sharing reported that sharing occurred in 1993. The total sharing reported was $348.48 million. Rate reductions and refunds were used in eleven of the states to implement the sharing, while infrastructure upgrades and rate credits were used in Idaho.[9] Oregon, another revenue-sharing state, had $8.3 million available and used a refunding mechanism for sharing.

---

[9] *Telecommunications Reports,* "Washington Proposes Changes in U S West's Sharing Plan," *Telecommunications Reports* 59, no. 40 (October 4, 1993): 12; and a telephone interview.

TABLE 2-11
AMOUNTS PRODUCED BY SHARING PLANS IN 1993

State or Company	Amount Shared (millions)	Type of Sharing
Pacific Bell[a]	$107.8	Returned in 1994 as monthly reduction on bills
GTE California[a]	100.5	Returned in 1994 as monthly reduction on bills
Alabama	15.7	Rate reduction
Kentucky	9.0	Rate reduction
Louisiana	21.6	Rate reduction
Mississippi	18.5	Rate reduction
Minnesota	3.7	Refunds
Nevada	0.58	Refunds
Rhode Island	1.8	Refunds
Texas	16.9	Refunds
Washington	33.2	Refunds
West Virginia	5.2	Rate reductions
Oregon	8.3	Refunds
Idaho	5.7	Infrastructure improvements and credits
TOTAL[b]	$348.48	

Source: *State Telephone Regulation Report*, "Ratepayers Gain $348M Under 13 States' Alternative Systems," *State Telephone Regulation Report* 12, no. 7 (April 7, 1994): 1 and 9-10; and telephone survey.

Notes: [a] Funds came from price-cap indexing of basic monopoly services under the California Plan. Neither telephone company had earnings large enough to trigger profit sharing.

[b] Rate case orders from traditional rate cases and earnings investigations in Wisconsin, Texas, Tennessee, South Carolina, New York, Illinois, and Florida reduced rates by $207.9 million during 1993. Rate increases of $39.9 were granted in Delaware, Missouri, New Mexico, Utah, and the District of Columbia.

The 1993 amount is greater than the (incomplete) data for 1992 shown in table 2-12, which indicates approximately $86 million available for sharing.[10] In one state with a sharing plan, the Maryland Commission ordered the telephone company to cut its rates by $28.6 million. Unfortunately, while it is tempting to draw generalizations from these data, it would be difficult to have much confidence in the results. The utilities in each of the states listed in table 2-11 differ in terms of their infrastructure investment plans, cost of service, whether or not an adjustment was made in the prices at the beginning of the AFOR, the amount of pricing flexibility, the strength of the economy in their service territory, and their sharing levels. Because each of these factors can effect the revenues or profits, simple comparisons should not be relied upon.

A plain "it can not be compared" recommendation is not, however, likely to be very satisfying to anyone. The amounts can be compared if an indepth analysis can be initiated. Twelve useful classifications which would be helpful in drawing a valid conclusion about the sharing levels in different states are presented below:

1. revenue v. profit v. nonsharing,
2. amount of and per/capita infrastructure expenditures,
3. depreciation expenses,
4. operating costs,
5. assorted general economic indicators relevant to their service territories,
6. sharing levels,
7. amount of oversight of sharing data,
8. degree of pricing flexibility,

---

[10] States with traditional RBROR also received "sharing" during this same time period. Wisconsin Bell, for example, was ordered to refund $22,500,000 to its customers in 1993. *Telecommunications Reports*, "Wisconsin Bell to Begin Refunding $22,500,000 to Customers," *Telecommunications Reports* 59, no. 13 (March 29, 1993): 40.

TABLE 2-12

AMOUNTS PRODUCED BY SHARING PLANS IN 1992

State or Company	1992 Amount Shared (million)
Alabama	$18.84
Kentucky	10.6
Louisiana	12.5
Mississippi	18.0
Nevada	N.A.
Oregon	N.A.
West Virginia	N.A.
Washington	20.0
Pacific Bell	11.8[a]
GTE California	<11.07[b]>
Idaho	5.1
TOTAL	$85.77[c]

Source: *State Telephone Regulation Report*, "Ratepayers Big Winners Under 17 States' Alternative Regulation," *State Telephone Regulation Report* 11, no. 4 (February 25, 1993): 1-3; and (October 4, 1993), 12.

Notes: [a] Pacific Bell's 1992 sharing would have been close to $90 million without the change in accounting for retirement benefits.
[b] No sharing occurred and rates were raised by $11 million.
[c] Rate increase not included in total.

9. competitiveness of its telecommunications markets,

10. adjustments made in starting pricing levels,

11. amount shared in previous years, and

12. length of time AFOR has been in effect.

A comparative analysis using these categories would allow some conclusions to be drawn. It is not clear how available the information required would be and whether pricing flexibility data could be easily compared.

A commission can do its own indepth analysis of its own state data. Although this might not be comparable to other states, it would at least have the advantage of being doable and more easily interpreted. Unless one of the above suggested classifications for the state changed, some confidence could be had that "apples were being compared to apples."

Quality of Service

State commissions have long been concerned about the quality of the telecommunications services received by business and residential consumers. This concern has been explicitly expressed in the orders and various agreements in state AFORs. States appear to want to ensure that the economic incentives contained in price-cap, profit or revenue sharing, pricing flexibility, and other types of incentive plans do not cause the QOS provided to decline, i.e., that a utility trade off quality for profits. Telephone utilities generally respond to this issue by saying that it would be economically irrational for them to allow their greatest single asset—their network—to deteriorate when faced with competition.

Twenty-one states were identified in a NRRI survey (see table 2-13) as explicitly including a QOS feature in their regulatory reform. Some states may have also had this as a consideration in their reform effort, but they did not necessarily include it in a final order.

Looking at the QOS policies of the states some differences are observable due to the history of the state, the regulatory goals, and the salience of QOS issues. An overview of the range of policies is presented below.

TABLE 2-13

NUMBER OF STATES WITH ALTERNATIVE REGULATION HAVING IDENTIFIED SERVICE QUALITY STANDARDS

Companies	Number of States
Ameritech	1
Bell Atlantic	3
BellSouth	4
NYNEX	4
Pacific Telesis	2
Southwestern Bell	2
U S West	5
TOTAL	21

Source: NRRI 1993 Survey.

- A state may have some type of requirement that the utility must continue to meet QOS standards during the time period the reform is in effect in order to remain eligible to use the reform. This statement of principle signals a commission's intent that service quality should not decline, but this policy may be difficult to implement. This is difficult because a commission might face many problems in reinstituting RBROR in order to correct even a serious QOS problem. Other less far-reaching remedies, however, may be available. In New York, NYNEX's ability to retain earnings over a 12.2 percent rate of return depends upon NYNEX's ability to meet certain QOS targets. Failure to achieve targeted levels could also result in penalties.

- Some states have simply stated that they would continue to monitor QOS.

- Some appear to have a complaint-driven monitoring process, while others have oversight systems that are based on standardized reporting routines. A complaint-driven state typically does not require reports or monitor quality directly. Commission staff and consumers can file complaints, which causes an investigation to be initiated. Some states do not have explicit standards and treat each quality complaint on its own merits.

  States with reporting requirements also follow-up on specific quality complaints but already have data that was filed by the utility. A commission may choose either to be reactive or proactive with the QOS data. A reactive strategy would wait for complaints and use the existing data as a starting point for resolving a complaint. A proactive approach would have staff monitoring the data to ensure that minimal QOS standards are being met. A proactive strategy may include having staff initiate actions that improve the existing level or that restore quality to acceptable levels.

- States have set-up monitoring schemes that check conformance with a set of new quality standards. These states may have updated or reaffirmed their QOS standards as a part of their reform activities.

- Some states have announced a process or an intent to upgrade their QOS standards after their reform begins. Here it appears that the state recognizes that there are technical engineering, competitive, market structure, and universal service issues of sufficient complexity as to warrant a separate proceeding. In these instances existing standards and reporting routines are kept in place until new standards are adopted.

- An automatic QOS adjustment factor was been proposed by utilities for some price-cap regimes. The basic price-caps formula generally does not include an automatic QOS adjustment. Quality issues or problems are handled by most price-cap states

without direct reference to the price-caps formula. Ohio Bell/Ameritech proposed that the following price-caps formula be used with an automatic service quality adjustment:

% change in price cap	=	change in inflation (GDP-PI)	-	% historical productivity offset	+/-	service quality adjustment	+/- exogenous impacts adjustments

In this formula the range of adjustment to the change in the price cap is increased or decreased automatically by the service quality score. If service quality is high, then a positive adjustment (expressed as a percent) would be made such that the cap would increase more than would otherwise have been the case. If the service quality is low, a negative adjustment would be made so that the percentage change allowed in the price cap is less than it would otherwise have been. The intent is to reward the utility with a higher price cap when service is good and to punish it with a lower cap when QOS is poor. A separate commission-approved chart, table, or formula could be used to determine the automatic QOS adjustment score to be used for each level of observed quality. The primary advantage of this approach is that it makes adjustments automatically without the need for separate investigations. The main disadvantage is that it may overcompensate the utility for achieving QOS levels that should be expected to naturally improve with the introduction of an improved technology. The most direct way to eliminate the overcompensation problem is to eliminate the positive incentive and to retain only the negative portion. This approach avoids distorting the basic price-cap efficiency incentives. In a nonprice-cap regime, an automatic adjustment clause would need to be examined in relation to the specific features of that state's reform in order to see if a similar distortion of efficiency incentives occurs.

Efficiency Incentives

Primary Efficiency Incentives

All pricing schemes contain efficiency incentives. AFORs and RBROR are no exception. Price freezes, price caps, pricing flexibility, and incentive regulation each has incentives that are different from traditional RBROR. These incentives are also different than the incentives for unregulated telecommunications firms. Listed below are the primary efficiency incentives for each of the major pricing reforms. AFOR states generally considered the incentives in each approach and chose the one that best met the needs of their state.

1. RBROR. Cost-plus nature gives an incentive to over invest in infrastructure in certain circumstances, to inadequately control operating expenses, and to minimize risk and innovation because profits are constrained. It does, however, contain a positive incentive to deploy and recover the cost of infrastructure through adequate earnings.

2. Price Caps. Price caps are designed to have increased efficiency because the main way that utilities can retain or increase their existing profit level with a price cap is to be more efficient than they previously were under RBROR. They must be more efficient than the general change of productivity-adjusted price increases for the economy as a whole.

3. Price freezes or moratoriums. Because prices cannot increase utilities must control costs or stimulate demand, such that they are more efficient than under RBROR.

4. Pricing flexibility. The incentive structure here has two parts. First the existing services are split (in various ways) between services without pricing flexibility and those with pricing flexibility. A different set of pricing rules is put in place for each. Generally, basic local exchange service for residential and small businesses are in the first group and competitive-type services are in the second group. In the basic group the incentives are driven by the type of regulation. In the pricing flexibility group, the economic incentive is to be more efficient than the utilities' competitors.

5. Incentive regulation. A very wide range of plans and features are grouped together under the rubric of incentive regulation. The overriding economic incentive is to respond to the particular features of the incentive plan in your home state and to have profits greater than those allowed under RBROR.

6. Sharing. While not truly a pricing plan, sharing is generally accompanied by some changes in the existing pricing rules. Often this means that a utility has pricing flexibility for competitive services, and the sharing means that if the designated competitive services are being sold in a market that does not have actual price competition, then the sharing mechanism functions as a partial disincentive to inappropriate pricing of regulated and unregulated services.

Secondary Efficiency Incentives

The various AFOR plans also contain other features that are designed to give additional efficiency incentives to the utility. These are secondary incentives that reinforce the primary economic incentives chosen by the state. These include base price adjustments, exogenous changes, reporting routines, infrastructure plans, and length of the AFOR. These items are examined below.

Exogenous Adjustments

A utility may be allowed to make exogenous adjustments under its AFOR. An exogenous adjustment allows a utility to petition a commission to have the financial impact of an extraordinary event included in and accommodated to in the utility's prices, and any revenue or profit-sharing obligations. Candidate exogenous events may include federally initiated separations changes, tax changes, acts of God, and accounting changes initiated by a regulatory body or the Financial Accounting Standards Board (FASB). The rationale for allowing these changes is that events in these categories are considered outside the ability of the utility managers to control and as such can not be directly effected by the efficiency-increasing incentives found in AFORs. Generally commissions indicated that it would be inappropriate to use exogenous adjustments to insulate the utility from every event in the exogenous categories. Instead, the logic of the exogenous approach has been to allow the utility or the commission staff to ask that an adjustment be made for "big ticket items" effecting the utility in one of the identified exogenous categories. All five states with price caps have an exogenous adjustment mechanism: New Jersey, Rhode Island, Delaware, North Dakota, and California. Six AFOR states allow for the possibility of exogenous adjustments.

States differ somewhat as to who can request an exogenous change, whether or not the exogenous impact must cross a specified threshold, who has the burden of proof, whether the adjustment is included in the price-cap formula, the timing of the exogenous adjustment, and the type of oversight employed by the commission.

Base Price Adjustment

A base price adjustment is an important efficiency incentive. As these adjustments are generally downward, they send a clear signal to the utility that it needs to be more efficient if it wishes to prosper. The logic flow here assumes that if AFOR reforms are needed, then some inefficiencies may already be included in existing prices. An adjustment derived from a cost study is typically used. Fourteen AFOR states have had a base price adjustment. Three out of four price-cap states report a base price adjustment.

Infrastructure Plan

As noted previously, a number of states have infrastructure plans. These plans serve as a benchmark against which the actual deployments can be compared. This functions as an incentive to the utility to deploy its infrastructure as promised in order to avoid criticism and more oversight in the future. The plans differ in terms of whether a formal document exists, whether actual commitments are made, and the amount of detail. These features strengthen the incentive but also increase transaction costs and may limit flexibility.

Length of AFOR

The length of the plan can function as an incentive, as can the structure of the renewal or evaluation process. Unfortunately, it is not clear how the incentives actually work. Some could argue that a long plan favors the utility, as it can have less oversight and push off any complaints by saying, "Lets wait til the end of the plan to talk about revisions." This is a powerful argument as it takes some time to see any reliable impact information. In this line of reasoning the utility receives assurance about the rules for a set time period and can act accordingly. This is a positive economic incentive for the utility. A counter argument is that only the utility benefits from the regulatory lag that occurs when the utility is granted a multi-

year reform. The thought here is that the reform is a compromise effected at the beginning of the time period and that circumstances may change and impacts may be adverse, but the utility can not have any rule changes. In this instance, a long time period would be a disincentive. These two opposing arguments may be unresolvable, other than from the practical need to look at AFOR plans at three or five year intervals and to continue to have commission oversight. The need for oversight is, in part, because both sets of incentives operate simultaneously and simple prudence dictates some monitoring. A true-up of sorts occurs when hearings or procedures are implemented to address continuing the AFOR. The strength of the positive and negative incentives can then be played out in a public forum.

Contextual Efficiency Incentives

The above primary and other efficiency incentives are muted or enhanced by a number of other factors. Some of these factors apply somewhat uniformly across all telecommunications providers and some do not. These incentive factors include

1. universal service obligations,
2. payment into universal service funds,
3. ability to receive universal service funds,
4. net inflow/outflow of universal service funds,
5. amount of sustainable competition,
6. extent of infrastructure deployment,
7. quality of service requirements,
8. holding company infrastructure deployment pattern, and
9. partnering and joint venturing strategies.

These contextual efficiency incentives are external to the AFOR and would impact a RBROR, or price cap, or unregulated telecommunications provider in different ways. These identified contextual incentives impact revenues and costs. They limit or increase the degrees of freedom a telecommunications provider has to operate within a particular AFOR. Any comparative analysis or evaluation should take these and other contextual incentives into consideration. A utility facing little competition would respond to price-cap incentives differently then would a price-cap utility facing strong competition.

Conclusion

Variety has been the hallmark of the regulatory reforms carried out by state utility commissions. Each state has sought to use the reforms in ways that promoted efficiency, encouraged the emergence of viable competition, advanced universal service goals, and furthered the wide spread deployment of new technologies and services. These reforms were carried out against a background that included significant changes in telecommunications market structure, federal regulation, and general economic conditions.

This survey (and any other similar survey) should be regarded as a snapshot taken at one point in time. Over time a number of states have tried more than one telecommunications reform, so any tabulation based on one survey is quite perishable and may not indicate overall trends. In some states legislation is pending or active proposals are before commissions that will change the type of reform.

State commissions (and legislatures) have gone through public and open reform proceedings. Many complex and interrelated issues were examined and negotiated solutions were generally fashioned. Accordingly, a simple comparison of states based upon one feature can only be regarded as illustrative, rather than definitively indicating an advantage or disadvantage enjoyed by one state or approach over another. A valid comparison would entail an indepth examination of the trade offs, goals, market structure, laws, and the state economy. All these important caveats aside, this present survey and analysis allows a state to judge where its reforms are in the context of other states. The survey results also help meet the need of states to learn more about their particular reform and to adapt their reform efforts

based upon the experiences of similarly situated states. Unlike conventional scientific experiments, the "state laboratories" conduct real-time, adaptive, and self-aware experiments that depend on a knowledge of the good and not-so-good results of reforms operating in other states.

CHAPTER 3

EVALUATION RESEARCH: AN OVERVIEW WITH APPLICATION TO PUBLIC UTILITY ISSUES

Evaluation research is a well-established and mature area of research within the social sciences. It has produced successful results based upon its solidly-grounded analytical methods. Just as other social science research methods have made valuable contributions to the regulatory analyses performed at state utility commission, so too can evaluation research. Econometric research, surveys, regressions analyses, expert panels, simulations, and forecasting are all good examples of common research tools used by commissions that have their origins in the social sciences.

Curiously, evaluation research has not been as widely used in regulation as one would expect. This is particularly unfortunate as evaluation research would seem to be able to produce the kind of policy-relevant research commissions need. Evaluation research is designed to answer the pragmatic "impact" and "does it work" questions facing state regulators as they consider or renew various pricing reforms, market entry and restructuring issues, and infrastructure deployment. State regulators care more about the consequences of their actions and whether commission actions successfully further state regulatory goals, than they care about defending a particular approach or doctrine.

Evaluation research is well-suited for the public interest style of commission decisionmaking. Below, an introduction to and a critique of evaluation research is presented. The critique's purpose is to give a regulatory audience a fair appraisal of its ability to meet the research and policy needs of commissions.

Introduction

Evaluation research has become a growth industry, as well as a legitimate social science concern over the course of the past three decades. What Lee J. Cronbach termed "the liveliest frontier of American social science" has spawned a proliferation of journals, books, and monographs devoted to the exploration of evaluation research issues.[1] It is interesting to note that during this period of intense involvement by academicians and evaluation consultants, the level of involvement of public managers was perceived by many to be limited. The evaluation literature during this period contained numerous references to underutilization of evaluation studies by practitioners.

A marked decrease in federal funds for evaluation and a reduction in large-scale social programs over the past few years have caused consultants and the academic community to debate the future role of evaluation research in the policy process. At the same time, retrenchment and the fear of additional cuts have prompted a reassessment of evaluation research on the part of many practitioners. In an effort to demonstrate the worth of their programs many public managers renewed their interest in evaluation research. The nexus between program evaluation and public management, at the state and local levels, represents an intellectually stimulating linkage that yields dividends in both theory and practice.

This chapter will synthesize a number of the salient issues surrounding evaluation research. Within a general overview of evaluation, outcome/impact evaluation will be accorded special attention. Specifically, it will:

- define evaluation research,
- discuss briefly the types of evaluation and their place in the public policy process,
- explore the conceptual and methodological issues surrounding evaluation research,

---

[1] Lee J. Cronbach and Associates, *Towards Reform of Program Evaluation: Aims, Methods and Institutional Arrangements* (San Francisco: Jossey-Bass, 1980), 13.

- assess the strengths and weaknesses of evaluation designs for outcome/impact evaluation,
- highlight utility-oriented evaluation research, and
- assess the future prospects for evaluation research.

What is Evaluation Research?

Definitions of evaluation research abound. Suchman has deemed it to be the "...determination of the results attained by some activity designed to accomplish some valued goal or objective."[2] Scriven defines evaluation research as "...a methodological activity which combines performance data with a goal scale."[3] Rossi and Freeman state that "evaluation research is the systematic application of social science research procedures for assessing the conceptualization, design, implementation, and utility of social intervention programs."[4]

Evaluation is best defined as a systematic effort to measure the performance or impact of a program in terms of the program goals. From a regulatory perspective a program could be a regulatory pricing reform, such as price caps, or flexible pricing, or revenue sharing. A systematic evaluation effort could, for example, use a classical scientific experimental design with control groups. Program goals are the standards against which performance or impact are judged. If a regulatory reform has the goal of increasing the rate of deployment of digital switches in rural areas, then success would be achieved when the rate was higher in those rural areas with reform than it was previously.

---

[2] Edward Suchman, *Evaluation Research* (New York: Russell Sage Foundation, 1967), 10.

[3] Michael Scriven, "The Methodology of Evaluation Research," in Carol H. Weiss (ed.), *Evaluating Action Programs* (Boston: Allyn and Bacon, 1972), 127.

[4] Peter H. Rossi and Howard E. Freeman, *Evaluation: A Systematic Approach*, 5th ed. (Newbury Park: Sage, 1993), 5.

The most significant element of agreement is the emphasis on systematic analysis that undergirds most definitions. Systematic analysis, rooted in the dictates of the scientific method, separates evaluation research from political evaluation efforts. Scientific inquiry makes explicit its assumptions, data, and techniques. Political evaluations make no such disclosures. Scientific inquiry can be challenged on these dimensions while remaining covert with political evaluations. Evaluation research can be replicated. Another researcher, with a different ideological orientation, can take the same set of assumptions, data, and techniques and rework the research. If the results of the two evaluations differ, the roots of the differences can be traced directly to underlying differences in assumptions or methodology, so that assessments can be debated fruitfully. Replicating political evaluation is more problematic as the rules of evidence are not clearly delineated. Many studies submitted to state regulatory commissions are properly characterized as political evaluations.

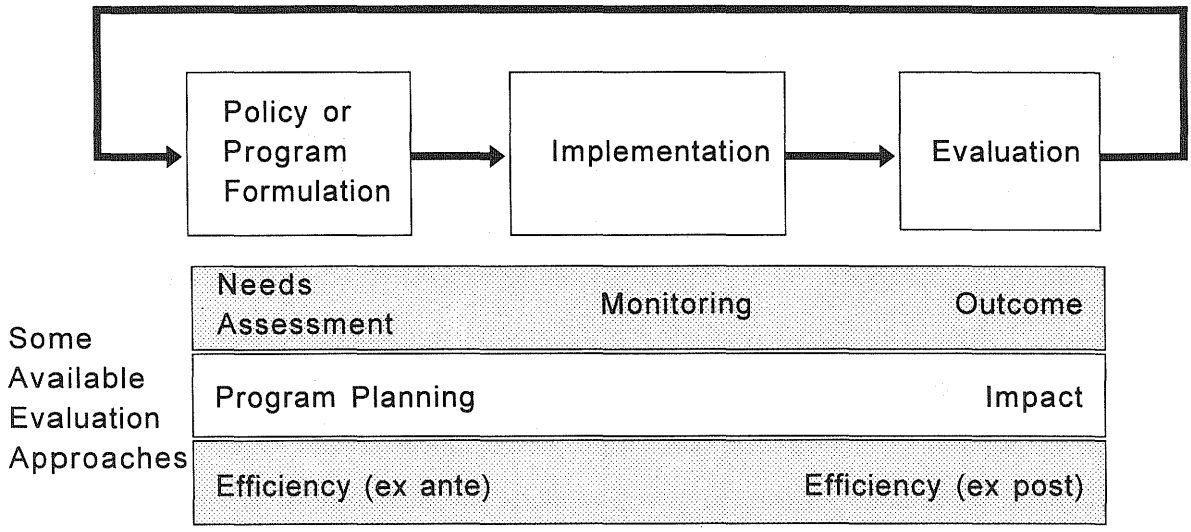
Scientific inquiry represents a way of thinking. It lays out a logic of justification and rules of evidence for decisionmaking that is different than evaluation predicated primarily on political factors or "common-sense." It permits the independent verification of its claims. A definition of science, often attributed to B.F. Skinner, as "the willingness to accept facts even when they are opposed to wishes" is key in recognizing the differences between evaluation research and other modes of evaluation.

In a democratic society evaluation research can supplement but will never supplant political evaluations. Politicians, regulators, and citizens will continue to make the important value judgments about the usefulness of public policy actions. Calls for accountability and the impetus to "reinvent" government have, however, raised the saliency of evaluation research in the decisionmaking process. Such calls indicate dissatisfaction with existing processes of design and implementation of public policy, suggesting both the need to evaluate systematically what happened as a result of a policy intervention and the importance of incorporating those lessons into the reinvention process. These goals require that the tools of scientific evaluation be deployed. Evaluation research will continue to be one input into the process but, increasingly, a more highly valued one.

At federal and state commissions regulation almost looks like it is being changed daily in response to rapidly changing conditions and goals. There is no shortage of proposed and actual regulatory reforms facing regulators and legislators. There is, however, a visible lack of reliable and objective information on the impact or success of the regulatory reforms. It is argued here that the use of program evaluation methods can allow regulators to obtain the valid and reliable information they need when considering, designing, implementing, or assessing a particular regulatory reform.

Types of Evaluation Research

A useful way to approach the discussion of the types of evaluation is to fit them within a public policy framework. If one assumes for analytical purposes that the regulatory and most public policy processes can be depicted as shown in figure 3-1, the types of evaluation can be anchored within each phase. It is important to note that a specific type of evaluation can be employed in more than one phase.



Source: Author's construct.

Fig. 3-1. Public Policy Process and Evaluation.

Policy or Program Formulation Phase

In the process of formulating public policy initiatives, problems and needs are defined and recognized, alternatives are assessed, and program design decisions are made. For a state regulator this stage may be when infrastructure modernization problems are recognized, alternatives are considered, and an order or plan is created that provides explicit incentives and timetables for accelerated infrastructure deployment. The evaluator has a wide variety of techniques that are relevant in the formulation process, both to provide evaluative information at this stage, as well as to provide information useful in the implementation and evaluation stages. The following section highlights a selection of those techniques.

Needs Assessment

Any sort of policy design will require that key players or stakeholders in the policy process be involved in the formulation of policy. Evaluation offers several techniques by which this information gathering may be systematized. Rossi and Freeman provide an excellent overview of the techniques available to perform a needs assessment.[5]

Expert Opinion or Key Informant Approaches

Experts typically appear in a regulatory setting as individual consultants that are paid to advance the position of one party and to discredit the positions other parties. Several expert opinion approaches seek to use a neutral consensus building process to build upon areas of common agreement. This emphasizes the expert portion of the participants more than the partisan portion. The ability to elicit less partisan responses is driven by the confidence that participating experts have in the facilitator, and the openness and the nonjudgmental nature of the process. Even on very contentious issues, partisans and experts alike are still more interested in accurately expressing their position on the nature of the problem than in critiquing others.

---

[5] Rossi and Freeman, *Evaluation: A Systematic Approach*, 5th ed. (Newbury Park: Sage, 1993).

Each of the processes briefly described below have experts or key informants provide information on issues or processes or outcomes. Key informants are not formal experts but have an expertise because they participated in a process and have a direct knowledge of events, timing, goals, and issues. Their expertise is that "they were there." Information provided by experts and/or key informants in these approaches has the most validity when a representative range of experts or key informants are used.

Key informants are selected to provide insight in the Nominal Group Technique (NGT) and for the Delphi Technique. The Delphi Technique is defined by Van de Hen and Gustafson as a "method for the systematic solicitation and collation of judgments [of experts] on a particular topic through a set of carefully designed sequential questionnaires interspersed with summarized information and feedback or opinions derived from earlier responses."[6] Participants do not have to meet face to face in the Delphi Technique. NGT as defined by Delbecq, Van de Hen and Gustafson requires a group meeting in which the following steps occur:[7]

- Silent generation and recording of issues and problems.
- Round-robin feedback from group members to record each idea in a terse phrase on a flip chart.
- Discussion of each recorded idea for clarification and evaluation.
- Individual voting on priority ideas with the group decision being mathematically derived through rank-ordering or rating.

The value of the data garnered through either of these two processes is dependent on the degree of knowledge of the experts selected and their representativeness of the population one wishes to assess. The adage that "where one stands, depends on where one sits" points up

---

[6] Andre L. Delbecq, Andres H. Van de Hen, and David H. Gustafson, *Group Techniques for Program Planning: A Guide to Nominal Group and Delphi Processes* (Glenview, Illinois: Scott, Foresman and Company, 1975), 10.

[7] Ibid., 6.

the potential biases attendant to a key informant determination of need. The California Public Service Commission, for example, once convened a large expert panel to estimate the growth of competition in various telecommunications markets. It did this, in part, to minimize the "dueling consultants" phenomenon and to take advantage of the collective expertise of a large number of telecommunications experts. Having each expert answer identical questions using specified assumptions can greatly narrow the range of predicted outcomes.

Community Forum

Analogous to a town meeting, the community forum approach elicits broader input into the definition of need. The representativeness of the participants and their willingness to share their views determine the utility of this approach. Community forums often are used to follow-up the key informant approach. Public hearings held by commissions may be viewed as a distant cousin of this technique. The intent in this approach is to have concerned parties state their needs. Because participants are self-selected, these forums and hearings can not be viewed as representing the whole population. Equally, there is no presumption that all needs can and will be responded to programmatically. The advantage of the approach is that it can provide an unfiltered image of the needs of the most concerned customers.

Indicators Approach

Census data and a wide variety of indicators are collected by the federal, state, and local governments. Crime, fertility, labor force participants, income data, economic activity, and disease are a few of the time series of indicators archived by governments. The number of households with telephones is one piece of census data familiar to state regulators. The existence of valid, neutral, and reliable indicators can aid in needs assessments by facilitating the delineation of a target population, and estimating the incidence and prevalence of a condition. The obvious drawbacks of indicators can be their lack of availability and reliability. The timeliness of census data is a problem for evaluation studies begun later in the decade. Regulators should find this type of information especially useful when facing programs dealing with specific populations, such as the elderly, the poor, the disabled, and those dwelling in rural areas.

Surveys

Querying individuals directly is a straightforward, often employed approach to needs assessment. The selection of the sample, the design of the questionnaire, the flow of the questions, the wording of the questions, and the analysis of the results are caveats that need to be considered in the conduct of a survey.

Asking individuals to state their needs or willingness to use a particular telecommunications service without also identifying the price of the service is a common failing in needs assessment surveys presented to commissions. Equally prevalent are surveys of nonrepresentative populations—such as telecommunications managers in large companies or high-end telecommunications users—which are presented as representative of the telecommunications needs of more general populations.

These techniques can assist regulators and policymakers in the definition of needs. They provide processes by which problems and concerns can be considered systematically in a nonadversarial setting. Properly and neutrally designed, these techniques can produce relatively untarnished information for regulators to assimilate and evaluate while formulating major policy initiatives, such as a regulatory pricing reform. It is important that this critical initial step in the public policy process be informed by the systematic tools of the evaluation analysis approach.

Program Planning

After a problem is defined and a decision is made to fix the problem, the selection of the particular intervention strategy, or solution, or option remains. Evaluators call the selection a program and define this as a set of related actions taken to achieve one or more specified goals. For a commission the program may be defined in an order (or series of orders) in legislation, or in a plan, or cooperative agreement. Traditionally, because state commissions acted in a responsive, quasi-judicial mode, they rarely used the term "program" to describe their actions. Increasingly, however, state commissions have specific focused actions, or programs, to achieve regulatory goals. As such, the lifeline, pricing reforms,

infrastructure modernization, and market restructuring activities of state commissions can be accurately characterized as programs intended to achieve regulatory goals.

The tools of program planning analysis can be useful in sorting out the various competing alternatives. In particular, the two tools briefly presented below can provide systematic information that can be evaluated by regulators. The techniques are primarily designed to separate rhetoric from reality. The first technique is more rigorous mathematically and the second is somewhat easier to accomplish. Policymakers can better evaluate the purpose of each program feature or component through using these techniques.

Multi-Attribute Utility Method (MAUT)

Edwards, Guttentag, and Snapper propose using the Multi-Attribute Utility Method (MAUT) as a tool in program planning.[8] Derived from decision theory, MAUT is a process that can be used in choosing among programmatic alternatives. The authors identify the following steps:

Step 1:	Identify the organization whose utilities, benefits or values are to be maximized. For whom is the evaluation being conducted?
Step 2:	Identify the issue or issues to which the preferences or utilities needed are relevant. What is the purpose of the evaluation?
Step 3:	Identify the entities to be evaluated. What are the choices or options?
Step 4:	Identify the relevant dimensions of value. Identify the criteria to be employed.
Step 5:	Rank the dimensions in the order of importance.
Step 6:	Rate dimensions in importance, preserving ratios.

[8] Ward Edwards, Marcia Guttentag and Kurt Snapper, "A Decision Theoretic Approach to Evaluation Research," in E. Struening and M. Guttentag (eds.) *Handbook of Evaluation Research* (Beverly Hills, CA: Sage, 1975), 140, and Luc Anselin and J. Stephen Henderson *Decision Support System for Utility Performance Evaluation* (Columbus, Ohio: The National Regulatory Research Institute, 1984).

Step 7: Sum the importance weights, divide each by the sum, and multiple by 100.

Step 8: Measure the location of the entity being evaluated on each dimension. Measurements are subjective judgments estimating the probability on a 0-to-100 scale that a given option will maximize each dimension.

Step 9: Calculate the utilities or values for each entity.

$$U_i = X_i W_i U_{ij}$$

where U_i = the aggregate utility for the ith entity;
W_j = the normalized importance weight of the jth dimension;
U_{ij} = the rescaled position of the ith entity on the jth dimension.

Step 10: Decide if a single act is to be chosen, the rule is to maximize U_i.

This approach works well when the decisionmaker is able to assign the required numbers with confidence. The policy analyst does not need to identify an exhaustive number of relevant dimensions or worry about the exact precision of the initial numbers assigned. It is the use of ratios that drives the analysis, so being "off" in the numerical value assigned to an important weight is not a problem unless the value lacks any validity. The approach is especially useful for choosing between programmatic alternatives because of a maximization decision rule.

Forward Mapping and Backward Mapping

Weimer and Vining offer practical policy advice for considering the implementation system in the program design phase.[9] Forward mapping is defined as "the specification of the chain of behaviors that link a policy to desired outcomes." Scenario writing, a form of forward mapping, is employed to make explicit the assumptions underlying the implementation models, as well as a flagging potential problem areas. This technique may be an especially valuable tool to convert the rhetoric and claims of advocates and consultants into

[9] David L. Weimer and Aidan R. Vining, *Policy Analysis: Concepts and Practice* (Englewood Cliffs, NJ: Prentice Hall, 1989), 311.

systematic documents suitable for analysis and comparison. Weimer and Vining identify the following elements in forward mapping:[10]

- A scenario is written that specifies the who, what, when, where and why.

- The scenario is critiqued. The plausibility of the actions required given personal and organizational interests is assessed. Avoidance tactics are identified. Compliance-inducing tactics are developed. The action and inaction of other actors directly or indirectly involved are assessed and addressed.

- The scenario is rewritten to improve its plausibility. Backward mapping is defined by Weimer and Vining as an approach that begins with the desired outcomes, ascertains the most direct way of achieving them and then maps action backward through the organization hierarchy to the highest-level policy that must be adopted to realize the desired outcomes.[11] Backward mapping is most valuable in developing policy alternatives that have a high probability of success, and forward mapping is most useful for anticipating problems in programs that are currently being implemented.

The advantage of this evaluative technique is that often programs and program components are assumed to have certain outcomes and to achieve particular goals. Although mapping necessarily lacks causal data, it is very useful because it forces both partisans and regulatory policymakers to justify program components on more than rhetoric. If a program component is to promote economic development, then the program designer should be able to map or show the sequence of events and actions that will take place from program initiation to economic development. Likewise, in backward mapping, it may be that a different

---

[10] Ibid., 313.

[11] Ibid., 311.

program would be designed if a chain or events is mapped starting with the desired result: economic development.

Both techniques allow regulators to have a reality check on the overall program or individual program components independent of the claims of others. Being able to evaluate and understand the planned output and impact is important at the program design stage, as well as at the monitoring and evaluation stages.

Economic Efficiency (ex ante)

In some instances, for evaluators, economic efficiency analysis in the program planning stage entails estimating the costs of policy options and comparing those costs to the benefits to be achieved by the proposed policy initiative. Two variants of efficiency analysis are commonly used. The most popular variant is cost-benefit analysis, in which the costs of the proposed policy are totaled and compared to the monetized, discounted benefits the policy is expected to yield. As shown below, cost-benefit analysis requires many difficult, controversial assessments, particularly regarding the assignment of the monetary equivalent for policy benefits that may include saving or enhancing human lives in ways that may be difficult to reduce to pecuniary values. When alternative policies are expected to have roughly similar benefits, the difficult issue of valuing benefits can be avoided by employing a cost-effectiveness analysis. For example, a decision to devote resources to the commercialization of a new drug might result in a cost-benefit analysis in which the benefits of the new drug were totaled and compared with the expected development costs to see whether development should proceed. In the process, important questions are raised about how those benefits should be measured. If, however, the new drugs were designed to compete with or substitute for an existing therapy, and if the two drugs were expected to have similar benefits, cost-effectiveness analysis would require that the cost of therapy with the new product be compared to that currently available. If the word "infrastructure" is substituted for the word "drug," in the preceding sentences, an immediate application to regulation can be seen.

Cost-benefit analysis is useful in forcing systematic consideration of an alternative's potential costs and benefits and in emphasizing that policy choices have consequences in the

form of foregone options. However, the formalization comes at a price, it requires that the analyst reduce to dollars benefits and costs that may be highly valued but difficult to convert to a common currency. The net value of rural economic development, for example, may be expressed in economic terms, but most observers would agree that this would paint an incomplete picture of the benefits to be obtained. More specifically, cost-benefit analysts must concern themselves with issues such as the following:

1. Which Costs and Benefits To Include

   The identification of costs including direct, indirect, and opportunity costs as well as the benefits to be considered can be problematic. Reliable *ex ante* cost estimates often are not available. The determination of costs and benefits is critical in this initial step in the cost-benefit analysis.

2. Monetizing Outputs and Outcomes

   The absences of a market for some outputs and outcomes prompts the consideration of shadow prices or derived prices. The valuation of human life presents a particularly thorny issue. How does one go about determining the value of a statistical life? How does one calculate the value of an improved delivery of safety services through a modernized infrastructure?

   A number of methods were developed, none of which has been found to be wholly satisfactory. For example, discounted future earnings can be used. The present value of lost future income resulting from a premature death can be estimated. Usually only labor income is examined. Those who are not employed, children, the elderly, and low-wage workers have a lower value placed on their lives. Willingness-to-pay measures are another alternative. Several variants of willingness-to-pay measures exist. Consumer polls and surveys have been used but preference revelation problems and the wording of questions have resulted in a wide range of estimates. Regulators know the inherent difficulty in studies that forecast costs, prices, and demand for new, or fledgling, or yet-to-be developed

telecommunications services. Analyses based on such derived data must be discounted and evaluated accordingly.

Further, the choices people make in the marketplace, reflected, for example, in the compensating wage differentials paid for risky occupations is another variant. Generalizing to a larger population from individuals willing to enter high risk occupations may understate the benefits from reducing exposure to risk. Policy planners may possess information about risks superior to that possessed by wage earners responding to risk differentials. In some instances, risk differentials may be compressed simply because employees possess few market alternatives to the risky occupation.

3. Choice of Discount Rate

In many government programs costs and benefits are realized over time. A fixed amount payable in the future is worth less than the same amount in the present. The choice of an appropriate discount rate is a significant decision in a cost-benefit study. The higher the rate at which benefits are discounted, the more difficult it is to obtain net benefits. Conversely, the lower the rate, the easier it is to justify a public program.

The analyst must decide whether society's discount rate should differ from those of private individuals and which of a wide array of potential discount rates should be employed. If real rates of discount are employed, the analyst must incorporate assumptions about rates of inflation, deflating future benefits accordingly. Nominal rates have exhibited very considerable volatility, with current rates (both real and nominal) far below those of the 1980s. Not surprisingly, no consensus exists on the current rate to use. The debate centers on whether to use the 10 percent urged by the Office of Management and Budget or a smaller value more in line with current market conditions. High values, which are achieved by reducing the value of deferred benefits in comparison to current costs, tend to make the policy planning process appear relatively myopic, and lead to focus on projects with large short-term impacts. Often researchers offer

computations for a range of discount rates, thus permitting the policymaker to incorporate whatever assumption seems appropriate. Under this approach, if it will take a jurisdictional utility ten or twenty years to ubiquitously deploy its advanced infrastructure then the benefits would necessarily need to be large to avoid being "discounted out of existence."

4. Choice Criteria

Several choice criteria are available. The Pareto criterion states that one social situation is better than another if at least one person is better off and no one is worse off in the preferred situation. It is surely reasonable to require that policies yield Pareto efficient outcomes in the sense that all opportunities to improve the lot of individuals without making someone else worse off are exploited. However, government policies typically involve redistribution, and accordingly the Pareto criterion will be violated by many proposed and desirable policy choices. Unfortunately, economists do not have reliable and well-accepted tools that permit comparisons of options resulting in *different* income distributions. One attempt to address this problem is the Kaldor-Hicks criterion, which stipulates that one social state is better than another if the preferred outcome makes it possible for those who benefit to compensate those who lose under the policy. The Kaldor-Hicks criterions is a "thought experiment"—compensation need not be paid, but the efficiency benefits of the policy choice must be large enough that the winners would be willing to compensate the losers and would still find themselves better off as a result of the policy's implementation. In regulation this could mean that if large sophisticated telecommunications users are disproportionately benefited by the universal deployment of an advanced infrastructure, they would still be ahead even if their rates were higher than the rates charged to residential and small business customers.

If distributional effects are simply ignored, more direct measures of a policy's attractiveness are available. The internal rate of return can be computed by finding the discount rate which makes the net present value (discounted benefits

net of discounted costs) of a project zero. More generally, distributional issues can be recognized explicitly as part of the choice decision among policies. That is, benefit-cost analyses can be disaggregated across persons, places, organizations, or other category judged to have distributional relevance, and the resulting analyses can then be weighed according to the distributional goals of the planning process.

The common theme across each of these techniques is in how to provide objective information in a way that allows better identification of problems and evaluation of the advantages and disadvantages of the available programmatic choices. None of these techniques is powerful enough by itself to identify and prioritize problems, or to unfailingly select the correct program. Rather, when used in an evaluative context, the information produced can be trusted because all parties can understand the strengths and weaknesses of the underlying data-generation methods. This kind of information serves as one input to a decisionmaking process. Political, economic, institutional, technological, and legal constraints are among the other inputs that a policymaker or regulator must filter through their value systems.

The evaluative context means that advocacy of a particular option or solution does not control the results obtained because the intent is always to let the "data speak for itself." Reports and testimony submitted by interested parties before a commission are not necessarily evaluation research even if one or more of the above techniques are used. It is the fair, systematic, and objective treatment of each option that is one of the key characteristics distinguishing evaluation research from the advocacy reports submitted by interested parties.

A second useful characteristic of evaluation analysis is the comparison of findings or results with regulatory goals. Having, for example, a low number of facilities-based firms available to challenge the incumbent LEC is not a problem, unless state regulatory goals call for increased local exchange competition. Absent goals, data can be used to prove many different points. In this example, the number of challenges is just a number, it "does not speak for itself" unless it can be done in relation to a goal. Evaluative research is particularly well-suited for analyses of conditions or performance in relation to goals.

Implementation Phase

Once a regulatory policy is formulated, putting the policy into action requires that considerable attention be focused on implementation decisions. While the assessment of how well a regulatory policy was implemented is the task of policy evaluation several of the techniques developed for policy evaluation bear directly on the process of implementation. These techniques fall under the heading of administrative monitoring and performance monitoring. The former refers to the compilation of descriptive measures of program activities and costs. Traditionally, much of the monitoring that has taken place in regulatory and other government programs has taken this form. Performance monitoring, on the other hand, emphasized evaluative measures focusing on outcome measures.

Program Evaluation and Review Technique (PERT)/Critical Path Method (CPM) Used for Performance Monitoring

While relevant in planning, PERT and the CPM are also useful in monitoring programs. Both are forms of network analysis that link together events that are designated by circles and activities, designated by arrows. The activities are usually measured in terms of the number of days need to accomplish them. The critical path indicates the longest time it will take to complete a project or program. Construction of a PERT/CPM with its serial logic and clearly demarcated events and activities is good discipline in the planning phase, but its principal use occurs during implementation where conformity with the time lines or budget estimates can be assessed. It would be useful, for instance, to know from using PERT information if the sequence of events connected with infrastructure deployment in a region was proceeding as planned.

Use of Aggregate Data for Administrative Monitoring

Primary sources for collecting implementation data are program records, observations, and self-report methods. Records, if kept for purposes other than the evaluation, can be an objective, accurate, and relatively inexpensive way of determining what is happening in a program. They can, however, be incomplete, inaccurate, and perceived as burdensome to collect if they are collected only for evaluation purposes and present confidentiality and ethical issues of access to certain kinds of records. An observation conducted by an outsider can provide a view of the program perceived as being impartial. Observation can be costly and inhibit the behavior of those being observed. Self-reporting measures, either through face-to-face interview or questionnaire, are used frequently. The interview allows researchers to collect data from those unable to complete a questionnaire and allows the interviewer the flexibility to probe questions in depth. Interviews are very costly and the specter of interviewer bias is possible. Questionnaires allow one to collect information from a large number of people over a wide array of issues. Obtaining an acceptable response rate is the single most difficult issue with questionnaires, once interviewer and question bias issues are resolved. Using this approach, a survey could be administered to vendors, telecommunications providers, and large users to see if interconnection policies were correctly implemented and to identify any previously unforeseen problems. This approach would be much less expensive than a typical hearing or a significant complaint investigation.

Evaluation Phase

The evaluation stage focuses on asking, what happened as a result of the program, service, intervention, or treatment? What would have happened in its absence? These are the key questions considered in outcome or impact evaluation. Indeed, outcome and impact evaluation are what is customarily thought of as policy evaluation. Though as we have seen above the techniques of policy evaluation are useful and utilized throughout the policy process. A host of issues must be considered as one attempts to provide answers to these fundamental evaluative public policy questions. For purposes of exposition, a discussion of these issues is organized below according to two categories: conceptual and methodological.

The classification is neither mutually exclusive nor exhaustive. Conceptual issues have methodological implications and methodological issues are rooted in the larger context of evaluation.

Conceptual Issues

Prior to undertaking an evaluation, the evaluator must confront numerous conceptual issues, such as deciding who should conduct the evaluation, issues dealing with which goals should be considered, establishing priority among goals, long-term versus short-term goals, and establishing a basis for comparison, are among the issues that must be addressed regarding the regulatory program selected for evaluation.

Which Program To Evaluate?

While it is true that all programs can be evaluated, *not all programs should be evaluated.* Joseph Wholey argues that scarce evaluation resources should be used for those programs ready to be evaluated.[12] He proposes conducting an evaluability assessment to ascertain readiness for evaluation. Good candidates have the following characteristics:

- clearly stated goals;
- plausible goals given resource and staffing levels;
- consensual objectives;
- reliable, valid data available for performance measures;
- steady-state implementation;
- decision regarding continuation, modification, or termination pending; and
- a management willing to use the evaluation.

---

[12] Joseph S. Wholey, "Evaluability Assessment," in Leonard Rutman (Ed.) *Evaluation Research Methods: A Basic Guide.* (Beverly Hills: Sage, 1983).

These steps serve as a diagnostic tool for commissioners and commission program managers. They indicate the areas that need attention before an outcome or impact evaluation is undertaken. Performing an evaluation if goals are unclear, or data are sketchy, or the program is experiencing budgetary or staff upheaval virtually guarantees that the evaluation will yield little useful information. Experience has indicated that it is better to cancel a planned evaluation unless most of the diagnostic points identified by Wholey are present.

Who Should Conduct the Evaluation?

There are two basic choices for conducting an evaluation. An outside evaluator can be contracted or the commission can perform its own evaluation. Outside evaluations tend to be costly but often appear to be more credible to other parties than evaluations conducted internally. Interestingly, recent research on the utilization of evaluation suggests that the results of internal evaluations are more likely to have an impact on the behavior of the organization under review compared to analyses done by outsiders. This suggests that commissions and other organizations feel more ownership of internally-generated evaluations, have a corresponding increased confidence in the results, and are more willing to implement recommendation. On the other hand, even though other parties may trust an external consultant more, commission (as with most organizations) have less comfort with externally-generated findings and may be less likely to implement recommendations.

While utilization may differ according to source, no empirical research has definitively established the superiority of either approach in terms of the quality of the evaluation product. After the decision about which program should be evaluated and who should conduct the evaluation, a number of additional issues must be addressed.

Whose Goals?

As goals are central to most definitions of evaluation research, a decision must be made as to whose goals will be considered. For most public initiatives, there are a number of stakeholders whose goals could be given consideration. Legislators, consumers, shareholders,

regulatory agencies, funding agencies, interest groups, service providers, and watchdog groups could all have goals with respect to a particular policy initiative. These goals, of course, can be contradictory with some maximizing efficiency and others championing equity. Whose goals should be considered? Careful judgments need to be made at this stage. Failure to consider the goals of relevant stakeholders can result in evaluations perceived to be illegitimate by some in the policy arena.

Goal selection may be a particular problem for a commission that does not have explicitly stated goals. Inspection of significant state commission orders may reveal both explicit and implicit goals. Clarification of the latter is likely to generate the most attention. A number of state commissions, for example, have participated in multi-party state-wide task forces that produced consensual goals regarding a modernized telecommunications infrastructure. These and other efforts can produce goal statements that can be used to guide a commission's evaluation research efforts.

Priority of Goals

Even when the relevant stakeholders and their goals are identified, the issue of goal priority remains. Are all the goals to be considered equal? Will they be rank ordered? What criteria will be employed in such a ranking system? While some of the techniques discussed in program planning and needs assessment (MAUT, Delphi and NGT) are relevant, there are no quick fixes for gaining consensus on mutually contradictory goals.

Temporal Dimension of Goals

State actions can have short, intermediate, and long-term outcomes and impacts. Which should enter into the evaluation? If the evaluation taps only the short-term outcomes or impacts of the initiative, the staying power of the effect is unknown as is the presence of any longer-term effects. Adding a longitudinal or time component to the evaluation design enables the researcher to determine the longevity of any behavioral change detected in the short-term evaluation.

This advantage is accompanied by a higher price tag. Tracking participants over a longer period of time, maintaining a database, and diverting staff resources to these tasks all contribute to the increase in both incurred and opportunity costs. Practically speaking, most programs are evaluated in the early stages of development, and the existing evaluation research results produced have a short-term orientation, leaving unanswered the extended consequences of public interventions. While a commission may ideally care about the effectiveness of a reform in increasing over a multi-year period the use of touchtone service, it may be difficult to consistently maintain the resources necessary to monitor and evaluate any increase.

Base for Comparison of Goal Achievement

If evaluation is an inherently comparative exercise, to what does one compare the results of a particular public program or regulatory initiative? How does one judge if a level of goal attainment is good, bad, or indifferent? For some programs, absolute standards were set either by legislative, judicial, or executive edict, an accrediting body, or professional associations. They serve as the yardstick by which comparisons are made.

The organization's performance prior to the intervention can also serve as the base for comparison. If data exist on comparable performance indicators within the organization for the period prior to the implementation of the program in question, the base for comparison is the organization itself in an earlier period. Absence of comparable data often undermines an organization's ability to use its past performance as the base for comparison. Within the organization, a program can be compared with other programs to assess its relative performance. Also, the performance of comparable organizations on the same policy intervention is another possible base for comparison. A program's actual performance can also be judged against a target set prior to implementation. The use of agreed-upon targets may well be the most appropriate approach for state regulatory commissions. Comparability though is difficult to achieve. A different mix of consumers and other inputs can make it difficult to compare the performance of one utility with another. Practical difficulties of obtaining comparable data from other organizations often inhibit this approach.

The existence of a group comparable to the one receiving the program, service, or intervention who do not receive the intervention is another powerful comparative technique. There are a number of ways to construct such a group. One could randomly assign units (organizations, individuals, regions, or groups) to the treatment group that receives the intervention or to the control group, but this is not always feasible. Naturally occurring groups who do not partake of the service or intervention also have been used as a base of comparison for those who do participate. While the research design section will detail the strengths and weaknesses of these approaches, it is important to understand that the choice of a relevant comparison group is not simply a technical issue. It represents an important conceptual issue that needs to be considered before embarking on the evaluation enterprise.

Methodological Issues

Causality

At their core, outcome and impact evaluation are concerned with making causal inferences. Programs are said to be the cause of the effects observed. Specific program incentives, for example, cause a faster rate of deployment. The criteria for making causal statements are well-established. There must be a linkage between cause and effect. Plausible alternative explanations must be ruled out. This final criterion, ruling out other explanations, poses a significant challenge. Causal statements are frequently made. A government program (X) is enacted. Outcomes (Y) are observed and attributed to the intervention (X → Y). Spurious and confounding variables (Z) could, however, account for the observed relationship between X and Y. For instance, increased rate flexibility (X) could be said to cause or be responsible for increased sales of telecommunications services (Y). A careful examination might, however, reveal that general economic growth (Z) is actually more responsible.[13] If a

[13] Suppose, for example, that Z influences both X and Y. The spurious influence of Z could, in such a case, entirely explain the observed covariation between X and Y. But if X and Y move together simply because of the joint influence of Z, a policy change in X, holding Z constant, may, and likely will, fail to exert the desired influence on Y. That is,

confounding variable explains some, but not all of the relationship, the result will be either an over or underestimation of the effect of X on Y. Without an adequate research design, invalid causal inferences can result.

Establishing the Counterfactual

Outcome and impact evaluation are designed to answer "what happened" as a result of the program or regulatory initiative. But before this question can be answered one must address the counterfactual question of what the situation would have been without the intervention. The comparison strategies identified earlier are critical in providing insight into this question. The choice of research design will determine how well answers can be provided to these two questions. A comparison of telephone companies with similar characteristics, except for pricing flexibility, would be one way to think about and establish a counterfactual.

Translating Goals into Measurable Objectives

Another methodological challenge is to convert goal statements that are often ambiguous and amorphous into objectives that can be measured. Universal service, affordable rates, and cost-causation are common regulatory goals that are typically difficult to measure. Three key features of translating goals into measurable objectives are reliability, validity, and availability. Measures must be reliable, meaning they must be stable. Measures must be valid, that is to say, they need to measure what one says they measure. Reliability is a necessary condition for validity. A valid measure is always reliable but a measure may be reliable without being valid. For instance, household penetration of telephones is a stable and reliable measure. However, it is not by itself a valid measure of the ability of disabled Americans to use the public switched telecommunications network.

---

such spurious factors can lead one erroneously to conclude program impact when the observed relationship was due totally to an extraneous factor.

Well-established techniques are available to assess the reliability and validity of measures. Multiple measures are often used to reduce the reliance on any particular measure. This strategy highlights the third consideration—availability and accessibility of data. Even if valid and reliable measures exist, they may be unavailable for the program under consideration. Faced with the task of constructing a new database and postponing the evaluation, researchers may substitute measures whose reliability and validity are suspect but which can be computed with data in hand.

Specifying the Independent Variable

Identifying the independent variable, the program X (pricing flexibility) whose Y impacts (increased sales of telecommunications services) are to be evaluated appears straightforward. In many cases, however, the program (X) is not a single entity, or price, but is instead composed of a number of components. That is, the pricing flexibility of program (X) may consist of competitive (X_1), partially competitive (X_2), and noncompetitive (X_3) prices, or an energy conservation program (X) may be comprised of home audits (X_1), weatherproofing assistance (X_2), and informational brochures (X_3). By aggregating them together in one X, it cannot be determined if they all contributed equally to generating the observed outcomes or whether observed outcomes were the product of a single component. From a program perspective it will not be possible from such an analysis to determine which part of the program should be emphasized, to get even better results, or deleted. From a resource allocation perspective with an improperly specified causal variable, it will not be possible to know if resources allocated to all component parts are necessary to achieve the desired results.

If attention is paid to the aggregation problem, these interpretation problems can be overcome. The key lesson here is that specifying the independent variable or cause requires careful thought as most programs are embedded in a complex set of variables. It would be equally as inappropriate to conclude that "everything is related to everything else" as it would be to prematurely settle on one particular variable. Identification of the causal variable or set

of variables requires asking "Why are we doing [or proposing] this?" Clarifying the "Why" will directly lead to a confident specification of an independent causal variable.

Statistical and Substantive Significance

In scientific experiments and in some evaluations a statistical significance test can be used as an objective criterion as to whether or not observed differences between the group receiving the service and a control group are statistically significant. Schneider and Darcy (1984: 574) question the choice of significance levels chosen in evaluation research. Typically, it is assumed that statistical significance is equated with whether or not the program had an impact. As Schneider and Darcy correctly point out, program impact is only one of the factors that affect the outcome of significance tests. Others include the number of cases, variation among cases, degrees of freedom, the appropriateness of the statistical measure used, and the hypothesis being tested. It is erroneous to conclude substantive significance from statistical significance. Moreover, the choice of a specific significant test level can create bias in favor of a null hypothesis.[14]

Why create such an enormous bias in favor of the null hypothesis? Suggestions of 10 percent and even 20 percent testing levels have been made. Schneider and Darcy propose that researchers report not only the probability that an impact of the magnitude observed was produced by chance (level of significance) but also the probability that an effect large enough to be relevant would have been detected if it had existed (the power of the test.) Decisions regarding the precision of and the policy relevance of the estimates required need to be made in the early stages of evaluation planning.

---

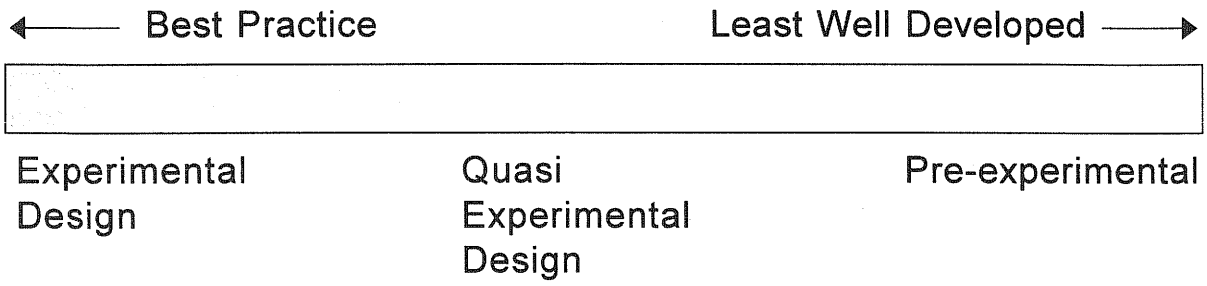
[14] To a research scientist the hypothesis would be that "pricing flexibility increases the sale of telecommunications services." The statistical significance test uses a null hypothesis that pricing flexibility does not result in increased sales. The .05 ensures that the null will incorrectly be rejected only 5 percent of the time.

Generalizability

To what persons, places, times, and settings can one generalize the results obtained in an evaluation? If the data suggest that a program is successful in Columbus, Ohio will it be successful in Oakland, California? If a particular pricing strategy achieves expected results with small service-oriented businesses, is it transferable to small manufacturing concerns? If a program worked in 1989, can it be expected to work in 1995? This issue of generalizability is often an important concern to policymakers. Balancing generalizability and causality must be reckoned with in the research design phase. It is in this phase that each of the issues identified above must be acknowledged and accommodated in a methodologically sound way.

Evaluation Research Designs

The evaluation research design provides a blueprint for the collection, analysis, and interpretation of data. The conceptual and methodological issues discussed earlier must be considered and resolved in the design phase. An expansive menu of options is available ranging from methodologically elegant experimental designs to one-shot case studies. If one envisions a continuum anchored on one end by the best practice method for inferring causality and anchored on the other end by the least developed, then experiments and preexperimental designs would occupy the polar positions, as illustrated in figure 3-2.



Source: Author's construct.

Fig. 3-2. Evaluation techniques ranked by ability to infer causality.

Evaluation Experiments

The experiment has been called the Cadillac of evaluation design. It represents the most powerful design available for inferring causality. The defining characteristic of an experiment is randomization. It is important to be clear about the definition of randomization. It does not connote a sloppy or haphazard process. Rather randomization in an experiment means that all units have an equal, nonzero probability of being selected for treatment. Randomly assigning individuals, companies, or geographical areas to a treatment or control groups does not guarantee that the two groups are identical. It does ensure that, in the long run, the two groups differ only through the operation of chance factors that behave according to well-established laws of probability. The purpose of the treatment group is to help answer the question—what happened as a result of this program? Did accelerated modernization result in an increase in demand? The control group is designed to answer the counterfactual question—what would have happened in the absence of the modernization program? For the area without the accelerated modernization, was there an increase or decrease in the demand for telecommunications services?

In order to conduct an experiment, the experimenter requires a great degree of control. Assignments of units, whether they be individuals, households, or organizations must rest not with the professional judgment of program managers, but with the randomization procedure of the experimenter. The experimenter needs to be involved in the program planning phase to build in randomization procedures before the intervention or new program is underway. An experiment, with its randomization requirement, may not be possible if a program is already underway when the evaluator is included.

There is a great deal of resistance to randomization and experiments from program managers and potential subjects. This resistance is not surprising given that the program is designed to have a beneficial impact on its subjects and placement in the control group is considered to be undesirable. Past reports of unethical behavior on the part of experimenters in the past contribute to skepticism.

In addition to the fears of abuse, a number of other concerns are often expressed about experiments. Program professionals can be reluctant to surrender their professional judgment

about program assignment to a randomization procedure. They might feel that it is unfair to deprive an individual of an opportunity to participate in a program deemed to be beneficial, in the name of randomization requirements. Hatry identified the following conditions in which randomization might be appropriate:[15]

1. When there is likely to be a high degree of ambiguity as to whether the outcomes were caused by the program if some other designs were used.
2. Some citizens can be given different services than others without danger of harm.
3. Some citizens can be given different services than others without violating ethical and moral standards.
4. There is substantial doubt about the effectiveness of the program.
5. There are insufficient resources to provide services to all potential clients; demand exceeds supply.
6. A decision on the program can be postponed until the experiment is completed.
7. Experimental conditions can be maintained reasonably well throughout the period.
8. Sufficient staff and resources are available to manage the experiment.
9. Client consent is not required or if it is can be obtained without invalidating the experiment.
10. Confidentiality and privacy of clients involved can be maintained.

Experimental Designs

The following is not an exhaustive enumeration of experimental designs. Rather it is a sample of those designs found in public sector evaluation research.

---

[15] Harry Hatry, Richard E. Winnie, and Donald M. Fisk, *Practical Program Evaluation for State and Local Governments* (Washington, D.C.: The Urban Institute, 1981), 42.

Classic Pretest-Posttest Design

The classic pretest-posttest design is denoted in figure 3-3.

Experimental Group	R	O	X	O
Control Group	R	O		O
	T_1	T_2	T_3	T_4

The notation is consistent with that of Campbell and Stanley (1963):[16]
R indicates random assignment.
O indicates the outcome measure: such as the demand for telecommunications service.
X indicates the intervention: accelerated network modernization.
T indicates time period.

Fig. 3-3. Classic pretest-posttest design.

What is of interest about the design is the net change between the experimental and the control group. The experimental group helps answer the question what happened as a result of the intervention. The control group answers the question of what would have happened in the absence of the intervention. The net change, the dependent variable of interest, is the difference between the two groups.

Recall that in order to infer causality, one needs to rule out rival alternative explanations. In the evaluation literature, these potential alternative explanations are referred to as threats to internal validity. These alternative explanations, or threats, are catalogued systematically. Evaluation designs are assessed by how well they rule out rival alternative explanations or threats to internal validity. In order to illustrate the nature of these threats, the following discussion uses the example of the classic pretest-posttest design and assesses how well it protects against these threats. Evaluators identified over twenty validity threats.

---

[16] Donald Campbell and Julian Stanley, *Experimental and Quasi-Experimental Designs for Research* (Skokie, Ill.: Rand McNally, 1963).

For purposes of this example, only a subset of those frequently occurring threats are considered.

History

The threat labeled "history" denotes events that happened between the pretest and posttest, outside the confines of the experiment, that may affect the pretest-posttest change. An example of this might be an increase in the inflation rate.

In the case of the classic design, there is little reason to believe that inflation would influence one randomly-assigned grouping and not the other. As the net change is of interest, any extraneous factors, such as inflation are neutralized.

Maturation

Maturation refers to the passage of time and accompanying changes of the subjects under study. If all customers become more exposed to information about new telecommunications services, perhaps this maturation itself, rather than the independent variable, best explains observed changes.

With reference to the classic design, because of random assignment there is little reason to believe that maturation factors will affect one group differentially as all are exposed to the new information.

Instrumentation

Instrumentation denotes any change in the measuring instrument between the pretest and posttest. This could occur after the start of the experiment if the U.S. Department of Commerce changed its definition of the Consumer Price Index (CPI), the Gross Domestic Product-Price Index (GDP-PI), or another inflation index.

If the experimenter has taken care to use the same measuring instrument on the treatment and control groups, any alternation in a measuring instrument, such as the CPI will affect both groups and will not be a concern in terms of comparison.

Testing

Testing refers to improved performance as a result of familiarity with a test or an experiment. Individuals score higher on the second round of testing.

As the same test is used for both experimental and control groups, any advantage gained due to testing will be reflected in both groups. As the net change is of interest, testing will not be a plausible threat to validity.

Selection Bias

Selection bias is the result of permitting membership in treatment and control groups to reflect nonrandom influences. Initial group differences can confound the outcomes. If early on large telecommunications users see an advantage in being in the group that has pricing flexibility, they may initiate successful strategies to be included in the experiment. This could distort the results.

As long as randomization is carried out properly so that individuals are randomly assigned to treatment and control groups, there is no reason to believe that a selection bias is at work. If randomization is not successful, then a good portion of the advantage gained by using control groups may be lost.

Regression to the Mean

Regression to the mean refers to the tendency for extreme scorers on the pretest to be less extreme on the posttest. Regression to the mean operates to increase the obtained pretest-posttest change scores among low pretest scorers. It operates to decrease obtained change scores among persons with high pretest scores. For example, if a treatment were applied only to subjects with particularly low telephone usage on an evaluation instrument prior to treatment, the regression phenomenon predicts that when remeasured their telephone usage scores would be higher even if the treatment had no impact.

Regression to the mean is not a plausible threat in the classic design because individuals were selected randomly, not on the basis of extreme scores. While the experiment which employs a classic design rules out more threats to internal validity than any other designs, there are rival alternative explanations that randomization does not address.

Experimental Mortality

Experimental mortality refers to differential attrition from one of the experimental groups. For instance, a job training program which placed onerous demands on participants might have very positive effects for subjects willing to satisfy those demands. But if those same demands caused many of the subjects to leave the program, an assessment based solely on the experiences of subjects successfully completing the program will overstate the program's impact. Firms with inelastic demand for telecommunications services might, for example, drop out of a flexible pricing experiment, whereas firms with price-sensitive demand may benefit sufficiently such that they stay in an experiment. The mortality of inelastic firms could hurt the validity of the observed results, if the results are improperly analyzed.

Randomization insures that the groups are equivalent at the start. It does not guarantee that at the posttest. The experimenter needs to track attrition in both groups in order to be able to reduce the plausibility of this validity threat. Regarding telecommunications services, it could be that companies in the treatment groups sought modernized services elsewhere, thus possibly contributing to a lower future demand.

Diffusion of Treatment

Diffusion of treatment is relevant when there is sharing of information between participants in the treatment and the control group. This is a problem as it compromises the ability of the control group to answer the question of what would have happened in the absence of the treatment as some of its members possess information intended for the treatment group.

Randomization does deal with this threat to validity. The experimenter needs to assess the probability of this threat and take measures to separate the groups if randomization is considered a potential problem. A mass media advertising campaign by the local telephone company could be seen by both the control and the treatment groups.

Compensatory Equalization of Treatment

Compensatory equalization of treatment refers to a response on the part of administrators who are reluctant to tolerate inequitable treatment and who therefore give

compensatory treatment to the control group. A commission might decide equity and cost-causation considerations require lower rates for the control group because its members are not receiving the benefits of the modern network infrastructure being installed for the treatment group. Lower rates, however, could also spark an increase in the demand for telecommunications services, thus making comparisons between the control and treatment groups difficult.

Randomization does not rule out this threat to validity. As in the diffusion threat, the experimenter has to make a determination as to the plausibility of this threat and build in safeguards.

Compensatory Rivalry and Demoralization

These validity threats are flip sides of a coin. Compensatory rivalry refers to the response of those in the control group who perceive they are receiving less desirable treatment and thus compensate by working harder than normal. A firm in the control group is not likely to accept the possibility of lower sales levels "in the interest of science" and could be expected to act affirmatively to overcome any advantage it perceives its rivals have obtained from participating in an experimental pricing program. Demoralization refers to the situation in which the response of the control group to a situation they perceive as being undesirable is to lower productivity. In either case, the control group does not give an accurate representative of what would have happened in the absence of the intervention.

Randomization does not guard against compensatory rivalry and demoralization. The experimenter must take steps to reduce them as validity threats. This may include asking questions or seeking data from the experimental and control groups about other changes in input and environment that they have observed.

The classic design experiment reduces more validity threats to internal validity than does any other design. It does not rule out all threats; no design can do that. It does, however, allow one to be more confident in making causal statements than does any other design.

Generalizability and Experiments

Generalizability, which is referred to as external validity in the parlance of evaluation research, is whether the lessons learned from the experiment can be extended more broadly. This is the key application issue of evaluation research. At one level, the evaluator asks whether the lessons derived from an experiment applied to a sample of business firms can be extended to the target population from which the sample was drawn. More broadly, generalizability asks whether the experimental findings are portable to other times, settings, and individuals. Randomization of subjects to treatment and control groups handles many problems of internal validity but does not guarantee that experimental results that satisfy the requirements of internal validity are relevant outside the confines of the experiment. For commissions this is an especially important issues as often times a commission will be given information about the successes a small number of firms experienced and will be asked to assume the same results will occur or be generalizable for all firms.

The issue of generalizing to the target population is the easier to deal with. Confidence in the generalizability of program results can be promoted if the following steps are followed:

1. A listing of the population to which one wants to generalize is obtained.
2. A group is randomly selected from that population.
3. That group is randomly assigned to a treatment or control group.

It is obvious that for a great many populations of interest, a listing of names is not available. For example, there is no central listing of school-aged children or convicted felons in the United States. Achieving a random sample is simply not possible in these instances. For those cases in which such a listing does not exist, the costs of conducting the evaluation enormously increase.

Because it is often prohibitive to select samples randomly from a target population, the second aspect of generalizability, namely the extent to which the results obtained in one setting are portable to another, must be addressed. For instance, a program appears to be

successful with college communities, but this does not mean it will be as effective with urban school districts or business training programs. Ideally, experiments must be replicated in different settings, using different population in order to ascertain the transferability of the program.

The relative importance of internal versus external validity needs to be assessed during evaluation planning. Oftentimes the program manager may be unwilling to incur the costs of actions designed to minimize threats to external validity or to insure the results can be generalized. Projections of the demand for new telecommunications services, for example, are often made using the most advanced and sophisticated telecommunications-intensive firms. This makes sense as firms not using advanced services would not be able to make valid projections. On the other hand, unless a discount factor is applied, this unusual sample group will produce results biased in favor of sophisticated versus average users. In this instance external validity is weakened in order to get data. In some instances, nonrandom sampling of the population to which generalizations will be made is either acceptable or unavoidable. For example, one could deliberately sample for heterogeneity, including in the sample a wide variety of types of persons and settings. Another way to deal with the issue of generalizability is to determine the average persons or settings to which one wishes to generalize and sample impressionistically from those groups.

Two variations of the classic design are frequently found in the evaluation of public programs as illustrated in figure 3-4.

Posttest Only Design

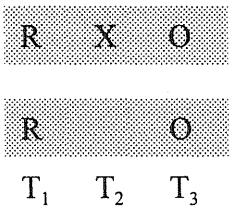


Fig. 3-4. Posttest only design.

Elimination of the pretest yields a posttest only design. The groups are randomly assigned as in the classic design. This design rules out nearly all the same threats to validity as the classic design, as well as insuring that there is no interaction between the pretest and treatment. It also could reduce the threat of experimental mortality as each group is measured only once. It permits only "between group" measurements, as there is no pretest measure, "within group" differences cannot be estimated. For example for each group, information would be lacking on maturation. The absence of the pretest measure takes away a check on the equivalency of the groups at the outset. This approach lowers costs and is frequently used.

Delayed Treatment Design

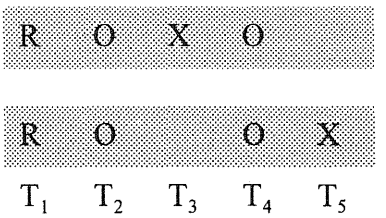


Fig. 3-5. Delayed treatment design.

This experimental design illustrated in figure 3-5 addresses the equity concerns often raised by withholding a treatment or service from a group. In this design the control group receives the treatment at a later date.

Quasi-Experimental Evaluation Designs

While classic design-type experiments may represent the "Cadillac" of evaluation design, considerations of cost and feasibility often permit nothing more than a good "four-door sedan." Quasi-experimental designs can be employed when random assignment is not possible. The category of quasi-experimental designs contains a wide variety of designs ranging from those that use matching or statistical controls to approximate the results obtained through random assignment to time series designs. Those that rely on matched or statistical

controls place a premium on the prior knowledge of the phenomenon under study by the experimenter and the state of theoretical development. To illustrate the issues to which the experimenter must attend, consider the following quasi-experimental designs in figure 3-6.

Untreated Control Group Design with Pretest and Posttest

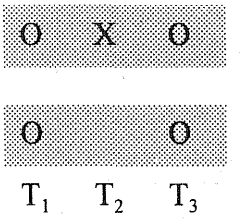


Fig. 3-6. Untreated control group design with pretest and posttest.

The absence of random assignment means that the experimenter must pay special attention to how the groups were formed. Threats of selection, regression to the mean, and selection-maturation must be ruled out. Random assignment in the classic design reduced the plausibility of these validity threats in this quasi-experimental design the experimenter must have detailed knowledge about selection procedures in order to rule them out. For example, if a program's managers are required to show positive outcomes for the programs they manage, then they have a strong incentive to select only those individuals appearing to have the strongest likelihood of success. Picking the fastest growing region or group of companies would be a natural outcome as the manager values increased sales versus increased scientific validity. The manager's selection would lead to an overstatement of the treatment's impact. If the experimenter is constructing the comparison group either through matching or statistical controls, a working theoretical model of the factors that affect the outcome is essential.

If a state had two utilities each serving a similar population, but only one utility used price caps, then a valid pretest and posttest comparison would be possible. If the price-cap utility served a predominantly urban area and the other utility did not, then the untreated control group design may not work.

Interrupted Time Series Design

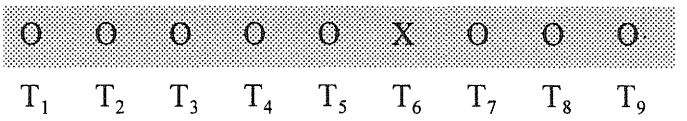


Fig. 3-7. Interrupted time series design.

Time series are collected as illustrated in figure 3-7. The existence of a repeated series of measures permits one to assess the volatility in the time series. The unit being measured serves as its own comparison group prior to the intervention (X). Measures taken prior to the intervention serve to address the counterfactual. Measures taken after the intervention (e.g., accelerated modernization) are intended to speak to the issue of what happened as a result of the program. Interpretation of results can be problematic unless one has a prior expectations about how program impacts will manifest themselves over time.

Thus, if utility operating revenues grew at 3 percent per year from T_1 to T_5, and then grew at 6 percent per year after the institution of pricing flexibility, this evaluation example would conclude (with many caveats) that the demand growth rate doubled after the institution of pricing flexibility. Having multiple observations (as opposed to one pre and one post) allows the evaluator to examine the trends for counterfactual explanations. It could be the case that while the prior average demand growth was 3 percent, the demand line was trending upward such that the year's demand immediately prior to the pricing flexibility was 6 percent. In this instance an evaluator would be reluctant to attribute all of the demand growth to the pricing flexibility granted.

If an individual is being measured repeatedly, reactivity problems could be a factor. Instrumentation, change in that measuring instrument, is always a concern in time series. The experimenter must be cognizant of any changes in definitions or record keeping.

Interrupted Time Series with Nonequivalent "No Treatment" Comparison Group

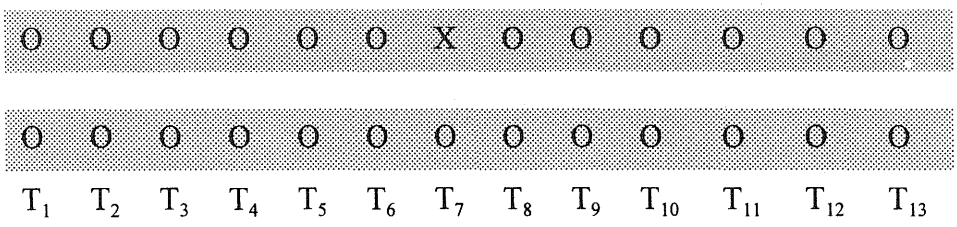


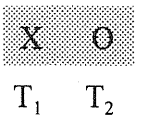
Fig. 3-8. Interrupted time series with nonequivalent no treatment comparison group.

A "no treatment" comparison group is added to the time series design as illustrated in figure 3-8. This provides another check on the counterfactual, as well as rendering history a less plausible threat than in the time series design. Insuring comparable data from a comparable entity is critical for this design. This approach would allow a better examination of trend lines and would, in the previous example, offer a better foundation for identifying the demand trend line.

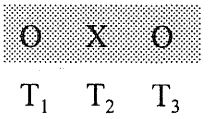
These quasi-experimental designs in figure 3-9, while often used in the public sector, rule out virtually none of the threats to internal validity highlighted earlier. History, selection, testing, instrumentation, maturation, and regression are plausible threats. The designs are often employed when evaluation issues are not considered in the program design phase, when evaluation is an after thought, when there are no data to permit a better design, and when "quick and dirty" is acceptable.

Preexperimental Evaluation Designs

One Shot Case Study



One Group Pretest-Posttest



Status Group Posttest

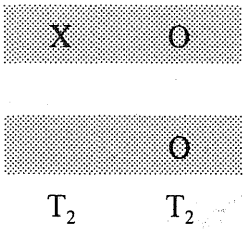


Fig. 3-9. Preexperimental evaluation designs.

The choice of a true experimental design or pre- or quasi-experimental design for a state commission can be determined by the answers to the following questions.

1. What resources are available for the evaluation?

   Experimental designs are the most costly analytical approach, partially, because of the need to select control groups. If available resources are estimated to be inadequate, then an evaluation design should not be used because poorly financed and constructed designs may produce erroneous results. Reliable information can be obtained by preexperiments, except the ability to infer cause and effect linkages is much less.

2. Should cause and effect be determined or are reasonable approximate impact measures desirable?

   If it is truly unknown whether a program or a part of a program produces the results observed, then an experimental design is needed. Alternately, if the program is not going to be modified, then a preexperimental design would be more appropriate.

3. What lead time exists to design the evaluation?

   It takes time to construct control groups and if the program has already begun or lead time is short, then a quasi or preexperimental design could be used. If adequate time exists, then a control group is desirable and preferred.

4. What constraints exist regarding random assignment?

   If random assignment of customers or areas is acceptable, then an experimental design is the preferred option. The other options presented above can provide useful results if random assignment is not feasible.

The analyst can weigh and sort the answers and see which design approach best fits the needs and resources available to the commission. The design choices are equally valid, but each design is more appropriate for certain circumstances and resource levels.

It is better to have an adequately funded design, with known limitations, than an inadequately constructed classic design experiment. This is because the design's flaws will be unknown and hard to discover.

Examples of Utility-Oriented Evaluation Research

Evaluation research has been embraced warmly by social policy researchers. This is partially due to the self-critical nature of evaluation research and partially to the ability of evaluation research to provide answers to the basic "does it work?" concern of program managers and policymakers. The self-critical nature of evaluation research allows the

researcher and the research users to easily understand the strengths and limitations of any findings. As indicated by the above review and critique, trade off of various legitimate concerns can increase or decrease the explanatory power of evaluation research. Being able to answer the "does it work?" question meets the public stewardship obligations of program managers and policymakers.

A voluminous body of literature is directed at assessing the performance of many social programs. In addition to the myriad social program evaluations, evaluations have been conducted for a wide array of other public programs, including some at state regulatory commissions. The purpose of this section is to highlight some evaluation research in areas within the domain of state commissions.

Walsh and Aleong reviewed twenty-two evaluations of residential energy conservation programs in order to assess their usefulness to utilities and regulators.[17] All of the evaluations identified energy savings as the primary objective. Cost effectiveness and widespread participation were secondary objectives. They found that most estimates of energy savings were based on simple pretest-posttest designs which were normalized for weather conditions (using the Princeton Scorekeeping Method). Control groups typically were not employed in the evaluations. Evaluation design issues were given little explicit attention. They found little consistency in assumptions and methodology across the cost-effectiveness studies reviewed. The consequence of these methodological inadequacies is that few definitive statements concerning the effectiveness of residential energy conservation programs can be made.

While the analytical standards that typify evaluations of energy conservation programs have been deficient, some studies have paid attention to measurement and design concerns. Nadel and Ticknor assessed four approaches to estimate energy savings: (1) engineering estimates, (2) comparison of program participant's pre and post daily killowatthour (kWh) use with that of a control group, (3) comparison of program participant's pre and post daily kWh

---

[17] Roberta Walsh and John Aleong, "The Design of Residential Energy Conservation Programs Under Least-Cost Planning: The Role of Evaluation," *Policy Studies Journal* 20: 102-111.

use without a control group but with a survey on changes in program participants' energy use patterns, and (4) conditional demand analysis.[18]

They found that while the engineering method was easy to use and required the least data and data analysis of the four methods assessed, it produced inaccurate estimates for customer subgroups. The daily kWh with the "control-group" method required several years of billing data and a carefully selected control group but produced more reliable estimates than the engineering model. The survey approach did not work well and was deemed an inadequate substitute for the control group. Finally the conditional demand method was the most complex. It required several years of billing data and powerful statistical programs.

Similarly Schultz and Eto examined a number of important measurement issues surrounding shared savings programs.[19] They discuss three options for estimating load reductions: (1) estimates can be fixed for each measure to be promoted by each and every program, (2) participants should agree prior to implementation on an explicit savings methodology, or (3) participants should agree that load reductions will be established after program implementation, based upon a particular methodology and schedule for monitoring.

Other measurement issues identified that are of interest for evaluation research are whether an analysis of customer bills from all participants or submetering the loads of high efficiency appliances represents the appropriate source for data on reliable energy savings. The measurement of avoided costs also is critical in a shared savings scheme as they directly affect shareholder earnings. All these issues define an increasingly important role for evaluation research in utility regulation.

---

[18] Steven Nadel and Malcolm Ticknor, "Electricity Savings from a Small Commercial and Industrial Lighting Retrofit Program: Approaches and Results," *Policy Studies Journal* 20, (1992): 48-56.

[19] Don Schultz and Joseph Eto, "Carrots and Sticks: Shared-Savings Incentive Programs for Energy Efficiency," *The Electricity Journal* (December 1992): 32-46.

Fowler addressed issues relevant in the benefit-cost analysis of utility conservation programs. He identified four ways to consider the question of whose costs and whose benefits should be considered. They are as follows:[20]

1. The participant's test compares costs and benefits just for the customers who are actually participating in the programs.

2. The nonparticipant's test compares costs and benefits just for a utility's customers who are not directly participating in the particular program being examined.

3. The all-ratepayer test compares benefits and costs for all the customers in the utility, both those who participate and those who do not.

4. The utility test compares benefits and costs just for the utility company, and distinguishes from the benefits and costs borne by the utility's customers.

Fowler applied the "all ratepayer" perspective to compare the costs and benefits of two new energy conservation and load management products in Sacramento and found them to be sound investments. He argued that a well-executed cost-benefit study provides a "level playing field" for comparing conservation/load management programs with the purchase-power/generation alternatives.

Newcomb employed a quasi experimental design to assess the cost effectiveness of conservation programs directed at residential customers.[21] In the early stages of energy conservation programs, estimates of the potential electricity savings came from engineering studies. These estimates often produced inaccurate estimates of savings actually realized by residential customers. Newcomb determined that actual field measurements of program participants were necessary in order to produce reliable estimates of electricity use. The vehicle for his study was the Low Income Electric Program (LIEP) in Seattle. The absence of reliable time series data precluded the use of a time series design. Weather fluctuations,

---

[20] Hugh Fowler, "Marketing Energy Conservation in an Environment of Abundance," *Policy Studies Journal* 20 (1992): 76-86.

[21] Tim Newcomb, "Conservation Program Evaluations: The Control of Self-Selection Bias," *Evaluation Review* 8 (1984): 425-440.

and rate increases resulting in reduced electricity consumption rendered a simple pretest-posttest design of little value. The volunteer bias evident in program participants made them different from the average low-income customer regarding education and attitude. A comparison drawn from nonparticipants would, therefore, be inappropriate. Newcomb opted to use the groups of people who signed up during the second year of the program's operation as the control group.

Three techniques were used to establish the comparability of the two groups. Data collected from the energy audit were used to compare the age, size of home, number of occupants, and type of space heating. A second source of data came from a mail survey to both groups to obtain information on other conservation measures taken and the reason for participation in the LIEP program. Finally, meter-verified energy consumption figures were derived for both groups for two months. All supported the equivalency of the two groups. He estimated that the program resulted in 3,422 kWh year of electricity conservation, valued at $485 (1981 dollars) over the thirty-year lifetime of the weatherization measures.

Hirst advocated a more pronounced role for evaluation research in demand-side management (DSM) programs.[22] Shared savings mechanisms with their incentives for utilities to minimize costs and maximize net benefits provide an excellent opportunity for evaluation research to make a contribution to the policy debate. He enumerated the purposes to which such evaluation can be put as follows:

1. document the energy savings, load reduction, and cost effectiveness of DSM programs,
2. show ways to improve programs by increasing participation rates, raising energy savings or cutting costs,
3. suggest ways to improve the design of future DSM programs,
4. support DSM budgets before the utility's budget committee, and
5. provide data to strengthen the company's load forecasts and resource planning.

---

[22] Eric Hirst, "The Role of Evaluation When Electric Utilities Get Financial Incentives for Their DSM Programs," *Evaluation Review* 117 (1992): 95.

In a similar vein, Selwyn argued for the primacy of the articulation of policy goals and measurements of benefits in his analysis of a state telecommunications agreement.[23] Narrowly construed engineering goals, expressed in terms of physical resources are deemed inadequate. Assertions of benefits are found to be inadequate substitutes for the quantification and communications of benefits to be derived from the new agreement. These goal clarification concerns fit comfortably within the evaluation research framework.

It is apparent from this brief review of evaluation studies for energy conservation programs that evaluations are still in the embryonic stage of development at state commissions. This should change when it is recognized that evaluation studies can have explicit findings that regulators can have confidence in because their logic is understood and less subject to the often unknown interpretation ambiguities associated with other less introspective analytical approaches. Applications of the tools of evaluation research can increase the usefulness of the analyses and the confidence that can be placed in them.

FCC Evaluation of Price Caps

When the Federal Communications Commission (FCC) instituted price caps for AT&T and for LECs, a provision for a performance review was included in the order. As of this writing the LEC evaluation is not complete, but the performance review of AT&T's experience under price caps has been completed and offers interesting findings. While the performance review may not score high when evaluated against the classic experimental design model, it is noteworthy in terms of the level of effort and the commitment made to review the pricing reform. From an evaluator's perspective the FCC price-caps review combines parts of a "before and after study" with time series data.

---

[23] Lee L. Selwyn, *A Policy Analysis of the Second Vermont Telecommunications Agreement* (prepared for the Public Contract Advocate State of Vermont, January 30, 1992).

Quality of Service

Evidence that service quality should increase when (1) new technologies are employed, (2) competition is faced by a carrier, and (3) price-cap rules are in effect is provided by the 1993 FCC price-cap performance review of AT&T. In this review both the FCC and AT&T saw quality improvements, even though no specific QOS adjustment factor was used for AT&T. In its analysis the FCC reported that the study by Schmalensee and Rohlfs, submitted by AT&T, had found "gains in service quality." The FCC also presented several charts that reveal an improvement in reliability and service quality.

The FCC did not draw strong conclusions regarding QOS because the data were not conclusive. In the case of AT&T, the FCC reaffirmed its quality and reliability oversight function and said it would rely on the authority it had under Section 208 and 214 to handle complaints or to initiate its own actions.

Ceiling Pricing

One concern regulators have is whether or not price-cap regulated utilities will set their prices for captive customers at the ceiling or maximum price allowed. Some have asserted that utilities would not use ceiling prices because of competitive pressures and because of their interest in avoiding reregulation. The 1993 FCC evaluation of AT&T's voluntarily selected pricing behaviors revealed that the ceiling price was used for the least competitive basket of services, Basket 1 (residential services). Examination of table 3-1 reveals that on average, AT&T used 99.68 percent of its price-cap ceiling for the less competitive Basket 1. As shown in table 3-1, a listing of the baskets in terms of the percent of price ceiling pricing authority used reveals a clear correlation with the aggregate amount of competition in each basket. Generally speaking, this type of pricing behavior seems to confirm the expectations of ceiling pricing under price caps where baskets face different degrees of competition.

TABLE 3-1

PERCENT OF PRICE-CAP CEILING PRICING AUTHORITY USED BY AT&T

Basket	Percent of Price-Cap Ceiling Pricing Authority Used by AT&T	Type of Services
Basket 1	99.6%	Residential and business
Basket 2	98.43	800 services
Former Basket 3	97.0	WATS and private line
Restructured Basket 3	95.0	Private line

Source: Authors' construct.

Profitability

The AT&T price-cap plan does not include profit sharing and the FCC's performance review shows that AT&T's interstate profits were higher in 1992 (a 12.77 percent rate of return) than they were in 1989 (an 11 percent rate of return). These higher profits occurred in markets that are thought to be more competitive in 1992 than they were in 1989.

Consumer Dividend

A consumer dividend is a regulatory device invented by the FCC in its AT&T price-cap proceeding in response to its partial concern that setting a productivity factor equal to the historical productivity levels would not necessarily force carriers to share the efficiencies derived from incentive regulation with consumers.

In order to address this concern, the FCC effectively adjusted the carrier productivity factor upward by 0.5 percent. The benefit to consumers is that the dividend assures that the consumers will be the first efficiency beneficiaries because prices will be at least 0.5 percent

lower than otherwise. Use of this tool represented a direct, tangible, and up-front formula-based device to ensure tangible consumer benefit.

Administratively the consumer dividend is added to the productivity factor in the price-cap formula, as shown below. Conceptually and procedurally it is compatible with or may be used as an alternative or as a complement to traditional profit sharing.

% change in price cap	=	% change in inflation	-	% offset productivity	-	consumer dividend	+/-	exogenous impact adjustment

In the FCC performance review of AT&T, half of the total consumer benefit identified came from the consumer dividend. The FCC performed the following calculation to determine consumer benefit over the three-year price-cap period.

Total amount by which AT&T has exceeded regulatory requirements by pricing below the cap	$891 million
Total consumer productivity dividend	$899 million
Total consumer benefit	$1,780 million

The fact that AT&T met the consumer dividend target, increased interstate profits, and paid attractive shareholder dividends during the three-year price-caps period suggests that productivity increases occurred. Based on these data the consumer dividend appears to have been the single largest source of benefit for the consumer. Importantly, the consumer dividend does not appear to have harmed the shareholder dividend or the interstate rate of return.

Given the prominence accorded the consumer dividend in the performance review, it is evident that the FCC felt that the dividend plays an important and visible role in assuring the public that the price-caps formula balances ratepayer and shareholder interests.

Future Prospects for Evaluation Research

In order to sustain the current rekindling of interest in evaluation research among practitioners, several conditions must be met. Performance, as measured by careful evaluations, must be shown to matter to decisionmakers. This requires that incentives be tied to evaluation results. Interest in evaluation research that is fueled only by organizational self-defense concerns will not likely be sustained after any perceived threat has subsided. Traditionally, incentives in the budget process have not reinforced the belief that performance on outcome measures in one year is critical in the determinations of next year's allocation. It is unreasonable to expect that radical change in the budgeting process is on the horizon. Past experiences with program planning and budgeting system (PPBS) and zero base budgeting (ZBB) indicate that the largely incremental budgeting process resists wholesale change in firms and governmental agencies. If practitioners are to incur the costs of evaluation research, commensurate benefits must be perceived.

Regulators are in a unique position as they oversee various regulatory reforms. Unlike other situations where decisionmakers have to rely on a budget cycle to make changes in a successful or unsuccessful program, regulators can initiate changes at scheduled and unscheduled time periods. Regulatory reforms often are instituted for a set period of time and no artificial budgeting type constraints exist on the range of actions open to commissions: although, in practice commissions favor incremental changes. Further, the general oversight and monitoring responsibilities of commissions are sufficient to allow intervention, if needed.

If signals are sent that evaluation research is important in the budget process, organizations might choose to invest in internal evaluation capacity. Research and evaluation departments in agencies, firms, and universities have felt the budget ax fall during lean periods. Investments in an organization's ability to carry out and react to systematic evaluations are essential if evaluation research is to develop as a credible input into decisionmaking. Even if an organization were to rely on external evaluation consultants, the internal capacity to communicate with outside evaluators is necessary for it to be an informed consumer.

In addition to building human capital, organizational commitment to the maintenance of data bases that can be used to answer the questions of outcome, impact, and efficiency of a program is needed. The existence of performance-based data will enhance the ability to make meaningful judgments about the value of public programs.

In order to unlock the potentially powerful contribution of evaluation research to the policy debate, evaluation must be considered in the program design phase. While the actual conduct of outcome/impact evaluation follows the implementation and formulation of a program, its genesis is in the initial stage of the policy process. Planning for evaluation allows for the possibility of the strongest possible causal inferences. If evaluation is not considered until implementation, the set of designs shrinks as does the time at hand to collect relevant data.

Conclusion

Evaluation research falls neatly into the traditional regulatory concept of "oversight." Use of evaluation research improves the ability of a state commission to exercise its oversight and monitoring obligations. Oversight occurs as a commission exercises its general powers and "watches for and reacts to" apparent troublesome actions of utilities. In an oversight mode a commission can take informal or formal investigations and implement any needed corrective actions. From an evaluation research perspective, the commission can be said to be assessing how the actual behavior of a regulated utility conforms to regulatory rules and goals and then taking the necessary corrective action. States with different goals or rules would, accordingly, differ regarding whether a particular utility behavior needs corrective action.

Evaluation research shows clear promise as a means of improving how commissions exercise their oversight obligations. Oversight generally suffers from its holistic perspective: it is predicated upon the commission constantly scanning its environment to detect, assess, and make corrective actions. This overarching and reactive style tends to have commission agendas seem to lurch from one crisis to another and tends to ignore successes.

The main weakness of oversight is that it is largely complaint-driven. While the tried-and-true investigatory processes triggered by complaints may yield satisfactory results, the essential weakness is that a commission does not know whether it is focusing on the right thing. All the commission knows is that it has properly reacted to a staff, consumer, or utility complaint. Evaluation research offers a state commission the ability to systematically compare behavior to goal achievement in a way that objectively identifies successes, as well as areas requiring action. The advantage of evaluation is that it allows a state commission to prioritize and to have confidence that it is addressing the right issues and not simply responding to the "squeaky wheel."

Evaluation research also lets a state commission design and seek consensus on the measures it will use to monitor performance and impact. As will be shown in other sections of this report, this feature is especially attractive as the usual way commissions see new data is from self-interested parties. The commission is then often in the position of determining which party's data is most trustworthy. It is better to have firm evaluation, or benchmark, or monitoring data that all parties can share and supplement. Consequently, any subsequent debate is more properly focused on the meaning of the data—how it effects goal accomplishment—rather than on the methods used to generate the data.

Evaluation research, of course, has very real limitations. The most important is the difficulty in identifying goals and reaching a consensus on the goals to be achieved by a particular pricing reform. Measuring and interpreting the impact of a pricing reform is another area of potential difficulty for regulators. Further, the ability to use true control groups is often lacking for regulators. As shown in this chapter, once recognized by a policymaker, these limitations can be addressed and resolved.

CHAPTER 4

USING QUALITATIVE METHODS TO EVALUATE REGULATORY REFORMS IN REGULATION

Regulatory pricing reforms, while seeming to be deceptively simple, are generally quite complex. Some of the complexity occurs because many important issues combine and are expressed in a single pricing rule. Price caps appear at first glance to look like a simple formula that requires only annual updates and minor arithmetic manipulation in order to determine the price cap. Yet the price cap base price and productivity factor are now widely recognized to be "one number" composites of a number of depreciation, separations, competitive pressure, management philosophy, regulatory environment, universal service, equity, and QOS factors.

The impact of regulatory pricing reforms, accordingly, is necessarily complex and may not be fully susceptible to measurement via traditional quantitative techniques. Qualitative, "nonquantitative," techniques were developed that can and have been used in regulatory settings to provide a better understanding of the impact of regulatory pricing reforms. Because most regulatory analysts are trained in quantitative methods, the use of qualitative methods is first discussed and then its use in regulation is presented.

Introduction

In the realms of science, the search for knowledge is acknowledged as the goal. It is postulated that there are definite, known laws governing our physical reality that dictate the way entities interact. Scientists strive to uncover those laws through research. In order to determine whether the results of research contribute to our knowledge of these theoretical laws, scientists have created an approach widely known as the scientific method. However, this method is only useful in examining that part of reality that can be meaningfully translated into some type of numbering system. The scientific quantitative method typically uses

statistical analysis, and checks results against established mathematical parameters to determine "statistical significance." This use of an objective, established standard against which research results can be measured gives all consumers of the research some assurance that these research results are "true," i.e., can be replicated and verified by others and are publicly disseminated.

For example, in the telecommunications industry, a researchable question that commissions often ask is whether there is a significant relationship between the type of alternative regulation in a state and the level of infrastructure development. Using quantitative measurement indicators, such as the number of fiber miles and/or the number of digital switches, it would be possible to scientifically establish the relationship between the type of regulation and investment in infrastructure. Statistical analysis of the data could be done by research scientists, whether they worked for a telecommunications company, a consumer advocates group, a state commission, or an investment house. Given that they all accepted the data as being valid, they should all get the same mathematical results. Comparing these results against existing statistical tables would indicate whether the results were statistically significant, and consequently, whether there is a relationship that can be statistically verified.

Because this type of research can allow diverse groups to arrive at the same conclusions, it has great appeal in decisionmaking situations. If decisionmakers are questioned, praised, or criticized for their subsequent decisions, they can point to the mathematical basis of the conclusions and justify their decisions because they are based on scientifically supported quantitative evidence. There is also the feeling that because there is tangible evidence, the decision—and subsequently, the decisionmakers are "right." The scientific approach yields reliable quantitative data that serves as one source of information that decisionmakers can use. The values of the decisionmaker, and the legal and institutional constraints faced are other inputs that will affect the ultimate decision. The appeal of scientific and quantitative methods is that the explicit and objective methods used mean the reliability and validity of the data can be easily checked.

As regulators are aware, use of a quantitative approach does not mean that all disputes are instantly resolved by a quantitative study. Rather, it means that once the biases of competing parties are discounted, it is possible to reach an agreement on the validity of the conclusions reached.[1] Parties can still disagree over whether a bottle is half-full or half-empty but not over the fact that verifiable quantitative data reveals the bottle to be at 50 percent capacity.

Qualitative Research

In addition to the quantitative studies used by commissions, qualitative research is another important class of studies. Strauss and Corbin offer the following definition:

> By the term qualitative research we mean any kind of research that produces findings not arrived at by means of statistical procedures or other means of quantification. It can refer to research about persons' lives, stories, behavior, but also about organizational functioning, social movements, or interactional relationships.[2]

Strauss and Corbin elaborate on when and why a researcher would consider qualitative instead of quantitative research.

> Some areas of study naturally lend themselves more to qualitative types of research, for instance, research that attempts to uncover the nature of person's experiences with a phenomenon, like illness, religious conversion, or addiction. Qualitative methods can be used to uncover and understand what lies behind any phenomenon about which little is yet known.[3]

---

[1]As any paid consultant or party can criticize a quantitative study endlessly, these criticisms should not be interpreted as indicating that the quantitative approach is intrinsically flawed. Instead, such criticisms should be regarded as arguments about the meaning of any findings.

[2]Anselm Strauss and Juliet Corbin, *Basics of Qualitative Research* (Newbury Park, CA: Sage, 1990), 17.

[3]Ibid., 19.

For regulatory purposes ***qualitative analysis can be defined as the use of systematic, nonquantitative procedures to obtain information about an event, entity, or object of interest.*** The approach stresses understanding of context, structure, process, and relationships. Qualitative techniques can be used to forecast, define, categorize, and compare.

Qualitative research is often used to explore an area; build an information base, understanding or framework, which can be subsequently used to construct quantitatively testable hypotheses. For example, before one could test for a possible relationship between infrastructure and the type of alternative regulation, there had to be an interest in exploring such a relationship, an interest that could have first been awakened by a utility testifying before a commission about why alternative regulation would stimulate more infrastructure investment than traditional rate-of-return regulation. This kind of utility testimony, or report, often is not quantitative, even though references may be made to, say, a business where sixty-seven new jobs were lost or created due to the presence or absence of infrastructure. At this stage the insight and vision of the utility manager establishes a qualitative model that states that a positive relationship exists between economic development infrastructure and the type of regulation. This approach relies on intuition. Its weakness is directly related to how strongly the values of the utility manager (in this example) have produced a qualitative model that does or does not match reality. It is not the fact that all parties have values, goals, and visions that cause any methodological problems. Rather, it is whether the utility's qualitative infrastructure-economic development model describes a linkage that actually exists.

The linkage can subsequently be examined through the use of quantitative and qualitative methods. In this example, this exploration might have involved talking to utility executives about future investment strategies, why they were or were not planning to change analog to digital switches, or how soon they were planning to do this. Additionally, state development officials, large users, and engineering consultants could be queried. At this stage, insights, informal decision rules, approaches, and constraints, are identified. A utility manager might say, "Traditionally, we wait for a user to request a new service, but we find that if we put facilities and services in place first that more users seem to appear." This qualitative statement may or may not ultimately prove to be true. It could be that a

subsequent quantitative analysis shows no difference between traditional and accelerated deployment. This second analysis might identify districts with traditional deployment and compare the demand to districts with accelerated deployment. This would be a quantitative study. A qualitative study would seek more information about the "more users seem to appear" notion. Ways to do this include talking to managers of telecommunications intensive firms about their decision rules, examining the assumptions and procedures underlying the utility's capacity expansion model, and identifying possible alternative explanations, such as an upturn in the local economy, or a change in the utility's marketing effort in an area. Such an exploration would be characterized as qualitative research. Out of this qualitative research, the following hypothesis might be constructed: the more telephone companies invest in infrastructure, the more demand occurs. This hypothesis could be tested using quantitative research methods. Used in this manner, the two approaches compliment each other.

But sometimes, utilities, consumer advocates, commission staff, residential consumers, and business managers make predictive statements before commissions that can not easily be subjected to quantitative analysis. Often, this type of testimony comes in one of three forms: (1) a case study or pilot project, (2) the personal experience of some individual or group, or (3) the construction of various scenarios (both heavenly and horrific) of what will happen in the state if the commission does or does not act in a prescribed manner.

Pilot projects are difficult to analyze quantitatively for a number of reasons. The one reason most often cited is the "Hawthorne Effect," which means that individuals or entities in a pilot project or experiment tend to perform better simply because they are selected as a pilot. A suburban housing development that is wired for broadband and whose residents are informed that they have been selected to participate in an innovative pilot program will generally act differently than if they did not know they were in a study. In this instance, not only do the residents behave differently, so too do the staff and managers of the utility because of their pride in being part of an innovative project.

Qualitative research approaches can, along with the program evaluation methods examined in Chapter 4, help regulators better understand what part of a pilot project can be accepted as (1) a "bell weather" or (2) a benchmark or (3) a prediction of future trends or (4) an indicator of whether some or all of the improvements cited are due to the excitement

surrounding an experiment. If a qualitative analysis revealed some of the following, then a Hawthorne-type effect may be present: charismatic management, crisis or do-or-die meetings, significant external funding, intensive training, high level of publicity, self-selection of participants, lack of a multi-year examination, and no systematic comparison with other firms. A qualitative researcher would be justified in discounting the purported results or impacts if any of the preceding factors were seen as unduly influencing the pilot.

Similarly, the testimony of the presidents of six telecommunications-intensive firms regarding the new jobs created and the revenues realized may also not be readily susceptible to quantitative analysis. Innovators tend to be uncritical enthusiasts of their favorite approaches. For example, some analysts have noted that firms that originally claimed success by following the Demming Total Quality Management (TQM) approach did no better than other firms, even though the managers felt their perceived successes were due to TQM.

With the six presidential testimonials a qualitative researcher might ask a contextual question, such as "Did you feel the utility was responsive to your needs?" A "yes" answer may actually undercut the testimonial, as it implies that all is well; unless a regulator feels that subsequently all new customers will not necessarily receive an equivalent level of responsiveness. In this approach, good insightful questions can overcome a lack of quantitative comparative data.

Using Qualitative Information

How influential is qualitative information? That depends on a lot of factors but it is fair to say that qualitative testimony can persuade commissions to adopt certain courses of action. Part of the reason for this persuasiveness may be that commissions do not always have ways to evaluate this type of readily available testimony. A second reason is that commission procedures allow experts to file qualitative testimony and unless their testimony is successfully challenged it can stand as evidence. Qualitative forecasts, scenarios, propositions, and hypotheses are commonly advanced in expert testimony. A third reason for this persuasiveness is that the expert can eliminate or discount any alternative explanations that would undermine the expert's conclusions. Absent good cross-examination, an expert's

qualitative statements may carry as much or more weight than an empirical quantitative study, because all flaws are eliminated or hidden or are not easily susceptible to discovery.

Lastly, qualitative information is often accepted because it is understandable. Many regulators and policymakers feel they can better understand and evaluate qualitative information because it can be filtered through their professional experience. This is often a more direct and useful approach for commissioners than having the commission's technical staff weigh and evaluate conflicting quantitative evidence presented by dueling consultants.

Consider, for instance, that many commissions have accepted the need for alternative regulation? Why have they done so? "Increased competition and a need for a more flexible regulatory approach to respond to competitive pressures have been cited as the primary reasons to consider alternatives to rate-of-return regulation."[4] This is a qualitative proposition and probably can not be empirically proven, although it is widely accepted.

To further illustrate this point, a company may state that unless it has unregulated profits, it cannot commit more financial resources to infrastructure upgrades. The company suggests that without upgrades, the state will not be able to compete with neighboring states in bringing in companies that can provide jobs. Others may make the counter argument that the company will still invest in infrastructure regardless of whether the company receives unregulated profits and that to allow a company to operate without some oversight of their profits will lead to anticompetitive practices.

How accurate are either of these qualitative predictions? Generally, the qualitative propositions advanced by the different parties are not immediately provable. This can often be overlooked as most qualitative testimony looks scientific because it contains citations of empirical data. Unless a formal quantitative method is used, the facts used for support should be looked at with skepticism. "Facts" not directly derived from a quantitative method do not necessarily have a verifiable causal link to the proposition being advanced.

---

[4]Ronald L. Howe, *Framing Paper Alternatives to Rate of Return Regulation* (Lansing, MI: Michigan Public Service Commission, Office of Planning, Policy and Evaluation, June 1989), 1.

In another example, it could be asserted that a state had a 5 percent increase in total employment due to the stimulative effect of its price caps pricing flexibility program upon demand. From an analytical perspective, however, the 5 percent should be regarded as an unlinked piece of data with no proven tie to the price cap reform. Either through additional qualitative research or by traditional quantitative research it would be possible to see if other factors explain the "5 percent," or if other price cap states had a similar growth in demand. Without independent analysis the "5 percent" assertion can only be regarded as a suggestive and as an unproven qualitative proposition.

With regard to any type of research, whether the conclusions of research are accepted or not is often dependent on the acceptance of the research data's validity. With quantitative data, there are accepted mathematical procedures to test data for validity and reliability. With qualitative data, in a proceeding it is possible that all parties are telling the truth as they perceive it, but that the differences in their underlying assumptions produce a line of myriad possibilities, all of which are possible, depending on factors such as the speed of technology development, consumer demand, the type of regulation, and the state of the general economy.

Qualitative information is usually accepted if the decisionmaker feels it has "face validity." That is, after the information has been filtered through the experience of the regulator, and exposed to criticism, then a gut-level trust or distrust of the information is formed by each commissioner, regulatory staffer, and party. Face validity simply means that "on the face of it" the information is accepted valid. Each participant in the regulatory process may not accord "face validity" to the same piece of qualitative information. The subjective nature of the validation process is a weakness of the qualitative approach. In response to this criticism, it is often argued that more damage is done when decisionmakers pretend to understand complex empirical studies that may in fact be flawed, than when they consciously accept and use qualitative information with all of its limitations.

Qualitative Methods

A number of methods were developed that can be used to analyze qualitative information. These methods allow the user to deal with the validity and linkage concerns raised above. Six of these methods are briefly described below and are suitable for use in a regulatory setting. A bibliography of qualitative methods references is presented at the end of this report.

Assumption Analysis

Researchers developed some easily applied methods for determining the value of qualitative data in terms of its validity. These methods were developed to address two objectives. The first is to determine the assumptions underlying the data. The second is to determine the accuracy of those assumptions and consequently, the statements made that are based upon them.

Let us return to our example of telephone infrastructure investment. A utility may say it will have to withhold, delay, or eliminate infrastructure upgrades unless it receives a certain type of alternative regulation. As a part of its reasoning, the company suggests that it is not economically in the company's best interests to make such investments unless it is allowed unregulated profits.

From a qualitative perspective, a researcher could ask: What are some of the possible assumptions upon which the company's statement is based? Do these make sense? Do they accurately reflect how economically rational business entities make decisions? The first question identifies assumptions and the subsequent questions address the real world fit of the assumptions.

Continuing the above example at least four reasonable assumptions can be drawn:

1. The utility believes it can continue to do business in an increasingly competitive market without making any significant infrastructure upgrades.

2. There is not enough customer demand to justify investment unless the utility gets some kind of an additional regulatory "break" or incentive.

3. The company believes they will make money on new services and wants the freedom to have profits at levels greater than those allowed under traditional regulation.

4. The state needs new services to attract new businesses and jobs and the need for new jobs will encourage the commission to adopt the requested regulatory reforms.

The next step in an assumption analysis is the most difficult and requires the analyst to individually examine each assumption for its fit with real world situations, logic, and incentives. To illustrate this it is necessary to continue the assumption analysis with the "continue to do business" assumption noted above. Implicit in the initial statement of the utility is the notion that if the utility does not get what it wants—unregulated profits—it will not make infrastructure upgrades. The analyst can ask, "How realistic is it that a utility facing emerging competition can delay or reduce its infrastructure building program?" Delaying strategies will hurt the competitive position of the utility and increase the pressure from large revenue, sophisticated customers to have the commission allow local exchange competition from new entrants. Because the primary asset of new entrants is their ability to use new technology for targeted groups, a delay strategy by the incumbent utility could cause a loss of market share. Based upon this simple and limited analysis of the fit of the assumption with real world conditions, an analyst could feel comfortable rejecting the "continue to do business assumption" because a telecommunications utility can not stop investing in infrastructure and still survive.

This type of analysis would be continued for each of the four identified assumptions. The outcome of the analysis is a better understanding of the strategic objectives of the utility and the practical range of tactical options available to it. In this example, a commission could conclude that if it decided to continue profit regulation, the utility would still have sufficient resources and incentives to continue its infrastructure modernization in the key economic sectors and areas of the state. It could also conclude that if it gave the utility the opportunity for unregulated profits, then the commission should have more input in the utility's infrastructure deployment plan, particularly with respect to rural areas.

Assumption analysis does not yield point predictions or have statistical significance tests. Its primary analytical advantage is that it forces the analyst to examine assumptions

underlying policy statements and to then identify the range of realistic options available to the commission. It is a qualitative, yet systematic, way of doing what some of the very best business and government strategists do as their mainstay. The public interest is well-served because assumption analysis can do a good job of separating rhetoric from reality and allow a better set of solution strategies to emerge. Interestingly, quantitative studies rarely have their underlying assumptions examined and suffer, accordingly.

Triangulation

This qualitative method may be familiar to regulators and involves having the analyst link assertions, assumptions, propositions, or cited facts and cited examples to at least two other independent information sources. The key here is the use of independent sources of information from entities not connected with the party being examined. A second-best source of data is from other parties in a proceeding, but this may be of limited usefulness for the purpose of triangulation.

The more independent the source of the information the better the results. Continuing the qualitative examination of the "continue to do business" assumption, an analyst could attempt to triangulate from positive and counter-factual directions. Positive triangulation involves thinking of supportive sources while counter-factual triangulation looks for the opposite. It is generally very useful to ask positive and counter-factual questions of each information source examined. As is the case with an assumption analysis, it is the creativity, energy, time frame, resource base, and technical expertise of the qualitative analyst that will determine the triangulation points examined.

In the "continuing to do business" example, an analyst could ask one or more of the Wall Street financial houses for their assessment of the advantages and disadvantages of a utility following a delayed infrastructure investment strategy. A second triangulation point could be seeking the opinions of high-end telecommunications users in the state regarding a delay and their response to a delay. Triangulation is better with the more corroborating information that can be found, so an analyst could also seek information from vendors, competitors, and consultants.

If, for example, Wall Street said it would be economic suicide for a delay strategy and high-end users said they could and would go elsewhere, it would seem safe to conclude that a delay strategy was posturing and unproductive rhetoric. If Wall Street said utilities can barely scrape up enough capital to replace obsolete equipment and high-end users said no viable supplier options existed, then a commission might have more confidence that the delay was a viable option for the utility, and the commission could react accordingly.

While Wall Street and the high-end users have some interests in common, their core interests and constituencies are different enough to permit the qualitative analyst to have confidence in the independence of their observations and opinions. Note there is no presumption that any one party is entirely objective or possesses the sole and true view. Rather, the thrust of the qualitative assumption analysis is to test and compare the utility's information against information provided by independent experts. The qualitative analyst needs to resolve comparability and proprietary issues normally associated with information produced by independent parties. It is the quality of the comparison that produces the primary analytical power of this method.

Inverted Propositions

A qualitative analysis can yield interesting results if the propositions or assertions or assumptions of a party are stated in an opposite fashion. The intent here is to free the analyst from flowery or persuasive wording by positing that the opposite will occur. Following the "continue business as usual" example, the inverted approach works as follows:

Original Assumption	The utility believes it can continue to do business in an increasingly competitive market without making significant infrastructure upgrades.
Plausible Inverted Assumptions (underlined)	1. The utility <u>does not</u> believe ... 2. The utility believes it <u>cannot</u> continue ... 3. The utility ... ~~without~~ <u>by</u> making significant upgrades ...

The original statement, proposition, or assumption can often be modified in more than one place. By inverting, a better understanding of the original assumption is generally gained because of the need of analysts to clearly state the opposite. Equally important, inverting allows a researcher to begin triangulation or assumption analysis. Further, in a regulatory setting the party making the original statement can be asked if it believes the inverted form to be never, occasionally, or always true. The utility's response will provide the analyst with additional information about the original statement and the range of available options.

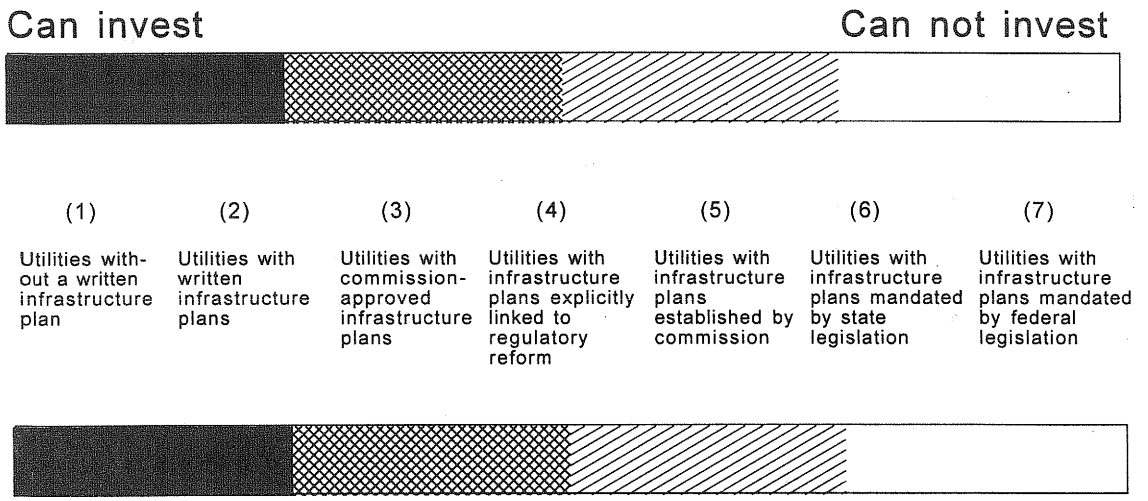
Continuum Analysis

A common feature of many qualitative methods is the interest in testing and examining words for their rhetorical content. The basic thrust is that value-laden and emotional words and concepts can blind an analyst to the full range of possible solution strategies or impacts possible. One way to minimize this problem is by constructing continuums that place the statements, assumptions, or propositions in a larger context. A continuum ideally has a polar type on each end and various gradations in between.

Several things happen when a continuum in constructed. First, positions of parties thought to be polar are often a lot closer in reality. This both increases the chance of consensus and can cause parties to appreciate the reasonableness of their adversaries in the context of potentially even more extreme adversaries or positions. Second, by having to make gradations or categories, the analyst (and the parties) can better articulate the key features separating the parties. This allows a focus on core versus periphery issues.

In the "business as usual" case, the parties may initially feel the illustrated precontinuum in figure 4-1 best reflects reality. A careful analysis of the position of the parties may reveal a more realistic continuum of choices, as shown in the second illustrative continuum.

This continuum suggests that the utility and commission may not be at polar positions, even where in disagreement. It could be the case that gradations three and four most accurately reflect where two parties actually are, but because each sees the other as a polar type they actually perceive that their positions are much further apart. Thus, this process



Source: Author's construct.

Fig. 4-1. Continuum of policy options.

focuses the debate and evaluation on whether an infrastructure investment should or should not be linked to a reform, rather than focusing on whether any infrastructure investment should take place at all.

The qualitative approach allows rough categories and frameworks to be formed that can place policies, statements, assumptions, and propositions in an ordered context. It gives benchmarks to compare items that are not easily susceptible to quantitative measurement. In this example it is more important to know that a written commission-approved plan is the key issue than it is to focus on details separating particular plans. Subcontinuums can be constructed, if needed, to handle finer or more detailed gradations.

The continuum approach allows a better evaluation because it forces the analyst to place items in an ordered context. This first step permits an analyst to have more confidence in subsequent analytical steps because of the firm and ordered foundation established using a continuum. Both quantitative and qualitative analyses can benefit from starting with a continuum approach.

Matrix or Category Analysis

All physical entities and concepts have features or attributes. A organization can be characterized as being large, or small; it can also be thought of as being centralized or decentralized. Also, all processes have inputs, throughput, outputs, and outcomes. A utility may have one or more of these features.

Matrix analysis takes the related attributes or outcomes and displays them in a common format. Consider that in prices caps, a lot of attention has been given to identifying the possible outcomes. Advocates have projected rosy outcomes and opponents have identified disheartening outcomes. The differences expressed have caused various parties to conduct a variety of empirical quantitative studies. Few of these studies have generated widespread consensus.

Using a qualitative perspective, a matrix analysis can be conducted that identifies price cap outcomes in a useful and reliable manner. Quantitative studies tend to make point predictions or set sharp empirical borders. Examples include "forecasting X new jobs under price caps" or "concluding that base rates should be reduced by Y," or that the "productivity rate for a utility is Z." Qualitative studies instead focus on categories and the correct ordering of categories. A qualitative researcher worries that the X, Y, Z point predictions can often be easily dismissed or overturned due to an analytical or data flaw. Quantitative studies are never perfect, and data problems and the choice of and implementation of a particular research method can have flawed choices that invalidate the point prediction. Qualitative matrix studies can be used to overcome X, Y, Z point prediction and other problems because the intent is to set up orderly categories that cover the full range of options rather than to make a particular prediction. The advantage is that an analysis does not rise or fall on the validity of a particular number. The offsetting and primary disadvantage is that no point predictions are produced.

This approach can be illustrated using the following price caps example. Note that a qualitative analysis, like a quantitative analysis, often starts with the analyst having an insight about a key relationship or feature of a problem. A quantitative analyst may hypothesize that A is correlated to B, and the analyst sets out to collect needed data. A qualitative analyst has

insight that, for instance, price caps have a small set of outcomes, and the analyst constructs a matrix of possibilities.

The price caps outcomes matrix in figure 4-2 is organized around the following interrelated ideas:

1. Only a limited number of outcomes are possible under price caps.
2. Two key variables determine price cap outcomes.
    a. Whether or not the initial or base price covers cost.
    b. Whether or not the price index tracks changes in the cost of providing service.
3. The base price either collects revenues greater than, equal to, or less than the utility's cost of providing service.
4. The price cap index—inflation minus an X productivity factor—either collects revenues greater than, equal to, or less than the utility's actual change in costs.
5. By combining the three possible changes in the base price and the three possible changes in the allowed price increase it is possible to construct a matrix of price cap outcomes.

In the matrix in figure 4-3, the issue is less about the accuracy of a particular cost/price analysis submitted by an interested party, and more about the understanding that a commission's price cap policy must be robust enough to cover all nine pricing outcomes. A good price caps policy must be designed to cover revenue surplus and revenue deficit situations and not be predicated upon an assumption that only one of the nine outcomes is possible.

In a matrix analysis the options are organized into rows and columns. It is the task of the qualitative analyst to derive outcomes that accurately correspond to the rows and columns. In Figure 4-2, Case A produces an increase in the surplus available because the initial base price is higher and the index formula is higher than cost. The matrix does not predict costs or cost/price relationships.

FIGURE 4-2

THE RELATIONSHIP OF CHANGES IN COSTS AND PRICES FOR A MARKET BASKET TO CARRIER PROFITABILITY

The Initial Base Price for a Market Basket Is:

The Allowed Price Increase for a Market Basket Is:	Greater than cost	Same as cost	Less than cost
Greater than the cost increase	Carrier surplus increases Case A	Carrier has new surplus Case D	Mixed results* Case G-1 Case G-2 Case G-3
Same as the cost increase	Carrier surplus remains the same Case A	Neutral Case E	Carrier loss remains the same Case H
Less than the cost increase	Mixed results* Case C-1 Case C-2 Case C-3	Carrier has new loss Case F	Carrier loss increases Case I

* See figure 4-3.

Source: Authors' construct.

FIGURE 4-3

EXPLANATION OF THE THREE MIXED OUTCOMES FOUND IN CASES C AND G IN FIGURE 4-2

Case C-1	Carrier surplus occurs because the actual increase in cost is less than the initial cost advantage
Case C-2	Neutral outcome because the actual increase in cost is essentially equal to the initial cost advantage.
Case C-3	Carrier loss as the actual increase in cost is greater than the initial cost advantage.
Case G-1	Carrier surplus as the gain from having a lower cost increase in cost is greater than the initial cost disadvantage.
Case G-2	Neutral outcome because the actual cost increase is essentially equivalent to the initial cost disadvantage.
Case G-3	Carrier loss occurs because the gain from having a lower cost increase is less than the initial cost disadvantage.

Source: Authors' construct.

The result of this type of analysis is that the researcher and policymaker have an explicit ordering of the outcomes likely to occur. In the second part of an analysis, different policy options can be played out against all nine outcomes to see how they work. For example, universal service, or accelerated infrastructure modernization would be less successful in Cases H, F, and I because available revenues would be lessened. The visible, simple, and explicit method allows the reader to independently validate the qualitative researcher's conclusions and decision rules.

The advantage of an attribute or outcome matrix is that many features that are not easy to measure can be deployed in a scaled, ranked, or ordered mode. This allows a great deal of information to be covered by a matrix. The ordered nature minimizes classification problems as the main thought is on the variation from "less-to-more" or from "large-to-small," as

opposed to whether a particular company, area, or policy is large or small. Additionally, this method uses features that already exist and does not necessarily require additional data gathering. The usefulness of the approach is governed by whether the display tells the reader about new features or outcomes that occur and would otherwise not necessarily have been known.

Comparative Analysis

Many qualitative studies depend upon and benefit from the use of quantitative data. Quantitative methods tend to use mathematical or statistical models to produce numbers that can be used for analysis. When qualitative analysis use "numbers" it generally means that data have been obtained from a secondary source. These sources include census data, surveys, governmental data, industrial data, and data from quantitative studies. Thus, a qualitative study may include reference of an X percent inflation rate, or a Y increase in debt, or a Z number of cancer deaths. A qualitative study occasionally includes the original tabulation conducted by the author. A study becomes quantitative when the method used to produce the data is a central feature of the analytical effort of the researcher.

In a regulatory context a qualitative study might include descriptive data about other states served by the holding company, or a listing of profit-sharing levels in price cap states, or an arithmetically derived calculation of the increase in revenues from a previous year. These numbers can serve as important parameters, benchmarks, and indicators in a qualitative study but are not the primary focus.

Conclusion

Qualitative methods can and have been used before regulatory commissions in general and in regulatory reform proceedings. Testimonials and pilot or case studies are perhaps the most frequent qualitative methods employed. The main advantage of qualitative studies is the insight provided and the main disadvantage centers around whether or not the same results can be independently replicated or verified by another researcher.

In a regulatory setting the replication problem is mitigated somewhat because of the ability to cross examine a researcher or witness. This allows the qualitative researcher's methods, assumptions, and goals to be publicly examined and evaluated. A second way the idiosyncratic biases of the researcher are reduced is that the qualitative study is generally one of many competing pieces of information that three or five or seven siting commissioners will evaluate in a particular proceeding. While commissioners do not necessarily validate or officially comment on each quantitative or qualitative study brought before them, each study does play a traceable part in any final order or stipulation. Viewed pragmatically, a qualitative study's methods, assumptions, and conclusions are validated to the extent that they influence the regulators.

Regulators do not have to choose between exclusively relying on quantitative or qualitative studies. Each can be used separately or in support of the other. Qualitative studies are strong on insight and are judged by the quality of the insight. Quantitative studies excel in reliability but can suffer from validity problems. Validity problems are often unrevealed in quantitative studies because of the readers' comfort level with the numerical results.

CHAPTER 5

ECONOMIC PERSPECTIVES ON ALTERNATIVE REGULATION

Introduction: Why Alternative Regulation?

Two key factors, technological innovation and criticism of traditional regulation, led to regulatory reform or alternative regulation of telecommunications firms.[1] Of the two, technological change was the more powerful force because technological change was sufficient to necessitate reform. Though criticism of traditional regulation undoubtedly exerted some force, as long as traditional regulation worked tolerably well, criticism alone was probably not sufficient to bring about radical reform. Although traditional regulation is not without flaws, it did result in deployment of what is arguably the world's finest national telecommunications system. Moreover, traditional regulation also made possible the achievement of universal access to telephone service (national telephone penetration rates were above ninety-four percent).[2]

Technological innovation changed the underlying cost structure of the telecommunications industry, created new market segments, allowed for possible delivery of a wider variety of services through the public switched network, and made competitive entry possible in market segments long protected by franchises and by cost or technical barriers.[3] Because

---

[1] Traditional regulation is called cost-of-service regulation, ratebase regulation, or rate-of-return regulation; it is the "straw-man" model of regulation which has been the subject of considerable criticism.

[2] The national household penetration rate for telephone service in the dwelling unit was estimated to be 94.2 percent as of July 1993. Furthermore, it was estimated that 95.6 percent of households had a telephone available to make and receive calls. See Alexander Belinfante, *Telephone Subscribership in the United States* (Washington, D.C.: Federal Communications Commission, December 1993), 14.

[3] Technological innovations include wireless telephony, fiber optics, and digital switching. Competitive entrants include interexchange carriers, competitive access providers, cellular

of technological innovation, natural monopoly arguments supporting regulation may no longer be valid.[4] Moreover, as the breadth of services delivered through the public switched network increases, accurately determining the cost of and allocating revenue requirements to discrete services becomes considerably more complex.

In addition, the telephone companies, faced with the inexorable force of competitive entry, have begun to offer advanced infrastructure deployment and educational packages as a *quid pro quo* in exchange for obtaining alternative regulation. The impact of competitive entry into formerly closed segments of telecommunications markets and the desire to promote economic development have led to a reexamination of the goals, methods, and outcomes of traditional public utility regulation, especially as applied to telecommunications.[5] Concomitantly, information infrastructure investment and advanced service development have become important components of state and national economic development policy.[6] If rapid

---

services, and cable television systems. New services (either offered at present or contemplated) include ISDN service, Class features (including caller ID, voice mail, call waiting, and call forwarding), video dialtone, and information services.

[4] For some discussion of the impact of technological innovation on competitive entry, see Phyllis Bernt, Hans Kruse, and David Landsbergen, *The Impact of Alternative Technologies on Universal Service and Competition in the Local Loop*, (Columbus, Ohio: The National Regulatory Research Institute, October 1992).

[5] Another factor pressing for regulatory reform is a desire to improve the quality of life by deploying the capacity to provide facilities for telemedicine and distance learning, both of which are applications of advanced technology and service capability. Several alternative regulation plans contain commitments from telephone companies to expedite deployment of the network facilities (digital switching, ISDN, Signalling System Seven, and fiber optics) necessary to provide these services.

[6] See Robert G. Harris, "Telecommunications Services as a Strategic Industry: Implications for United States Public Policy," in Michael A. Crew, editor, *Competition and the Regulation of Utilities* (Boston, Massachusetts: Kluwer Academic Publishers, 1991), 97-119; Francis J. Cronin, Paul Hebert, and Elisabeth Colleran, "Linking Telecommunications and Economic Competitiveness," *Telephony* (September 7, 1992): 38-42, Patricia M. McGovern and Paul Hebert, "Telecommunications and Economic Development," *Telephony* (November 2, 1992): 26-31; and Francis J. Cronin, Elisabeth K. Colleran, Paul L. Herbert and Steven Lewitzky, "Telecommunications and Growth: The Contribution of Telecommunications Infrastructure Investment to Aggregate and Sectoral Productivity,"

deployment of advanced infrastructure and services is desired and if competitive entry is to be encouraged, new regulatory regimes may be required.[7]

Goals of Regulation

Utility regulation (at least in the public interest view) is designed to serve as a surrogate for the discipline of competition in markets that are not likely to produce, on their own, a workably competitive outcome or for which there is a public interest in the general availability of reliable service on reasonable terms.[8] Around this general notion of regulation as a substitute for competition, a combination of social and economic goals has evolved.

Telecommunications Policy 17, no. 9 (December 1993): 677-90.

[7] Although a state's legislators and regulators are justified in not wanting their state to fall behind its neighbors in the race to reach the information highway, caution should be exercised when considering suggestions that their state will be ahead of the technology curve if and only if regulation is reformed according to the telcos' wishes. It is simply not possible for every state to simultaneously be "first" in the race. Furthermore, telcos will deploy advanced technology when it makes business sense to do so; deployment before such technology is needed will lead to excess capacity and carrying charges for the company and possibly upward pressure on rates.

[8] The notion that some markets, if left alone, will not approximate a competitive market include conditions of "natural monopoly" and markets which could lead to destructive competition and potentially unstable supply. For a thorough discussion of these concepts see William W. Sharkey, *The Theory of Natural Monopoly* (New York: Cambridge University Press, 1982). Other markets can fall under regulation not because they are natural monopolies but because either the competitive market produces a result that is not acceptable (perhaps because of problems with perceived distributional equity, information asymmetries, externalities, or public good characteristics) or the competitive industry competes closely with a regulated industry (trucking and railroads, for example).

A listing of the widely accepted goals of regulation would include the following (somewhat overlapping) goals:[9]

(1) ensuring utilities the opportunity to be financially viable,

(2) preventing utilities from earning monopoly profits,

(3) controlling ruinous or destructive competition,

(4) controlling undue price discrimination, especially when discrimination works against the interests of the most vulnerable customers,

(5) preventing cross-subsidization, especially of competitive services by monopoly services,

(6) preventing the exploitation of monopoly power,

(7) promoting universal service,

(8) ensuring adequate reliability and quality of service,

(9) encouraging efficient production and consumption patterns,

(10) promoting innovation,

(11) encouraging the transition to competition where appropriate, and

(12) promoting economic development.

In addition, another goal of regulation is that it be administratively efficient and feasible. Regardless of the method of regulation, its administrative, reporting, compliance, and monitoring costs must be relatively low and implementation must be feasible. Moreover, the process of regulation should be open and understandable, regulatory decisions should be reasonably predictable, and rates should be sufficiently stable to allow for long-range planning by producers and consumers. These last goals may lead to some reluctance to make major

[9] Goals are listed in no particular order.

changes in the level or structure of rates. These goals are also related to the goals of equity or fairness, especially as the regulatory process itself must be perceived as being fair.

The Problems with Traditional Regulation[10]

For more than thirty years, economists, especially those associated with the "Chicago School," have been critical of the rationale, structure, method, and results of traditional regulation. Criticisms of traditional regulation have been numerous and varied. When faced with a "technical" (natural) monopoly, Friedman argued that to choose among private monopoly, public monopoly, and public regulation was to choose among evils. Furthermore, he stated that, "if tolerable, private monopoly may be the least of the evils."[11] Stigler and Friedland conducted an early empirical investigation of the effects of regulation on prices, and found no significant effect.[12] Independently, Stigler, Posner, and Peltzman presented discussions that were critical of regulation (or at least the likelihood that it will promote the public interest rather than the private interests of regulators and/or the regulated firms).[13] Jarrell presented an empirical test of Peltzman's theory that public utility regulation was sought by and primarily benefitted the industry.[14] Demsetz criticized the rationale for

---

[10] Readers familiar with the academic literature on traditional regulation may wish to skim this section and go directly to page 128.

[11] Milton Friedman, *Capitalism and Freedom* (Chicago: University of Chicago Press, 1962), 28.

[12] George J. Stigler and Claire Friedland, "What Can Regulators Regulate? The Case of Electricity," *Journal of Law and Economics* 6, no. 1 (1962): 1-16.

[13] George J. Stigler, "The Theory of Economic Regulation," *Bell Journal of Economics and Management Science* 2, no. 1 (Spring 1971): 3-21; Richard A. Posner, "Theories of Economic Regulation," *Bell Journal of Economics and Management Science* 5 no. 2 (Fall 1976): 335-357; and Sam Peltzman, "Toward a More General Theory of Regulation," *Journal of Law and Economics* 19, no. 2 (October 1976): 211-240.

[14] Gregg A. Jarrell, "The Demand for State Regulation of the Electric Utility Industry," *Journal of Law and Economics* 21, no. 2 (October 1978): 269-295.

regulation by commissions and proposed the use of franchise bidding instead.[15] In addition, Averch and Johnson used neoclassical economic modeling to conclude that traditional regulation leads to overcapitalization and noncost-minimizing behavior.[16]

Inefficient Pricing and Incentives

There is little doubt that, as applied, traditional regulation is an imperfect surrogate for competition. There are several reasons for this. First, regulation has goals (universal service and equity are examples) that are not directly considered in competitive markets.[17] Although it has probably done an adequate job in reference to the first seven goals, traditional regulation, with its focus on ratebase, rate-of-return, revenue requirements, and administrative

---

[15] Harold Demsetz, "Why Regulate Utilities?" *Journal of Law and Economics* 11, no. 1 (1968): 55-65.

[16] Harvey Averch and Leland Johnson "Behavior of the Firm Under Regulatory Constraint," *American Economic Review* 52, no. 5 (December 1962): 1052-1069. Although this tendency has been associated with Averch and Johnson, the same conclusion was reached in Stanislaw Wellisz, "Regulation of Natural Gas Pipeline Companies: An Economic Analysis," *Journal of Political Economy* 71, no. 1 (February 1962): 30-43.

[17] Promoting universal service and equity may lead regulators to adopt policies that result in outcomes that economists would categorize as inefficient. Thus, if efficiency is the standard against which outcomes are measured, regulatory outcomes are almost certain to be judged harshly. For an accessible treatment of some of the issues surrounding the trade off between efficiency and equity see Edward E. Zajac, *Fairness or Efficiency: An Introduction to Public Utility Pricing* (Cambridge, Massachusetts: Ballinger, 1978). Zajac observes (at age twenty-one) that:

> . . . the regulated sector . . . is the mirror image of the market sector. The principal economic-theoretic concerns in the market sector . . . are of tangential concern in . . . ratebase, rate-of-return regulation . . . On the other hand, matters of fairness and justness—of relatively little theoretical concern in a pure market economy—loom large, either explicitly or implicitly, and are of primary importance in regulatory proceedings.

The trade off between efficiency and equity is also discussed in James C. Bonbright, Albert L. Danielson, and David R. Kamerschen, *Principles of Public Utility Rates* (Arlington, Virginia: Public Utilities Reports, Inc., 1988), 179-192.

due process has long been viewed as ineffective, slow, cumbersome, inefficient, and sometimes perverse.

Traditional regulation tends to rely on historic, average accounting cost pricing rather than more efficient marginal cost-based pricing.[18] Traditional regulation also tends to rely on aggregation and averaging, which sets uniform rates for large groups of customers, even though the costs of serving them may vary considerably. Moreover, traditional regulation tends to use implicit redistribution powers to create subsidies for certain favored customer categories or services. In telecommunications, the most notable of these subsidies are those that are believed to flow to residential access. It is widely thought that business customers subsidize residential customers, toll usage subsidizes local access, and urban access subsidizes rural access, all in the name of promoting universal service. These subsidies are not likely to be sustainable in the face of competitive entry, and they possibly may have created pricing anomalies that led to uneconomic bypass of the public switched network.[19]

Second, traditional regulation does not generally create an incentive structure that effectively mimics competitive markets. Critics have argued that in addition to leading to excess investment and, concomitantly, excessive costs, the bureaucratic and "cost-plus" nature of traditional regulation leads to reduced incentives for innovation and cost cutting. Indeed, the cost-plus nature of traditional regulation and lack of competitive pressure can lead to inefficient behavior. This form of inefficiency may result in gold plating (designing the

---

[18] The argument is sometimes made that historic accounting cost pricing is more accurate and less confusing that marginal cost pricing. This may be because marginal cost pricing is based on estimates of future costs and because economists disagree about the correct measure of marginal cost. This may mean that those in favor of using historic accounting cost would rather be being precisely wrong than approximately right. Another argument favoring historic accounting cost-based prices is that they are somewhat easier to link to a revenue requirement than are marginal cost-based prices. Nevertheless, if average costs and marginal costs differ, historic accounting cost prices are not capable of transmitting correct signals to either consumers or producers. Thus, such prices tend to create distortions.

[19] Uneconomic bypass results when an entrant's costs are above those of the regulated firm but it is able to offer service at rates below the regulated firm's rates. This could happen if the regulated firm's rates include contributions for various subsidy mechanisms that do not similarly burden entrants.

system for more durability or reliability than is optimal), managerial slack (a lack of incentives to aggressively cut costs), or allowing managers the freedom to pursue personal objectives other than profit-maximizing or cost-minimizing behavior.[20]

Another problem with traditional regulation is the lack of incentives for risk taking. Although a utility may keep the rewards of successful risk taking during the period between rate cases, such rewards will most likely be redistributed to ratepayers at the next rate case. It is questionable whether such short-lived interim benefits provide strong incentives. Furthermore, unsuccessful risk-taking may produce penalties as a result of cost disallowances or prudence tests. Thus, traditional regulation lacks either the competitive pressures or the rewards for innovation that exist in competitive markets.

---

[20] Economists distinguish between two forms of efficiency: technical and allocative. Technical efficiency is the narrower concept: a firm is technically efficient if it produces the maximum output given the inputs used (no waste). Allocative efficiency means producing output at the least total cost. This requires choosing the correct mix of inputs over time and may involve altering the current mix of inputs. Moreover, allocative efficiency implies technical efficiency but the converse need not be true. However, a firm that is technically inefficient cannot be allocatively efficient. One form of technical inefficiency was discussed in Harvey Leibenstein, "Allocative Efficiency vs. X-Efficiency," *American Economic Review* 56, no. 3 (June 1966): 392-415. Another form of inefficiency is discussed in theories of expense preference behavior, which suggest that managers might favor certain types of expenses over others (fringe benefits and perquisites, for example). Expense preference theory was developed in Oliver E. Williamson, *The Economics of Discretionary Behavior: Managerial Objectives in a Theory of the Firm* (Englewood Cliffs, New Jersey: Prentice-Hall, 1964). Williamson's theory was applied to public utilities in Michael A. Crew and Paul R. Kleindorfer, "Managerial Discretion and Public Utility Regulation," *Southern Economic Journal* 45 (1979): 696-709. Reliability issues are considered in Michael A. Crew and Paul R. Kleindorfer, "Public Utility Regulation and Reliability with Applications to Electric Utilities," in Michael A. Crew, editor, *Issues in Public Utility Pricing and Regulation* (Lexington, Massachusetts: D.C. Heath, 1980), 51-75; and Michael A. Crew and Paul R. Kleindorfer, "Some Elementary Considerations of Reliability and Regulation," in Michael A. Crew, editor, *Problems in Public Utility Economics and Regulation* (Lexington, Massachusetts: D.C. Heath, 1980), 143-65.

Information Asymmetry

Additionally, because regulators generally depend on the utility for most information about costs, demand conditions, etc., a principal-agent or information asymmetry problem can arise. Stated simply, regulators are unlikely to be able to obtain or process sufficient information to make optimal judgments regarding the efficiency of the utility's operations and managerial effort. Also, regulated firms may have incentives to use strategic behavior, for example, controlling the flow of information to mislead regulators.[21] Traditional regulation may be second-best when applied to technologically stable natural monopolies. However, under conditions of information asymmetry, traditional regulation may be "last-best" when applied to multiproduct firms engaged in technologically evolving and potentially competitive markets.[22]

---

[21] The information asymmetry problem is not unique to public utility regulation; it can occur in many settings. One cause is information overload: regulators must simultaneously deal with multiple utilities that operate in several industries. Another cause is strategic behavior by the utilities: regulation tends to rely on adversarial proceedings in which each party has strong incentives to present its case as convincingly as possible. This may lead to some "gaming" of the system.

Numerous authors have made proposals for dealing with this problem. These include: Ingo Vogelsang and Jörg Finsinger, "A Regulatory Adjustment Process for Optimal Pricing by Multiproduct Firms," *The Bell Journal of Economics* 10, no. 1 (Spring 1979): 157-171; Martin Loeb and Wesley A. Magat, "A Decentralized Method for Utility Regulation," *Journal of Law and Economics* 22, no. 2 (October 1979): 399-404; David P. Baron and Raymond R. DeBondt, "On the Design of the Regulatory Price Adjustment Mechanism," *Journal of Economic Theory* 24, no. 1 (February 1981): 74-90; David P. Baron and Roger B. Myerson, "Regulating a Monopolist with Unknown Costs," *Econometrica* 50 no. 3 (September 1982): 911-930; Jean-Jacques Laffont and Jean Tirole, "Using Cost Observations to Regulate Firms," *Journal of Political Economy* 94 (1986): 614-41; Tracy R. Lewis and David E. M. Sappington, "Regulating a Monopolist with Unknown Demand and Cost Functions," *RAND Journal of Economics* 19, no. 3 (Autumn 1988): 438-57; and Tracy R. Lewis and David E. M. Sappington, "Inflexible Rules and Incentive Problems," *American Economic Review* 79, no. 1 (March 1989): 69-84.

[22] One reason for this is that the traditional regulatory process of determining cost allocations, setting revenue requirements, and developing tariffs is simply too cumbersome for an environment characterized by an ever expanding array of jointly produced services and constantly evolving technology.

It must be noted that public utility regulation has not been the only form of regulation under attack: a broad movement developed aimed at deregulating or reforming regulation of other sectors of the economy, such as transportation (airlines, trucking, and railroads) and financial services (commercial and investment banking, savings and loan institutions, brokerage services). A discussion of the general movement towards deregulation or regulatory reform and some empirical estimates of the results of deregulation in several sectors may be found in Winston.[23]

For these reasons, various alternate regulatory methods are now being proposed, implemented, and evaluated. Alternative regulatory methods include incentive regulation, price-cap regulation, social-contract regulation, full or partial deregulation (which may be service-by-service deregulation or allowing limited price flexibility), and opening markets to competitive entry.[24] Although the specific goals of the various alternative regulatory methods differ somewhat, they are generally aimed at eliminating or reducing some of the problems thought to be associated with traditional regulation and/or promoting additional social policy or economic development goals.[25]

---

[23] Clifford Winston, "Economic Deregulation: Days of Reckoning for Microeconomists," *Journal of Economic Literature* 31, no. 3 (September 1993): 1263-1289. Winston finds that deregulation has had largely beneficial results. He also suggests (at 1286, note 48) that:

> In practice, the choice is between some form of imperfect competition or imperfect regulation. Given this choice, the accumulated evidence . . . suggests the burden of proof should be on those who argue price and entry competition is not workable.

[24] As implemented, various forms of alternative regulation exhibit considerable overlap, and many are hybrids or combinations of several regimes. Social contract regulation may include a combination of price freezes, price caps, and partial deregulation. Allowing competitive entry can be considered to be a form of alternative regulation that alters the traditional implicit social contract under which utilities were granted exclusive franchises. Price-cap regulation for some services is often combined with partial deregulation and incentive regulation in the form of a sliding-scale through which the utility shares its profits (above a certain rate of return) with its customers. Also, technology and infrastructure deployment commitments are sometimes part of alternative regulation plans.

[25] Most of the goals of regulation are not genuine goals, in and of themselves. The goal of regulation is (or should be) to promote overall social welfare. Regulators cannot do this

Critiques of the Critics

Although numerous analysts have been fervently critical of public utility regulation, others are not totally critical of regulation. A rebuttal to some of the Chicago-School critiques may be found in Trebing.[26] The assumptions of the Chicago School are questioned by Evans, who views the belief in antiregulation rhetoric as being almost an act faith for Chicago-School critics.[27] The historical validity of Peltzman's analysis is questioned by Priest.[28] Phillips sees pluses and minuses in both regulation and market models.[29] The empirical evidence of the Averch-Johnson overcapitalization bias is mixed, especially when the effect of demand uncertainty and/or the effect of risk on capital market values is considered.[30] Also, the guaranteed rate of return on ratebase envisioned in the Averch-Johnson analysis is not reflected in real-world implementation of rate-of-return regulation. In addition, regulators' traditional concern for the health of incumbent firms can be traced to their concern for the dependable and reliable utility services, and this concern

---

directly. They can, however, choose a set of intermediate goals or indicators that they believe are reasonably consistent with social welfare goals. Among the broader goals of regulation are economic efficiency, economic growth and development, and equitable or fair pricing and distribution policies.

[26] Harry M. Trebing, "The Chicago School Versus Public Utility Regulation," *Journal of Economic Issues* 10, no. 1 (1976): 96-126.

[27] R. G. Evans, "Slouching Toward Chicago: Regulatory Reform as Revealed Religion," presented at the Symposium on Regulation," University of Toronto Law School, Millcroft, Ontario, October 28-30, 1981.

[28] George L. Priest, "The Origins of Utility Regulation and the 'Theories of Regulation' Debate," *Journal of Law and Economics* 36 (April 1993): 289-323.

[29] Almarin Phillips, "Regulation and Its Alternatives," in Chris Argyris, et al., *Regulating Business: The Search for an Optimum* (San Francisco: Institute for Contemporary Studies, 1978), 152-172.

[30] See Edwin A. Rosenberg, *Input Choice Under Uncertainty: An Application of the Capital Asset Pricing Model to the Electric Utility Industry* (Unpublished Ph. D. dissertation, North Carolina State University, 1984).

leads to ensuring the viability of incumbents and a reluctance to put faith in unknown and possibly unreliable entrants. Also, as stated by Waterson, the critics of regulation cannot have it both ways:

> Assuming that regulation has the potential to promote social benefits, either it is effective, but in the wrong directions, or it is ineffective: it can scarcely be both. We also have the paradox that . . . those same people who tend to argue that the common law promotes efficiency also reject the idea that regulatory law does the same thing.[31]

Recently, Shepherd considered the criticisms of traditional regulation and made some comparisons with alternative forms of social control. He concluded that when compared with real-world implementations of other plans, traditional regulation might not result in large inefficiencies. Moreover, immediate and total decontrol may not be advisable: rather than leading to effective competition, premature deregulation may lead to entrenched dominance if underlying market conditions allow an incumbent firm to maintain considerable market power.[32] It has also been argued that, some of the details of traditional telecommunications regulation (certificates of public convenience and necessity, detailed tariffs for each and every service, and some cost-of-service study requirements) may no longer be necessary. However, if the primary role of regulation is to actively protect consumers from the excesses of utility behavior, regulators should still be watchful for cross-subsidies, cost, profit, and risk shifting to or from monopoly markets and other potential abuses.[33] As an institution of social policy, traditional regulation most likely serves various public purposes, although alternative institutions might serve the same purposes more efficiently.

---

[31] Michael Waterson, *Regulation of the Firm and Natural Monopoly* (New York: Basil Blackwell, 1988), 8. Citation in original omitted.

[32] William G. Shepherd, *Regulation and Efficiency: A Reappraisal of Research and Policies* (Columbus, Ohio: The National Regulatory Research Institute, July 1992).

[33] Douglas N. Jones, "Has Traditional Regulation Outlived Its Role in Telecommunications?," presented at New England Conference of Public Utility Commissioners Symposium, The Balsams, Dixville Notch, New Hampshire, June 29, 1993.

Analysis of the Effect of Alternative Regulation Plans

Analysis of the effects of alternative regulation may be difficult. Changes in regulatory regimes or policies affect individuals and firms, whose actions, in turn, impact the effect of alternative regulation. In addition, alternative regulation is an area in which implementation may be ahead of theory; and implementation is certainly ahead of empirical analysis.

Nevertheless, evaluative analysis is both necessary and possible. Different jurisdictions will adopt different degrees and types of alternative regulation at different times and for different companies. Evidence will accumulate and lead to some conclusions as to the effect of various plans. For example, Mathios and Rogers performed an empirical study of the effect of price-cap regulation on AT&T's intrastate, interLATA prices in multi-LATA states. They found that, compared with states that imposed traditional regulation, average prices were lower in states that allowed price flexibility in conjunction with a cap on average prices.[34]

As compared with traditional regulation, price-cap regulation has potential benefits, including providing incentives for the firm to achieve cost savings, eliminating the Averch-Johnson bias towards capital inputs; restricting the firm's ability to finance predatory moves in competitive markets by raising monopoly prices; reducing the administrative cost of regulation, and allowing greater pricing flexibility for individual services, which encourages the firm to move toward socially optimal Ramsey prices. However, price-cap regulation also creates potential problems or difficulties including (1) determining the correct price index and productivity adjustment upon which to base changes in the cap; (2) allowing greater pricing flexibility, which may lead to predatory pricing if monopoly and competitive services are included in the same basket of services, (3) creating incentives for reduced service quality or system reliability, and (4) allowing the firm to move toward Ramsey pricing by exploiting differences in demand elasticities between services or customers. Because the link between

---

[34] Alan D. Mathios and Robert P. Rogers, "The Impact of Alternative Forms of State Regulation of AT&T on Direct Dial Long Distance Telephone Rates," *RAND Journal of Economics* 20, no. 3 (Autumn 1989): 437-53.

investment spending and revenue requirement is broken, price-cap regulation may also reduce incentives for LECs to upgrade service in some smaller, rural exchanges.

Note that one of the advantages of price-cap regulation (the move toward Ramsey pricing) is also considered to be a disadvantage. This is because Ramsey pricing, though optimal from a total welfare or efficiency standard, may violate implicit norms of equity or fairness by justifying relatively larger price increases for low-elasticity (monopoly) customers than for high-elasticity (competitive) customers. Another of the advantages (limiting the firm's ability to adopt predatory pricing) can become a disadvantage depending on how the plan is designed.

Ex Ante and *Ex Post* Analysis

Evaluating the effectiveness of various regulatory methods requires that the methods be subjected to both *ex ante* and *ex post* analysis. *Ex ante* analysis is needed to examine the incentives created under the alternative regulation plan. Such incentives are given to various groups, such as the LEC, regulators, consumers, and competitors. Moreover, the evaluation should attempt to ensure reasonable incentive compatibility.[35] *Ex ante* analysis can be used to determine the extent to which the goals of the alternative regulation plan are likely to be achieved. *Ex post* analysis is needed because outcomes must be compared with goals, and corrective actions must be designed, if needed. Because various alternative regulation plans are already in operation, and some are nearing their *ex post* evaluation and analysis phase, the results of *ex post* analysis of existing alternative regulation plans can provide important information about what works and what does not work. Information on the achieved results under existing alternative regulation plans should be an important part of the information considered in the *ex ante* analysis of proposed alternative plans.

---

[35] A regulatory plan is incentive compatible if it leads the regulated firm to act in a manner that leads to the outcome desired by regulators. This means that the firm's natural profit maximizing behavior is harnessed to create some socially desired result.

Ex Ante Criteria for a Desirable Alternative Regulation Plan

The normal assumption is that a utility's primary motivation is to earn and retain profits. Therefore, alternative regulation plans tend to allow the utility to earn and retain a higher level of profits, if certain goals are met. Equivalently, alternative regulation may rely on the invisible hand of self-interested behavior to induce the utility to increase its profits, thereby achieving certain goals along the way. Regulators have definite goals in mind when they institute an alternative regulation plan. Those goals and the incentives given the utility should stand in a clear and logical relationship to one another. A good alternative regulation plan harnesses the utility's self interest to achieve the goals regulators have in mind. When considering the merits of an alternative regulation plan, regulators might consider several criteria in making an *ex ante* analysis. Among these are the following:[36]

- If incentives are used, they should address those aspects of a utility's performance that are under the control of its management.

A utility should be neither rewarded nor penalized for the effects of actions that are beyond its control. However, a utility may be rewarded for reacting promptly and/or correctly to changes in its environment or opportunities. Similarly, the utility may be penalized for reacting slowly and/or incorrectly to changes in its environment or opportunities.

- The utility should be given a clear expectation as to how its performance under the new form of regulation will be evaluated and how rewards or penalties will be conferred.

---

[36] These criteria are adapted from the discussion in William P. Pollard *et al.*, *Rate Incentive Provisions: A Framework for Analysis and a Survey of Activities* (Columbus, Ohio: The National Regulatory Research Institute, November 1981), 55-66.

Although this may lead to some gaming of the plan, the utility needs to be told what measures will be used to assess its performance and how performance measures are linked to the reward structure.[37]

- Application of the alternative regulation plan should result in a positive net benefit to the utility's consumers and society as a whole.

The net benefits, in terms of cost savings and/or increased consumer surplus should be larger than the benefits that flow to the utility. Although alternative regulation that benefits the utility and does not harm consumers could still provide positive net benefits, regulators should attempt to ensure that also consumers reap some visible benefit. In addition, because the adoption of an alternative regulation plan may result in some consumers being worse off than before, it must be clear that the winners' gains more than offset the losers' losses.

- The information necessary to evaluate the desired behavior should be free from tampering and ambiguity.

Generally, regulators must rely on the utility for information upon which to assess the outcome of the alternative regulation plan. Therefore, some possibility exists that the utility will find it advantageous to manipulate the reporting and/or interpretation of information to be used in evaluating its performance. Although it is not always easily accomplished, care must be taken to ensure that the data used in the evaluation is complete, objective, easily verifiable, and understandable.

- The goal and method of application should have neutral effects and have few unintended adverse consequences.

---

[37] For some further discussion of this see the following section, Planning for *Ex Post* Analysis.

Regulators should be aware that basing performance evaluations on one criteria may promote behavior that is detrimental to another goal. Profit incentives based on cost control may lead to less maintenance and/or a degradation of service quality. Plans that call for rapid deployment of advanced technology or services by telephone companies may result in excessive costs if sufficient demand does not materialize or if there is a significant technological advance. In addition, such deployment may give telephone companies a first-mover advantage that may prove difficult for competitors to overcome. Similarly, incentives based on rapid responses to installation and service orders may lead to cost increases, because additional installers and service technicians may be required. Therefore, performance measures should be multi-dimensional to limit unintended adverse consequences.[38]

- The alternative regulation plan should address and eliminate disincentives that exist in traditional regulatory practices.

Alternative regulation is intended to encourage efficient behaviors or promote social goals to a greater extent than is possible under traditional regulation. Therefore, a good alternative regulation plan induces behavior that better comports with these goals than does traditional regulation. Regulators should consider the means by which and the extent to which the alternative regulation plan promotes social goals and addresses the deficiencies of traditional regulation.

---

[38] For discussion of ways in which performance incentives might lead to adverse, unintended consequences, see Robert J. Graniere, Daniel J. Duann, and Youssef Hegazy *The Effects of Fuel-Related Incentives on the Costs of Electric Utilities* (Columbus, Ohio: The National Regulatory Research Institute, November, 1993). Graniere, Duann, and Hegazy actually found fortunate or positive unintended consequences in the electric utility case. However, the possibility of adverse unintended consequences must not be ignored and regulators would be advised to prepare for them.

Ex Ante Analysis

In an ideal situation, regulators explicitly identify their goals and how effective different methods or programs would be in achieving these goals. One such *ex ante* analytical approach, for example, would require regulators to undertake the difficult task of determining goals they wish to pursue and the relative weights or importance attached to the goals.[39] When goals have been assigned weights, each proposed solution or method can be scored in terms of its ability to achieve the goals, and an overall performance or merit score can be constructed for the various methods. The merit scores of alternative regulation methods can be compared with each other and with the merit index of traditional regulation, which could be used as a norm (with its score set to equal 1.00) since it is the existing standard.

Another, more feasible approach requires only that the impact of various alternatives on regulatory goals be described or measured relative to traditional regulation.[40] Regulators would choose the alternative that provides the preferred mix of outcomes. Furthermore, analysts might be able to deduce regulators' preferences based on the alternatives chosen.[41] One possible difficulty with this approach is that, although preferences may map uniquely to alternatives, the mapping from alternatives to preferences may not be unique. In other words,

---

[39] This approach may not be easily implementable because regulators may be neither willing nor able to list goals in a consistent rank order or other order from which weights can be developed. Nevertheless, analysis can be based on the assumption that regulators, in principle, have consistent and stable preferences. A hypothetical illustration of this approach is found in Raymond L. Lawton, "Regulatory Reform: What Is It and How to Tell What Works," presented at the Twenty-Fourth Annual Conference of the Institute of Public Utilities: *Shifting Boundaries Between Regulation and Competition in Telecommunications and Energy*, Colonial Williamsburg, Virginia, December 16, 1992.

[40] When there are multiple objectives, some of which may conflict, decisionmaking will not usually be a simple task. The problem might be made more manageable through the use of scorecards.

[41] Another possibility would be to develop a "hedonic demand model" or "characteristics demand model" for alternative regulation. Hedonic demand models are based on the notion that most goods (here alternative regulation plans) have multiple attributes, and consumers (here regulators) choose the one with the best mix of outcomes.

there may be more than one set of preferences that lead to the same regulatory regime being chosen.

It is not likely that regulators will be able (or willing) to state their goals or preferences in a form that is amenable to conducting an academic researcher's vision of an ideal *ex ante* evaluation and analysis. On the other hand, regulators should be willing and able to state the broad intent of alternative regulation. This broad intent would likely include such items as enhanced deployment of advanced technology and enhanced services (including distance learning, telemedicine, and information highway onramps), maintaining or improving QOS, keeping local access prices as low as possible (possibly by freezing them for several years), opening local and intraLATA markets to competition, and easing the LECs' and the commissions' administrative burden of regulation. Likewise, telephone companies proposing alternative regulation should be willing and able to explain how customers will benefit relative to traditional regulation.

One method of comparing various plans includes the following steps:

1. On each of the goals, determine whether the impact of each form of alternative regulation is positive, negative, or neutral relative to traditional rate-of-return regulation.[42]
2. Using regulators' preferences and weights to be applied to different goals, choose the form of regulation that has the most positive net impact. This may be difficult because many impacts will be qualitative rather than quantitative. Moreover, regulators may find it difficult to determine or compare the magnitude of the impacts.

This approach can help avoid and discount the tendency of reform advocates to unfairly compare an ideal form of their favored reform to traditional regulation by selecting only certain goals. If all proposed reforms are compared against the same set of goals, then a more useful and balanced evaluation analysis can result.

---

[42] A comparison table that could be used for this purpose was presented in Stephen C. Littlechild, *Regulation of British Telecommunications' Profitability* (London: Department of Industry, 1983), 37. This table is reproduced in Michael Waterson, "A Comparative Analysis of Methods for Regulating Public Utilities," *Metroeconomica* 43, no. 1-2 (1992): 205-226.

Rate-of-Return Considerations Under Alternative Regulation

Regulators must balance several opposing concerns when considering alternative regulation plans. Almost all alternative regulation plans allow the firm to earn higher profits, provided it operates more efficiently, meets social goals, or takes advantages of market opportunities. Although regulators will not be as concerned with rates of return under alternative regulation, as they have been under traditional regulation, several reasons exist for regulators to consider achieved rates of return.[43]

Why Care About Rate of Return Under Alternative Regulation?

Even under alternative regulation, it may be legitimate and necessary for regulators to monitor rates of return and take corrective action if they become unacceptably high. Rates of return are one indicator of whether monopoly profits are being earned. Rates of return must be monitored unless markets become sufficiently competitive or contestable to justify elimination of the entire concept of economic regulation, which has as a central purpose the control of monopoly power. High rates of return that result from highly efficient and entrepreneurial activity in competitive markets, or from exploitation of some unique resource, may not be objectionable. High rates of return in competitive or contestable markets will attract entry, putting downward pressure on prices and on achieved rates of return. High rates of return that are sustained over time and result from barriers to entry, which may include legal restrictions, or exploitation of monopoly customers, are not benign.

In addition, the ability of regulators to make a long-run commitment to alternative regulation and hold the firm to the bargain may have political or legal limits. If alternative regulation creates an opportunity for unconscionably high rates of return, political considerations may lead to actions to limit them. At the other extreme, regulators may be

---

[43] Much of the following discussion assumes some form of price-cap regulation, but similar results may be obtained under other forms of alternative regulation where rates of return are allowed to vary.

required to act under the *Hope* test if the utility becomes unprofitable.[44] For this reason, it is possible that alternative regulation adopted at the legislative level is preferable to alternative regulation adopted at the agency level. This is because it may have more long-term credibility.

How Is Rate of Return Controlled Under Alternative Regulation?

The inability to totally disregard the earned return of a utility operating under alternative regulation plans may be addressed by the inclusion of profit-sharing arrangements in many alternative regulation plans. Such plans put both upper and lower bounds on the utility's earnings. Another mechanism regulators have adopted for dealing with this situation is to adjust the productivity offset in price-cap regulation. This can be done implicitly (by setting a relatively high productivity target), explicitly (by including a "consumer dividend,") or by using both methods. Yet another mechanism available to control the utility's rate of return is to require the company to make infrastructure upgrades earlier than market conditions would otherwise dictate.

These mechanisms create a difficult situation because, though possibly necessary, they tend to transform alternative regulation into back-door rate-of-return regulation. For example, one of the putative benefits of price-cap regulation is that, because price ceilings are based on an exogenous index, the link between the firm's reported costs and its prices is cut. This is intended to reduce the problems of cost-plus regulation and information asymmetry. By doing so, price caps are thought to promote efficiency and allow diligent firms to prosper. However, if a high achieved rate of return causes regulators to raise the productivity factor,

---

[44] See *Federal Power Commission versus Hope Natural Gas*, 320 U.S. 591 (1944). In that decision (at 603) the Court held that the return to the equity owner "should be sufficient to assure confidence in the financial integrity of the enterprise, so as to maintain its credit and attract capital." The extent to which this standard imposes upon regulators the obligation to throw a life preserver to a foundering utility operating under alternative regulation has not been settled.

the price cap becomes partially endogenous. This is so because superior performance can result in a higher productivity offset and can lower future profits.

Moreover, if price-cap regulation is combined with profit sharing, regulators will still be required to analyze costs to determine the profit level.[45] This leads to bargaining between the firm and the regulators as to the productivity factor and as to what costs are allowed when determining the level of shared profits. The potential for regulators to adjust price caps to achieve back-door rate-of-return regulation by "recontracting" the alternative regulation plan was recently considered by Sappington and Weisman, who are concerned with the effects of incentives to recontract that may affect regulator behavior. Recontracting generally makes it difficult for the utility to meet the raised standards and reduces the incentives for the types of behavior regulators are trying to promote.

Among the potential problems described by Sappington and Weisman are the following: (1) regulators may be tempted to disallow some costs so that the utility's measured earnings are higher, which creates more excess profits to share; (2) regulators may turn price-cap regulation into rate-of-return regulation by raising the productivity offset to limit profits; and (3) if basic residential rates are frozen or controlled by a price cap, regulators may be more willing to allow competitive entry. This puts downward pressure on rates for the utility's more competitive services without putting upward pressure on rates for monopoly services.[46]

One desired outcome of alternative regulation is that it increases the chance that superior performance will be rewarded and inferior performance will be punished. Sappington and Weisman also offered several proposals to improve the likelihood of achieving this outcome including: (1) using automatic rate stabilizers (for example, interest

---

[45] See Braeutigam and Panzar, "Effects of the Change from Rate of Return to Price Cap Regulation."

[46] See David E. M. Sappington and Dennis L. Weisman, "Designing Superior Incentive Regulation: Accounting for All of the Incentives All of the Time," *Fortnightly* 132, no. 4 (February 15, 1994): 12-15.

rates) to adjust target prices;[47] (2) using benchmark or yardstick competition to judge performance; (3) giving firms a menu of choices of regulatory regimes, some of which might allow the firm to keep a larger fraction of gains if it elects a higher productivity hurdle (as the FCC has done for small LECs); (4) rewarding a firm for being willing to face more intense competition (this would operate similarly to the treatment of firms opting for a higher productivity hurdle); (5) limiting the potential for regulators to recontract; (6) using revenue sharing rather than profit sharing (because this lowers the regulators' incentives to disallow costs in order to increase shared profits); (7) having legislation that defines the plan; and (8) making plans as comprehensive and complete as feasible (although it is difficult to cover all contingencies).[48]

Planning for *Ex Post* Analysis

A good alternative regulation plan should have a built-in evaluation and analysis component. Ideally, the evaluation plan would be adopted more-or-less simultaneously with the adoption of the alternative regulation plan. The assessment plan will be an integral part of an alternative regulation plan, and would provide for collecting the data necessary for *ex post* analysis. The total package might include the following:

- A clear statement of the goals or objectives that the alternative regulation plan is intended to achieve.

A clear statement of the plan's goals can provide a reality check. No alternative regulation plan can do everything, and unrealistic claims or goals can be seen to be hyperbole or overselling. Also goal conflicts can be better sorted out if they are clearly specified. Possibly

---

[47] Automatic rate stabilizers might act analogously to purchased gas or fuel adjustment mechanisms. The rate stabilizers could track changes in major inputs and adjust rate ceilings to reflect such changes.

[48] See David E. M. Sappington and Dennis L. Weisman, "Designing Superior Incentive Regulation: Modifying Plans to Preclude Recontracting and Promote Performance," *Fortnightly* 132, no. 5 (March 1, 1994): 27-32.

a prioritization or hierarchy of goals could be adopted in this stage. A time path for some goals might also be adopted.

- A clear description of the policies or tools that the plan will use.

This part of the plan would help determine whether the alternative regulation plan is likely to be effective. There must be some belief that the tools are sufficient to meet the objectives of the plan.

- A clear identification and description of the measures or indicators that will be considered in the evaluation process.

The measures or indicators must be related to both the goals of the plan and the tools or policies of the plan. Some of the measures may operationalize the plan's goals and objectives; other measures or indicators may be intermediate to the goals themselves.

- A clear description of the feedback mechanism to be used if the results achieved under the alternative regulation plan do not meet the goals or expectations.

A good plan would be self-healing or fault tolerant. This means that the plan could incorporate a nearly automatic mechanism for making *moderate* midcourse corrections based on comparison of the achieved measures of relevant indicators relative to the goals established for them. Gross failure to meet goals could require more analysis—for instance, to determine whether the goals or tools were inappropriate.

Ex Post Analysis

Alternative regulation needs to be analyzed and evaluated after it has been in operation for some period. Regulators should not put alternative regulation into place without allowing for periodic review and possible adjustment. However, the timing of reviews poses another dilemma: reviews that come too soon, or too often, and allow for major adjustments or recontracting, may reduce the incentives for the sorts of positive behavior that regulators are

attempting to promote. Although interim or tracking reviews can and should be done on an ongoing basis to identify the shortcomings of the plan, a thorough review should come no sooner than the three-year mark in order to allow the utility, its customers, and its competitors to adjust. A thorough review after, say, five years would allow all parties to adjust their behavior; the five-year period would also allow for the plan to be in operation over a complete economic business cycle of the national economy.

The timing of the initial review, the factors that will be considered, and the method by which the plan will be evaluated should all be determined (at least tentatively) when the alternative plan is adopted or very soon thereafter. This allows for a planned and orderly process of collecting and analyzing data, and will produce a more useful review. Economic factors or indicators that might be considered for analysis are discussed in a later section.

It has generally been argued that alternative regulation will have an impact on telephone utilities, consumers, competitors, regulators, and others. After alternative regulation has been in effect for a reasonable period, regulators should be able to ascertain the extent to which this impact has been positive and significant. Put simply, alternative regulation is claimed to result in more efficient behavior by utilities (which are rewarded with higher profits), more innovation, faster deployment of advanced services and technologies, more competition, a less costly regulatory process, lower overall costs to consumers, and enhanced economic growth and development. Moreover, it is claimed that all this can be accomplished without degradation of service quality. Indeed, it is often argued that service quality may even improve as advanced services and equipment are deployed. If this is so, analysts should observe differences in behavior and outcomes among states that have alternative regulation and those that do not, and between holding company operations in states with alternative regulation and in states without alternative regulation. Also, in states with alternative regulation, differences should be observed between the behavior of individual companies before and after adoption of alternative regulation. These differences may also be found to depend upon the type of alternative regulation in place and the length of time it has been in place. One difficulty with this approach is that companies may behave as if they are under alternative regulation even though they are actually under traditional regulation. Given the general movement toward alternative regulation, and the fact that alternative regulation may

be "on the table" at commissions or legislatures in states that have not, as yet, adopted it, companies may rationally believe or expect a shift to alternative regulation and alter their behavior in anticipation. Such anticipatory shifts would make it more difficult to determine the effect of alternative regulation.

Economic Indicators or Measures

If alternative regulation is successful, regulators should be able to observe measurable and favorable impacts on several indicators or variables of interest. Useful indicators, which should respond favorably to the adoption of alternative regulation, include those that provide information on the utility's financial health and efficiency, the QOS available to consumers, the vitality of competition in telecommunications markets, regulatory costs including delay, and overall societal measures including growth and development.

Firm Indicators

If, as claimed, alternative regulation allows utilities to be more profitable and gives them incentives to be more efficient, there are several variables regulators might observe. One key indicator is the level of profits or the earned rate of return of utilities operating under alternative regulation. Similarly, if profit sharing is adopted, then another indicator is the dollars available to be shared with consumers. Other indicators attempt to measure the absolute efficiency or improvements in the firms' efficiency in producing outputs at least cost. Such indicators include the costs incurred by the firm, employment levels, and increases in productivity of the labor and capital used by the firm. Another area of interest is whether the firm is becoming relatively more efficient in the X-efficiency sense (that is, is the firm moving relatively closer to producing at least cost). Also of interest are the changes in the prices that the firm charges for different services. Regulators might want information on the rate of price increase for monopoly services as contrasted with the rate of increase for competitive services. Other indicators that regulators should consider are the firm's investment spending on new infrastructure and on innovation (including R&D expenditures).

Quality of Service Indicators

It is also claimed that alternative regulation can increase profits without resulting in a decrease in service quality. Regulators attempting to evaluate the results of alternative regulation may wish to analyze indicators including the speed of enhanced service availability (ISDN, for example) and advanced technology deployment (digital switching, SS7, fiber in the local loop); penetration measures for advanced services; traditional quality measures, such as call completion time, call failure rates, switching speed, line noise, signal quality, number of outages; and other measures, such as installation and repair time, billing errors, and numbers of customer complaints.

Market and Other Indicators

Alternative regulation may also have an impact on the number and viability of competitors, the rate of competitive entry, and on trends in LEC/regional holding company market share in competitive markets. Other indicators include items such as a shortened delay due to the regulatory process and a reduction in the administrative costs of regulation (for the utility and for the commission). These costs include those resulting from rate cases and other formal proceeding costs and those resulting from decision lag and delay. Finally, all other things being equal, if the claims of proponents of alternative regulation are correct, economic growth and development should be more rapid under alternative regulation than under traditional regulation.

Measuring the Effects of Alternative Regulation

In general, if alternative regulation has the claimed or desired effects, firms or jurisdictions operating under alternative regulation should show better results or outcomes for some variables of interest than those without alternative regulation, and the rates of change in these variables should also respond favorably after the adoption of alternative regulation. Although it is not likely that all variables will respond in a uniformly favorable manner, if alternative regulation is successful, many of these measures should be seen to respond favorably. Ideally, the utility will be more profitable and more efficient; consumers will have

more services available, at lower cost, possibly from multiple providers; there will be more competition in telecommunications; service quality will not deteriorate, and it may improve; the direct and indirect costs of regulation will decrease; and economic growth and development will be faster than it would be under traditional regulation. In order to determine whether, and in what direction, these measures have responded to alternative regulation, a number of techniques may be used. Several quantitative techniques that are available to analysts are discussed briefly in the next section.

Empirical Techniques

A variety of empirical techniques are available for analyzing the impact of alternative regulation. These include both quantitative (statistical) approaches, which are briefly discussed in this chapter, and qualitative approaches, which are discussed in Chapter 4. Quantitative approaches vary from relatively simple to very complex; the available techniques include univariate and multivariate time-series and cross-section analyses. Although many good reasons exist for preference of clear and simple empirical analyses, complex analyses may also be required in order to distinguish the effects of alternative regulation from the effects of the many other forces that act upon the telephone industry.

One caveat must be emphasized: in general, causation cannot be proved. Though in some time-series analyses a limited form of causation can be inferred,[49] causation cannot normally be established; only association or correlation can be demonstrated. Even with good analysis, it is difficult to prove causation in the sense of proving, for example, "alternative regulation caused more rapid deployment of advanced technologies and services." Notwithstanding this caveat, if a body of empirical data consistently indicates strong association or correlation between alternative regulation and an outcome, some presumption of

---

[49] In time-series analysis, limited forms of causation can be inferred based upon the behavior of two variables over time. One such form of causation is "Granger Causation." See George G. Judge, R. Carter Hill, William E. Griffiths, Helmut Lutkepohl, and Tsoung-Chao Lee, *Introduction to the Theory and Practice of Econometrics*, second edition (New York: John Wiley & Sons, 1988), 767-770.

causation may be developed. Moreover, the presumption will be stronger to the extent that the results of numerous studies are similar and other possible causative factors are considered.

Linear Models: Analysis of Variance and Regression Analysis

Among the commonly used quantitative or statistical techniques are analysis of variance, regression analysis, and covariance analysis. These techniques are related in that each is an application of what statisticians call the "general linear model," and each can provide information concerning the effect of alternative regulation on variables on interest.

In a simple one-way analysis of variance, the independent variable could be the presence of alternative regulation—with alternative regulation being given the status of a classification or category variable. The observations on the variable of interest would be split into two groups based on whether alternative regulation was in force. If alternative regulation has an effect, the average or mean value of the variable of interest would be different in the two groups. In its simplest form, analysis of variance is based on a calculation of an overall mean or average value for the dependent variable for the entire observation set and for each of the subgroups—firms or states operating under alternative regulation and those operating under traditional regulation.

One way to present the results that does not require a formal analysis of variance is to display a chart showing the group means side-by-side. Such simple diagrams provide a useful visual comparison. However, since there is natural randomness in results for firms or states regardless of the type of regulation in place, differences in group means may be due to randomness or chance. Therefore, a formal analysis of variance includes a test of the hypothesis that the observed difference between the group means is due simply to chance. If there is a statistically significant difference in group means, it may be attributed to the type of regulation imposed, provided that other factors have been considered.

More complicated analysis of variance (two-way designs or factorial designs, for example) may use additional categorical variables to explain differences in outcomes. Often, however, other explanatory variables are numerical (number of access lines, area of service territory) rather than indicators of categories or classifications. In this case regression analysis or covariance analysis can be used to account for the effect of the numerical variables.

Because it is capable of including multiple classification variables and numerical variables, regression analysis is more general than analysis of variance. Regression analysis can include classification variables by treating them as "dummy" or shift variables, which are assigned a value of 1 if an observation came from a firm under alternative regulation and are assigned a value 0 if an observation came from a firm under traditional regulation. A regression equation that includes both dummy and numerical variables might be in the following:

$$Y_i = \alpha + \beta_1 D_{1i} + \beta_2 D_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \epsilon_i .$$

In this regression equation, Y represents the variable of interest (also called the dependent variable); the subscript i indicates that this is the i^{th} observation or data point; D_1 and D_2 are dummy variables, which are either 0 or 1; X_3 and X_4 are numerical variables that the analyst believes influence Y; and ϵ_i is the random error term for the i^{th} observation. D_1 and D_2 could indicate the type of alternative regulation in place. For instance, D_1 might indicate alternative regulation, and D_2 might indicate whether profit sharing was used.

Since the analysis is intended to identify the effect of alternative regulation, the inclusion of the numerical variables, X_3 and X_4, is designed to create, as nearly as possible, a nonexperimental analogue to a controlled experiment. Although the aim of the study is not principally to measure the effect of X_3 and X_4, their inclusion in the regression allows the analyst to take into consideration their effects on the dependent variable. By doing so, the effect of alternative regulation can be better isolated and measured. For example, the dependent variable, Y, might be a measure of the deployment of an advanced technology in a state or service area; X_3 might stand for the proportion of access lines in metropolitan areas; and X_4 might stand for the average personal income in the area.

The regression coefficients, β_1, β_2, β_3, and β_4, indicate the estimated effect on the dependent variable of a one-unit change in the independent variables, holding the values of the other independent variables constant. In this regression model, if D_1 indicates the presence of alternative regulation, and if the estimated value of β_1 is statistically significant, then alternative regulation may have effected the variable of interest. If regulators hoped that

the variable would respond positively, and if β_1 is positive, alternative regulation can be believed to have had the desired effect.

This regression equation is intended to be representational. In an actual regression model there might be many dummy and numerical variables. However, if additional explanatory variables are added to the right-hand side of the equation, it may become difficult to calculate and interpret the regression coefficients. This equation would be used to estimate the effect of alternative regulation on one variable of interest. Since regulators are likely to be interested in the effect of alternative regulation on many variables, numerous such regressions would have to be estimated to obtain a complete picture of the impact of alternative regulation. Regression models are adaptable to a number of circumstances. Although the equation presented above is linear in form, regression analysis is highly adaptable. Certain nonlinear functional forms can be estimated by transforming the independent variables, the dependent variables, or both. Although the calculations are more complex, regression techniques are also adaptable to purely nonlinear models.[50]

Cross-Sections and Time-Series

A regression of the type discussed above could be used in either cross-section studies or time-series studies. In cross-section studies, the analyst would collect contemporaneous measurements of the dependent and independent variables for a group of firms or jurisdictions and use the model to estimate whether the type of regulation imposed affects the dependent variable. The aim is to detect differences in whatever variable is of interest (profits, prices, technology deployment, QOS, and so on). Examples of simple cross-section studies include those presented by Milton Mueller, who compared results in Nebraska after deregulation with results in those U S WEST states that still had traditional regulation.[51]

---

[50] For a discussion of nonlinear estimation see A. Ronald Gallant, *Nonlinear Statistical Models* (New York: John Wiley & Sons, 1987).

[51] Milton L. Mueller, *Telephone Companies in Paradise: A Case Study in Telecommunications Deregulation* (New Brunswick, NJ: Transaction Publishers, 1993).

In a time-series study, the analyst would compare the behavior of the dependent variable before and after the change of regulatory regime. If alternative regulation has the predicted results, a favorable or desired change in the time path of the dependent variable should be detected. One form of time-series analysis involves analyzing the behavior of the dependent variable for one firm or one state over time—using a dummy variable to identify the portion of the time series during which the firm was under alternative regulation. One problem with this type of analysis is that, because alternative regulation has only recently been implemented, there may be too few observations or data points on individual firms or states for this analysis to be effective in obtaining a good estimate of the effect of alternative regulation.

Another form of time-series modeling that might prove useful is an "event study," which examines the effect of a switch to alternative regulation on a number of firms, even though the switch may have happened at different times.[52] For an event study, the analyst would examine the behavior of one or more variables of interest for a group of firms or states for several years prior to and several years after implementation of alternative regulation. The timing of the implementation need not be the same for all observations; rather, the change in the dependent variable after implementation of alternative regulation is estimated for a number of firms or jurisdictions, under the assumption that a predictable response pattern can be found.

If it were possible to pair each firm experiencing an event such as a shift to alternative regulation with a firm that was not shifted to alternative regulation, a baseline could be established for the behavior of the dependent variable. However, this is not likely to be feasible. Event studies require information on the behavior of many firms, most of which would be outside an individual Commission's jurisdiction. Even so, regulators in individual states could still benefit from examination of the behavior of variables of interest in other jurisdictions.

---

[52] Event studies are widely used in empirical finance to examine the behavior of various firms' stock prices before and after announcement of mergers, buyouts, or stock splits. Although each firm experiences these events at different times, common patterns do emerge, and anomalous behavior can be detected.

In addition to separate cross-section and time-series studies, models exist that combine cross-sections and time-series in a single estimation to estimate the effect of alternative regulation across firms and over time. Combined cross-section, time-series models may prove to be useful, but estimation of mixed models is more complex than either cross-section or time-series models by themselves.

Benchmark Studies

One use of empirical analysis might be to assess the efficiency of telephone companies in producing their outputs at least cost. If adoption of alternative regulation leads firms to become more efficient, regulators should observe that firms operating under alternative regulation should become relatively more efficient over time compared with firms under traditional regulation. Benchmark studies can be used to estimate what it "should" cost firms to produce a certain mix of services.[53]

If benchmarks can be developed, firms under alternative regulation can be compared with those under traditional regulation. If alternative regulation is successful, firms operating under alternative regulation would be expected to become relatively more efficient than those operating under traditional regulation. In other words, firms under alternative regulation would move relatively closer to the benchmark over time. However, benchmarks may not be easy to develop, and such studies are not likely to be without controversy, as those firms judged to be relatively inefficient are not likely to accept such results.

---

[53] An early attempt at developing benchmarks using multiple regression analysis is William Iulo, "The Relative Performance of Individual Electric Utilities," *Land Economics* 38, no. 4 (November 1962): 315-26. Another possible source of benchmarks would be simulation studies such as David Gabel and Mark Kennet, *Estimating the Cost Structure of the Local Telephone Exchange Network* (Columbus, Ohio: The National Regulatory Research Institute, October 1991).

In addition to the regression approach used by Iulo and the simulation approach of Gabel and Kennet, data envelopment or frontier models may prove useful.[54] Data envelopment and frontier models differ from traditional regression models in that, instead of estimating an average cost of producing a set of outputs, they estimate a minimum cost frontier or envelope.[55] These techniques may offer regulators a means by which they can judge whether firms do, in fact, move closer to the efficiency frontier or best practice envelope after alternative regulation is adopted. If this proves to be the case, alternative regulation may have achieved at least one of its objectives.

In traditional regression, the data points for individual firms are scattered above and below the estimated regression line. In a frontier approach all the data points lie on or above the frontier, indicating that no firm has been able to beat the frontier. In other words, because the frontier identifies the minimum cost of producing given levels of output, no observation lies below the frontier. In frontier models, it is possible to estimate and compare the relative

---

[54] A thorough discussion of the theory underlying estimation of frontier cost and production functions and data envelopment analysis may be found in Rolf Färe, Shawana Grosskopf, and C. A. Knox Lovell, *Production Frontiers* (New York: Cambridge University Press, 1994). The seminal work in data envelopment analysis is A. Charnes, W. W. Cooper, and E. Rhodes, "Measuring the Efficiency of Decision Making Units," *European Journal of Operational Research* 2, no. 6 (November 1978): 429-44. Applications of these techniques to measure efficiency in various contexts may be found in Loretta J. Mester, "How Efficient are Third District Banks?," *Federal Reserve Bank of Philadelphia Business Review* (January-February 1994): 3-18; Toshiyuki Sueyoshi, "Stochastic Frontier Production Analysis: Measuring Performance of Public Telecommunications in 24 OECD Countries," *European Journal of Operational Research* 17, no. 1 (January 1993): 1-13; and Toshiyuki Sueyoshi, *Measuring Efficiencies and Returns to Scale on Nippon Telegraph & Telephone in Production and Cost Analyses*, mimeo, (Columbus, Ohio: College of Business, The Ohio State University, August 1993).

[55] In the usual regression equation the sum of the estimated errors is zero, so that (approximately) as many individual observations will be associated with positive as negative prediction errors. In frontier or envelope analysis the model estimates the limits of the data, so all the errors are one-sided. For example, a frontier cost function would constrain all errors to be positive, so that no observation lies below the cost frontier. This implies that no firm in the data set was able to produce its output at a lower total cost than that implied by the frontier. The frontier, thus, represents the set of efficient combinations of cost and output. That is, no firm can produce its output at a cost lower than the frontier indicates.

efficiencies of various firms by comparing their percentage distance from the frontier. If alternative regulation is successful in promoting efficiency, firms should be observed to move closer to the frontier after adoption of alternative regulation.[56]

Multivariate Analysis

Most of the approaches previously mentioned rely on univariate analysis, which attempts to estimate the effect of alternative regulation on individual variables one at a time. Univariate analysis is commonly used and is easily interpreted. Another approach, multivariate analysis, attempts to simultaneously and jointly estimate the effect of alternative regulation on multiple measures of interest.[57] One advantage of multivariate analysis is that it accounts for relationships among several dependent variables. It might, therefore, detect differences in the vector of means of a set of variables of interest when the differences in the means of the individual variables of interest would not be considered significant in univariate analysis. This might be especially useful when trying to model the effect of alternative regulation on QOS, or technology deployment, both of which are multidimensional concepts. It might be the case that a significant difference exists between a group of variables under alternative regulation, yet the differences would not be identified under univariate estimation. A disadvantage of multivariate analysis is that interpretation of results may be more difficult than in the case of univariate analysis, because of the added complexity.

---

[56] Note that utilities operating under traditional regulation may have been engaging in noncost-minimizing behavior. To the extent that such behavior has been common, empirical estimates of the frontier may overstate the true minimum cost or production.

[57] Multivariate analysis can be distinguished from univariate techniques, such as multiple regression. Multiple regression estimates the effect of multiple independent or explanatory variables on a single dependent variable. Multivariate analysis estimates the effect of single or multiple independent or explanatory variables on a set of dependent variables. Estimation of a system of simultaneous multiple regression equations would be one multivariate technique. Other multivariate techniques include multivariate discriminant analysis and multivariate analysis of variance and covariance, each of which is a generalization or extension of a univariate technique. For a discussion of these techniques see Richard A. Johnson and Dean W. Wichern, *Applied Multivariate Statistical Analysis* (Englewood Cliffs, New Jersey: Prentice-Hall, 1988).

For example, multiple discriminant analysis might divide or sort observations into two groups, those under alternative regulation and those under traditional regulation. Statistical procedures could then be used to determine whether the vector of mean values of a set or group of variables of interest is different between the two groups. If a difference is found, the model can be tested by assigning each observation to the group to whose mean vector the observation is closest. If the model is useful, assignment of this type will be correct more than pure chance would allow.[58]

Suppose that a sample includes twenty-five observations on a set of QOS measures from firms under alternative regulation and twenty-five observations from firms operating under traditional regulation. Random coin-toss assignment of observations to the two groups would be expected to be correct fifty percent of the time. If the discriminant model correctly assigns ninety percent of the observations to the correct group, an analyst could believe that there was, indeed, some difference between the two groups. Furthermore, if other factors were properly accounted for, the difference might be attributed to different regulatory regimes.

Summary Thoughts and Conclusion

There is a worry concerning empirical analysis that must be recognized. Given abundant data, numerous possible sets of explanatory variables that might be used, and a variety of estimation techniques, it is difficult for the analyst to avoid falling prey to data mining or data torturing—estimating many models until "pleasing" results emerge. An anonymous saying is that "if you torture the data long enough, they will confess." Indeed, easy access to powerful computers makes it almost too easy to sift through the results of numerous estimations in order to find one or two that confirm the analyst's prior beliefs. In

[58] Discriminant analysis is similar to regression techniques using "limited dependent variables." In these techniques, the dependent variable might be a categorical or classification variable, rather than a numerical variable. For information on these techniques see G. S. Maddala, *Limited-Dependent and Qualitative Variables in Econometrics* (New York: Cambridge University Press, 1983).

that case the analyst is using statistical analysis much as a drunk uses a lamppost: "for support rather than illumination." Some time ago, a scholar offered a warning that is still useful to remember:

> If an economist is clever enough or persistent enough, he can always find an equation that fits the available data fairly well; he may also convince himself that it is a theoretically reasonable equation. The danger lies in being too clever or too persistent, and finding an equation that fits the available data well enough but is nevertheless wrong because it describes temporary or accidental features of the available data, rather than the enduring systematic features.[59]

This warning is intended neither to denigrate empirical analysis nor to deter regulators from engaging in it. Instead, it should serve as a reminder to regulators not to rely on any single empirical study. There are numerous possible analysis and evaluation techniques, none which, by itself, is likely to be sufficient to reach a definitive conclusion as to the effect of alternative regulation. Moreover, regulators need to have answers to many questions about the effects of alternative regulation; therefore, multiple analyses will be required, and a variety of techniques will be employed. Because the effects of the various state alternative regulation plans are of interest to regulators in other states, and data requirements are substantial, regulators might consider joint or even national studies to assess and monitor the effects of alternative regulation. In time, sufficient evidence will accumulate at both the state and national level for regulators and legislators to assess the impact of various alternative regulation plans. Until such evidence is accumulated, regulators and legislators should treat alternative regulation as experimental and be willing to modify it, as necessary, to ensure that it meets the needs of all parties.

Applied economics and regulatory economics have developed methods to allow the analytical power of theoretical economics to be meaningfully applied in a real world setting. This chapter adopts the perspective of applied economics and reviews the need for, and criticism of, traditional regulation. The pricing reforms proposed and implemented as

---

[59] See Carl F. Christ, *Econometric Models and Methods* (New York: John Wiley & Sons, 1966), 9.

solutions are also examined from this perspective. Basic benchmarks that analysts can use to monitor the empirical analytical techniques underlying economic studies have also been presented. Armed with this information commissions should be able to better understand and analyze the economic analyses submitted by contending parties about the successes achieved and/or failures evidenced via a state's pricing reform.

CHAPTER 6

MEASURING COMPETITION

Introduction[1]

One of the factors driving the move toward alternative regulation is the current trend toward the expansion and intensifying of competition. As noted in prior chapters, technological developments have eroded the strength of natural monopoly claims in many market segments. In addition, as technological barriers to entry were lowered, regulatory policies that based prices on aggregation and averaging and that created internal subsidies to benefit certain users made competitive entry inevitable. This became especially true as legal and regulatory barriers to entry were also reduced. Various entrants either have entered or are poised to enter local access markets. These include interexchange carriers (IXCs), cable television system operators, competitive access providers (CAPs), cellular providers, personal communication system operators (PCS), and others (electric utilities). Everyone seems to "want in" on the envisioned explosion of services and usage of the telephone network. LECs see competitors everywhere. They argue that (1) they are just one of the players in the new markets, (2) they no longer have significant market power, (3) they should be free to compete with all new entrants, and (4) regulatory restrictions on their offerings are unnecessary and anticompetitive.

Existing or potential competitors in the local access market argue that allowing the LECs (especially the Regional Bell Holding Companies) into cable television, information services, or interLATA toll services would lead to remonopolization of those markets. Each of these interest groups has some plausible arguments to make. They all involve the question

---

[1] Readers familiar with basic competition concepts and issues may choose to go directly to page 168.

of defining competition and determining whether the public interest would be well-served if various entities are allowed to enter.

In regulatory reform proceedings competition is an important issue that is invariably covered. This regulatory examination has several dimensions. The first is the notion of competition as a preferred end-state, or goal, toward which the telecommunications market is trending by itself and through the assistance of regulators. Different parties have different opinions on how fast and through which process the telecommunications market will evolve into a competitive market. The second dimension is the general tendency in regulatory pricing reforms to segment the originally fully regulated telecommunications market into regulated, partially regulated, and unregulated sectors. A third, but closely related, dimension is whether the market segmentation is done before empirical data are available indicating the presence of sustainable competition or if services are classified *ex ante* and assigned to the fully, partially, and deregulated sectors.

Whether services are classified based on actual competition or by stipulated competitive categories, commissions properly have devoted a great deal of attention on how to measure and conceptualize competition. Interested parties have generally deluged commissions with studies, task force papers, data, testimonials, and reports about the presence or lack of competition. Incumbents have usually reported that competition is facing them on all fronts and they ask for regulatory relief in the form of a particular regulatory reform. Challengers present reports and data that celebrate the emergence of fledgling competition, but warn that the incumbents retain sufficient market power even under proposed regulatory reforms to effectively eliminate competitors and to keep only as much fringe competition as is needed to prevent reregulation. Challengers say such an outcome will keep prices high, will slow innovation, and will delay widespread deployment of advanced infrastructure and telecommunications services. Incumbents counter by saying such fears are self-serving attempts to gain additional restrictions on the ability of the incumbent to compete.

It is the intent of this chapter to present a brief review of the benefits of competition and the methods of determining whether a market is competitive. Also presented is a proposal for defining partial or high levels of competition in telecommunications markets. In addition, some other recommendations for commission oversight of the evolution of

competition are presented. By presenting this information in a nonadversarial setting, various competitive concepts and measures can be examined without advancing the interests of parties in a particular proceeding. Armed with this information commissioners and senior commission staff should be better able to evaluate and appraise competition information submitted in the course of a regulatory reform proceeding.

Competition

Assumptions of the Competitive Model

One problem is that competition means one thing to a policymaker and another to theoretical economists. *Perfect competition* for an economist describes a market organization in which all firms produce homogeneous, perfectly divisible output; there are no barriers to entry or exit; producers and consumers have full information; there are no transactions costs; there are no unpriced externalities in production or consumption; and all participants are price takers in the sense that they do not believe that their individual behavior can affect prices. Competition implies that individual firms believe that the own-price elasticity of demand for their product is large in absolute terms—that is, small increases or decreases in the price they charge (holding constant the prices charged by competitors) will have large effects on the amount of the demanded output. In addition, there can be no collusion among groups of buyers or sellers to control price or output. In this view of competition there are numerous buyers and sellers, none of whose production or consumption is large relative to the total market.

In reality, competition never looks like this and for most policymakers looks more like "workable competition," which is described later on in this chapter.

Competitive Outcomes

In general, if a market is competitive certain results would occur. First, if all economic activity is undertaken in competitive markets, an economically efficient allocation will result. This means that all goods are produced at their minimum cost or, equivalently, given the resources available, total output is maximized. Under these conditions, no reallocation of production or of the factors of production could produce greater output.

Second, an efficient allocation of resources maximizes the sum of consumer and producer surplus. In a simple demand and supply model, consumer surplus can be identified as the area between the market demand curve and the equilibrium price of a product. Consumer surplus is a measure of the difference between the maximum total amount consumers would be willing to pay for the equilibrium quantity of a product rather than do without it and their actual total expenditures on the product. Similarly, producer surplus can be identified as the area between the market supply curve and the equilibrium price of the product; it is a measure of the difference between actual total revenues received by the producers and the minimum total amount producers would be willing to accept for the equilibrium quantity of a product rather than forego all sales.

Third, in a competitive market there are no long-run economic profits. Each firm earns just enough to keep it in the market. The firm covers its total economic cost, including a normal return on investment. A competitive market is dynamic rather than static: firms enter and exit the market depending upon whether they believe they can cover their total costs.

Uses of the Competitive Model

Analysis of the competitive model forms the basis for much of economic theory. One use of the competitive model, or other models, is to make inferences about the response of market participants when confronted with changes in costs or demand conditions. This predictive use of the competitive model compares the comparative static properties of the competitive model with those of other models.

The competitive model is also often used in a normative sense to evaluate suboptimal models or results. Because the competitive model indicates the best possible outcome (using economic efficiency as the norm), the results of market organizations that are not competitive can be compared. This normative use of the competitive model compares the long-run equilibrium under competition to that under other market structures.

Perfect Competition vs. Workable Competition

Pure or perfect competition is a theoretical paradigm rather than a real-world market structure or organization. All existing market structures produce outcomes in terms of prices, quantities, and resource allocation that diverge from the competitive ideal. Although other market structures do not comport completely with the competitive model, some may be considered as being approximately or workably competitive. The concept of workable competition is straightforward. If some of the restrictions of the perfectly competitive model are relaxed, the outcome may still be reasonably close to the competitive result. Among the assumptions that can most likely be relaxed are such requirements as large numbers of buyers and sellers, homogeneous or undifferentiated products, and ubiquitous information availability. Among the assumptions of the competitive model that are crucial is that there be no significant barriers to entry or exit.[2] One observer described workable competition as follows:

> Competition ... may be regarded as effective or workable if it *offers buyers real alternatives* sufficient to enable them, by shifting their purchases from one seller to another, substantially to influence quality, service, and price. It *requires the presence in the market for several sellers, each of them possessing the capacity to survive and grow, and the preservation of conditions which keep*

[2] A barrier to entry is any factor that prevents a firm from quickly and costlessly entering or exiting a market. Normally, high profits in a market attract entry, which tends to lower profits to a normal level. If there are barriers to entry, high profits may be persistent. Examples of barriers to entry include the existence of patents and licensing or franchise restrictions. Other barriers to entry might include attempts by incumbents to raise entrant's costs or the necessity to overcome an incumbent's brand recognition and customer loyalty. The existence of significant economies of scale which makes the cost minimizing scale of operation large relative to the size of the market may also create a barrier to entry.

> *alive the threat of potential competition from others.* In brief, *competition may be said to be effective or workable whenever it operates over time to afford buyers substantial protection against exploitation by sellers.*[3]

Although the concept of workable competition is straightforward, a precise determination of whether a market is workably competitive may be difficult. Some years ago George Stigler, in discussing the concept of workable competition, observed that

> How close an industry should be to competition . . . to be workably competitive has never been settled. Indeed the criteria . . . which deserve most weight in any application of the concept have not been agreed upon. Two competent persons who study a particular industry can disagree on its workable competitiveness, and there exists no analytical basis for eliminating the disagreements.[4]

Public Policy and Competition

Notwithstanding the difficulties encountered in defining workable competition, one view is that public policy should attempt to ensure workably competitive structures or promote approximately competitive outcomes unless the cost of achieving such structures or outcomes exceeds the net benefits. Thus, antitrust policy and utility regulation—though through different mechanisms—aim at moving structures or outcomes more towards the competitive ideal or at least inhibiting the most egregious anticompetitive behavior. Antitrust policy focuses on limiting the ability of firms to create or exercise monopoly power. Regulation focuses on harnessing monopoly power to achieve public ends.

---

[3] The concept of workable competition has been credited to John M. Clark, "Towards a Concept of Workable Competition," *American Economic Review* 30, no. 2 (June 1940), 241-56. This explication comes from Clair Wilcox, *Competition and Monopoly in American Industry* (TNEC Monograph 21, 1941) as excerpted in Joel Dean, *Managerial Economics* (New York: Prentice-Hall, 1951), 55-56.

[4] George J. Stigler, *The Organization of Industry* (Homewood, Illinois: Richard D. Irwin, 1968), 12.

An idea that must be stressed is that simply having numerous firms in a market does not by itself make the market workably competitive. If one firm dominates the market because of its relative size or technological advantage such that it sets the price and/or technology standard for its competitors, the market cannot be said to be workably competitive.

One difference between the common use of the term "competition" and economists' more formal use of the term, is that, in common use, competition refers to the active rivalry and chess-match behavior common in the real world of business. The common use of the term may be exemplified by two furniture stores across the street from one another. More likely than not, the managers will have an active sense of competition and rivalry. They may compete with "sales," special terms, free delivery, and through other means. Each will likely believe that the prices and terms offered by the other affects their own ability to attract customers.

As economists use the term, competition is impersonal or anonymous: no participant takes any notice of the actions of any other individual player. The economist's concept of competition may be exemplified by the actions of neighboring farmers. Although they are competitors, neither has any perception that the actions of the other affects their ability to sell their crop or the price they will receive.

Competition may also be visualized as an evolutionary process rather than a static result. Competition is a force that acts upon market participants to lead them to seek more efficient methods of satisfying consumers' wants. The ability of firms to capture, if only briefly, the gains from "building a better mousetrap" provides positive incentives for innovation and efficiency. Competition also provides negative incentives: firms that do not seek efficiency and adopt innovations will not prosper, and they may perish.[5]

---

[5] The idea of competition as a force exerting pressure on firms to be efficient producers and that the lack of competitive pressure may lead to noncost minimizing behavior has been explored by Harvey Leibenstein, among others. See Harvey Leibenstein, "Allocative Efficiency vs. X-Efficiency," *American Economic Review* 56, no. 3 (June 1966): 392-415.

Measures of Competition and Monopoly Power

Market-Level Measures

It is sometimes taken as axiomatic that an industry comprising numerous firms—none of which has sufficient impact to appreciably affect market price—will behave more in accord with the competitive ideal than an industry composed of a few large firms—each of which may have some impact on the market.[6] Thus, summary measures can be developed that provide *prima facie* evidence as to the degree of competition in a market. Two common market-level measures have been used, the n-firm concentration ratio and the Herfindahl-Hirschman Index (HHI). Both are useful in investigating competition in the context of regulatory reform proceedings.

Concentration Ratios

The n-firm concentration ratio, CR_n, is simply the percentage of total industry sales (or other indicators such as capacity, output, or employment) attributable to the n largest firms. Commonly, four-firm or eight-firm concentration ratios are used. The four-firm concentration ratio, CR_4, would be calculated by setting n = 4 in the following formula:

---

[6] For instance, much of traditional industrial organization economics (both theoretical and empirical) has been built around the Structure-Conduct-Performance paradigm. In this paradigm—often associated with the work of Edward S. Mason and Joe S. Bain, the structure of an industry (number of firms and their size distribution, degree of product differentiation, cost structure, and other measures) influences the conduct of the firms in the industry (extent of rivalry, advertising, vertical integration, research and development spending, and pricing, among others) which determines the performance of the industry (production and allocative efficiency, and equity, for instance). For a concise discussion of the Structure-Conduct-Performance paradigm see F. M. Scherer and David Ross, *Industrial Market Structure and Economic Performance*, third edition (Princeton, New Jersey: Houghton Mifflin, 1990), Chapter 1.

$$CR_n = \sum_{i=1}^{n} S_i \ .$$

In the summation formula S_i could represent the i^{th} firm's share of market or industry sales or some other measure of firm size (capacity or employment). Only the shares of the n largest firms (four in this case) would be used in the calculation. As an example of the use of concentration ratios in public policy, the Antitrust Division of the U.S. Department of Justice classifies markets as follows: (1) "unconcentrated" if CR_4 is less than 50 percent, (2) "moderately concentrated" if CR_4 is between 50 percent and 70 percent, and (3) "highly concentrated" if CR_4 is greater than 70 percent.[7]

---

[7] Although this classification was developed to aid in determining whether the Justice Department would intervene in proposed mergers, it provides an indication from the unregulated sector as to what might be considered a workably competitive market based on market structure indicators.

The Department of Justice's guidelines focus on the effect of a proposed merger on an industry or market's Herfindahl-Hirschman Index (HHI), but other factors (entry conditions, the health or viability of the merger partners, and the nature of the product) can be considered in determining whether or not to oppose a merger. The classification scheme used critical HHI thresholds rather than concentration ratios. However, the HHI thresholds were found to corresponded roughly to four-firm concentration ratios of 50 percent and 70 percent, respectively. Department of Justice, Antitrust Division, *Merger Guidelines* (Washington, D.C.: U.S. Department of Justice, Antitrust Division, June 14, 1984), Sec. 3.1.

Additional discussion of the 1984 merger guidelines may be found in numerous sources including: Lawrence J. White, "Antitrust and Merger Policy: Review and Critique," *The Journal of Economic Perspectives*, 1, no. 2 (Fall 1987): 13-22; Ernest Gellhorn, *Antitrust Law and Economics in a Nutshell*, (St. Paul, Minnesota: West Publishing Co., 1987), 113-121 and 354-374; John S. McGee, *Industrial Organization* (Englewood Cliffs, New Jersey: Prentice-Hall, 1988), 220; and John S. Horning, Raymond W. Lawton, Jane L. Racster, William P. Pollard, Douglas N. Jones and Vivian W. Davis, *Evaluating Competitiveness of Telecommunications Markets: A Guide for Regulators*, (Columbus, Ohio: NRRI, January 1988), 47-52. A state-level view is given in Susan Beth Farmer, "Market Power and the National Association of Attorneys General Horizontal Merger Guidelines," *Antitrust Law Journal* 60, no. 3 (Developments 1991-92): 839-48.

Herfindahl-Hirschman Index

The HHI is calculated by summing the squares of the market shares of *all* firms in an industry or market. Thus, the HHI makes use of more information about the size structure of an industry than do concentration ratios.[8] The HHI uses more information about the structure of a market than does CR_n. In addition to considering the shares of the largest n firms, the HHI is influenced by the distribution of shares among those n firms. For a given CR_4, the more unequal the shares of the four largest firms, the higher will be the HHI. It is given by the following formula:

$$HHI = \sum_{i=1}^{n} S_i^2 \ .$$

In a pure monopoly market, the HHI would equal 10,000 (100 squared). In a perfectly-competitive market, the HHI would approach zero as a limit. As a specific example, if there were one hundred firms each with 1 percent of the market, the HHI would be 100. Department of Justice guidelines for the HHI indicate that unconcentrated markets have an HHI of less than 1,000 (implying that the largest firm's share would be less than 32 percent). Moderately concentrated markets have an HHI between 1,000 and 1,800 (implying that the largest firm's share would be less than 43 percent). Highly concentrated markets have an HHI greater than 1,800.[9]

---

[8] It may be impractical to determine the market shares of all firms in a market, especially if there are numerous small producers. Luckily, HHI values, at least in a qualitative sense, are not very sensitive to the shares of small participants (less than 5 percent) so that interpretation or classification of markets is not likely to be affected by their omission.

[9] Among the possible industry structures that would give a HHI of 1,000 would be an industry composed of ten firms, each with 10 percent of the market. Such an industry would have a CR_4 of 40 percent and, thus, would also be classified as unconcentrated under that criteria. An industry with five firms each controlling 20 percent of the market would have an HHI of 2,000 and would be classified as concentrated. In addition, its CR_4 would be 80 percent so that it would also be classified as concentrated using that criteria. The two classifications are not always in agreement, so it is possible that an industry would be classified differently under the two measures. For example, an industry with one firm with 40

Applying Concentration Ratios and Herfindahl-Hirschman Indexes

In real markets, market power is a matter of degree rather than strict presence or absence. Therefore, analysts have attempted to devise simple market measures that can be used to discriminate between markets that are workably competitive and those that are not. Unfortunately, the empirical results are mixed, and no simple and uniformly acceptable measures have, as yet, emerged. There are a number of reasons for this lack of consensus.

It can be difficult to define a market strictly in terms of the product because similar products may be in different markets, and dissimilar products may be in the same market. For example, a Rolls-Royce and a Hyundai are both automobiles and are both sold in numerous countries, but few analysts would think of them as being in the same market. At the other extreme, seemingly different products may be in competition with each other. For example, although the first image of competition in the fast food market may be that it is among McDonalds, Wendy's, Burger King, and other hamburger chains, clearly KFC, Pizza Hut, Taco Bell, and others are in the market. Casual dining restaurants such as Bob Evans, Shoney's, Frisch's, and the "mom and pop" restaurants and diners are also in the market to a certain extent. In addition, many convenience foods sold in grocery stores also compete with "fast food" for the consumer's dollar. What this implies for the analyst is that calculating concentration measures and/or Herfindahl indices may be difficult because the market may be much broader than simply "fast food hamburgers."

Another difficulty is that the same product or service may be sold in more than one market or to more than one customer group, and the extent of competition may differ depending on the type of buyer.[10] Moreover, measuring market shares may depend on the

---

percent of the market and with ten other firms, each with 6 percent, would have a CR_4 of 58 percent and an HHI of 1,960.

[10] For example, business and residential customers use some of the same telecommunications services but the degree of competition and, thus, market power varies across the two groups.

analyst's choice of geographic limits. Should a market be defined as SMSA-specific,[11] statewide, regional, national, or international? Clearly, the measures of market structure will depend on the scale chosen. These are not easy questions, and the choice must be made carefully.

Furthermore, there may not be a strict correspondence between market measures and performance. Market behavior may be affected by potential entrants, as well as current players. If this is the case, the link between observed CRs and HHIs and performance may be weak. Therefore entry conditions must be considered when attempting to forge a link between existing industry structure and performance measures. Other things being equal, a market will behave more like the competitive ideal if barriers to entry are low than if they are high. This point has also been expressed forcefully in the analysis of contestable markets in which the behavior of incumbents is constrained by the *potential* entry of firms that might not be current participants. The more contestable the market, the closer to competition will be its performance. Indeed, if a market is perfectly contestable there need not be large numbers of players for the efficient outcome to emerge.[12]

---

[11] Standard metropolitan statistical area as defined by the Department of Commerce.

[12] The concept of contestability focuses on the existence of sunk costs (costs that cannot be recovered if a firm enters and exits a market) as being the major barrier to entry, assuming that there are no government-enforced restrictions on entry and exit. In a contestable market, the structure that emerges will be the least-cost structure. See William J. Baumol, John C. Panzar, and Robert D. Willig, *Contestable Markets and the Theory of Industry Structure*, revised edition (New York: John Wiley & Sons, 1986).

Two views on the applicability of the theory and the contestability of the inter-LATA toll market are found in William G. Shepherd and Robert J. Graniere, *Dominance, Non-Dominance, and Contestability in a Telecommunications Market: A Critical Assessment* (Columbus, Ohio: The National Regulatory Research Institute, March 1990). See also Patricia M. Worthy, "'Unregulation' as Regulatory Reform in Telecommunications," *Public Utilities Fortnightly*, (March 30, 1989): 9-13.

However, just as there are few examples of perfectly competitive markets, there are few examples of perfectly contestable markets. Nonetheless, the ability and willingness of potential entrants to engage in "hit-and-run" entry[13] must be considered. In telecommunications, the activities of switchless resellers and aggregators and some providers of customer premises equipment (CPE) may be evidence of some degree of contestability of some segments of the market. However, those segments which require large fixed capital investments and time to develop a customer base and become profitable seem to be less like the contestable ideal because of the necessary commitment to the market and because the quick profits from hit-and-run behavior are not possible.

The competitiveness of individual markets also depends on the number, size, and sophistication of consumers, as well as the amount of information available to them. A market in which sophisticated consumers spend time searching for the best combination of price and service will be more effectively competitive than a market in which consumers make their choices based on factors such as habit and brand name.[14]

Firm-Level Measures

In addition to market-level measures, several firm-level measures have been suggested. These include the largest firm's market share, the relationship of price to marginal cost (or the Lerner Index), and simple profitability measures, such as the rate of return on equity. These will be considered in turn.

---

[13] "Hit-and-run" entry is entering a market to seize upon opportunities for profits before the dominant firms react, then leaving quickly and costlessly.

[14] For example, the telecommunications market for business customers, especially for large, multiline customers, is likely to be more competitive than the residential market because large users are better able to devote resources to selecting the best alternative. Indeed, large users may have telecommunications managers whose primary function is to arrange for cost-effective telephone services. For more on this point see Roger D. Colton, "Consumer Information and Workable Competition in Telecommunications Markets," *Journal of Economic Issues* 27, no. 3 (September 1993): 775-91.

Market Share

The first is simply the largest firm's market share. If the largest firm in an industry or market has a "high" market share, and if its share is considerably greater than that of its nearest rival, it can be presumed to have market power. Using such an analysis, a cautious position might be to suspect the existence of market power if the largest firm accounts for more than half the total market.

Lerner Index

The competitive model indicates that the long-run competitive equilibrium will result in prices equal to marginal costs. This is because, whether competitive or not, firms maximize profits by producing the output level at which marginal revenue equals marginal cost, and in competitive markets marginal revenue and price are equal. If a firm has market power and produces the output level for which marginal cost equals marginal revenue, a related measure of market power is the price-cost margin. Specifically, the Lerner Index, L, is given in the expression below. In calculating the Lerner Index, P is the market price, and MC is the firm's long-run marginal cost.

$$L = \frac{P - MC}{P}.$$

$L = 1/\eta$ where η is the absolute value of elasticity of demand for the firm's output. In perfect competition the elasticity of demand for a firm's output approaches infinity; therefore, L will approach zero in competitive markets. A high Lerner Index implies a low elasticity of demand, and some market power. In applying the Lerner Index, since marginal cost is not usually known, a proxy such as average variable cost can be used to approximate L. Note, however, that this approximation will be biased upward because short-run fixed costs are not included in average variable cost.

Profitability or Rate of Return

One of the predictions of the competitive model is that competitive firms in long-run equilibrium will not earn economic profits (that is, they will not earn more than their risk-adjusted cost of capital). Moreover, the competitive model predicts that, absent barriers to entry, high profits will attract competitive entrants and reduce profits to the competitive level. Based on this prediction of the competitive model, if firms are consistently able to earn more than their cost of capital, there may be evidence of market power. Indeed, traditional rate-of-return regulation focused on attempts to determine the competitive cost of capital and allow the utility the opportunity to earn that amount but no more.

Other Measures

There are many potential indicators of monopoly power.[15] Some are developed from theoretical analysis. Others are based on *ad hoc* reasoning in the sense that a simple causal link is established between the existence of competition or monopoly and an observable outcome. Relying too much, however, on *ad hoc* measures may lead to an erroneous conclusion.[16] Public policies that assume that a large market share or a high rate of return is uncontrovertible evidence of monopoly power or predatory behavior may punish the diligent

---

[15] For some discussion of alternative measures see Michael R. Baye and Dennis W. Jansen, "Industry Performance Indices and the Economics of Information: New Perspectives and Caveats," *Review of Industrial Organization* 7, no. 1 (January 1992): 83-90. See also: Alexis Jacquemin, *The New Industrial Organization: Market Forces and Strategic Behavior*, (Cambridge, Massachusetts: The MIT Press, 1987), Chapter 3; Dennis W. Carleton and Jeffrey M. Perloff, *Modern Industrial Organization*, (Glenview, Illinois: Scott, Foresman, 1990), Chapter 12; and Scherer and Ross, *Industrial Market Structure*, Chapter 11.

[16] The fact that the truth of proposition X is sufficient for proposition Y to be true should not lead to the conclusion that if Y is true then X must also be true. Making such a leap is the fallacy of affirming the consequent. Even if we believe that firms with monopoly power tend to earn high rates of return or economic profits, we cannot necessarily accept as true the statement that firms that earn high rates of return therefore have monopoly power; there may be other causes of high rates of return.

and virtuous, as well as the guilty. High rates of return and large market shares are not bad in and of themselves, especially if they result from a firm's ability to be more innovative or efficient than other producers. Appropriate public policy should consider the historical evolution of industries and markets, and action should be taken only against firms that achieve or perpetuate an advantageous position through means that contravene legitimate policy goals.

The historic evolution of competition in a market and the dominant firm's reaction to emerging competition may also provide information as to the competitiveness of the market. If the dominant firm has engaged in anticompetitive practices or exercised its monopoly power in the past, or if it is in a position to impact its rivals' costs, prices, or ability to serve their customers, competition may be more difficult. Unless they have changed their stripes, firms that have a history of attempting to impede competition in the past, either by blockading entry or by engaging in predatory pricing, may be likely to do so in the future.

Competition Facing Local Exchange Companies

The telecommunications market is, in reality, a disparate group of submarkets or segments. The segments may be categorized by class of customer (residential, small business, large business, other communications firms) and type of service (access, switched services, transport, enhanced services, and information services). LECs see themselves as competing with a number of different types of players. In some segments, competition is real, effective, and comes from numerous sources. In other segments competition is not effective, and, although competition probably will increase over time, it would be unwise to act as if competition is an accomplished fact when it is not.

It is useful to remember that there is a difference between the dominant firm having competitors and the market being truly competitive. Moreover, though any hint of competition may unsettle a firm conditioned to being a pure—though regulated—monopolist, having a competitive fringe, which captures small parts of some segments, is not equivalent to

the market being competitive.[17] For example, among the things that are not strong evidence of competition are expressions of "interest" in physical interconnection with central offices by a competitive access provider (CAP) and IXC. Expressing interest is not the same as requesting colocation and being willing to pay for it. Undoubtedly, some such interest will result in actual requests for colocation and interconnection, but effective and true competition can come only from actual players' willingness to make investments and commit themselves to the market.

In general, LECs are the only full service or end-to-end providers in the various markets for telecommunications services and equipment in their service territories. All of a customer's telecommunications needs except interLATA toll service can be provided by the LEC—and the RBOCs desire to enter the interLATA toll market. Although LECs face numerous competitors, each such competitor is a niche player attempting to serve narrow segments of the market. Moreover, the LECs have the most complete set of offerings, and customers purchasing some services or equipment from other providers generally continue to obtain others from a LEC. In addition, some of the LECs' competitors, in providing service to their own customers, also obtain services from the LECs.

There is no doubt that the trend is toward more competition in various telecommunications submarkets, and all the players are wary of each other. Various players seem to be eager to enter each other's markets while jealously guarding their own core markets. For example, newspapers and other information service providers feel threatened by potential offerings of the regional holding companies and have sought protection from them—or alliance with them in some cases. Cable television operators sense a threat from the RHCs' desire to deliver video services, and the RHCs feel threatened by cable systems' potential to offer switched telephone services. Enhanced service providers are on their guard and are considered a threat by the RHCs. CAPs and IXCs are also competitors for certain market segments. In addition, intermodal competition from wireless services pose some threat to

---

[17] For more on competitive fringe models see Horning, et al., *Evaluating the Competitiveness of Telecommunications Markets*, 34-37; Scherer and Ross, *Industrial Market Structure*, 221-226; and Carleton and Perloff, *Modern Industrial Organization*, Chapter 8.

LEC services, and the AT&T/McCaw Cellular combination may raise the level of competition in some segments. Moreover, combinations of players (the U S West/Time Warner affiliation, and the AT&T/McCaw Cellular merger) may create some concern that a few large multimedia firms will dominate the coming "infotainment" age in which previously diverse technologies (such as those embodied in video, telephony, and computers) converge to become variations on a digital transmission theme.

As noted above, there is a difference between having competitors and having a competitive market. Even if competitors are affiliated with large national or international firms, those firms are spreading their resources over multiple geographic areas, and LECs may have a significant home-field advantage. Moreover, opening a market to competition does not immediately eliminate the dominance of a former monopolist. The effective elasticity of supply of a LEC's competitors is important in this context. Absent regulation, LEC price increases will be held in check to the extent that competitors are able to increase their output considerably in response. Unless competitors are able to quickly increase their output, the LEC may not face effective competitive pressure.

Some segments are well along the road toward workable competition. For example, Centrex services face competition from numerous CPE vendors. Some enhanced services are also offered in the context of a competitive market. For example, voice mail service competes with private voice mail, private answering and paging services, PBX voice mail, and customer-owned answering machines. In addition, although substitutability is not perfect, voice mail competes with other services, such as call forwarding, and call waiting. Each of these has advantages and disadvantages as compared with voice mail. Speed calling also faces competition from various sources. However, competition in local access, switching, and transport is neither fully nor ubiquitously developed.

Policy Toward Competition

LECs face competition in some of their market segments; moreover, the degree of competition and the number of competitive segments is likely to increase over time. There are, however, relatively few segments that can now be classified as fully or highly

competitive, especially in light of LECs' current market shares. Competition requires that competitors have sufficient market presence (size and market share), resources (both technical and financial), and commitment to create a workably competitive environment. Competition also requires that there be no unnecessary barriers to market entry and that incumbent firms do not have undue advantages. Finally, competition requires that consumers have sufficient information about the availability of services from various providers and that they seek out the best price-performance combination.

LECs often propose that they be given additional freedom to offer new services and to compete in partially and fully competitive market segments. They should be given such freedom, but they should not be allowed to maintain or to raise barriers to entry, drive competitors from the field, or marginalize them. In order for competition to develop, current and potential competitors must have an opportunity to establish themselves. The following recommendations, if implemented, can help create such an opportunity. After implementing these recommendations, commissions would not control prices to the extent that they have in the past (except for basic monopoly services). Rather, they would act as a referee in the market, creating an environment conducive to competition and efficient production of telecommunications services, enforcing equal access, and nondiscriminatory pricing rules.

Adopt Criteria for Competitive Classifications and Allow LECs To Compete.

Although competition is emerging in many segments of telecommunications, it is too early to classify most segments as fully or highly competitive. Policy should be aimed at promoting competition where possible and efficient, and allowing LECs to compete with entrants.[18]

---

[18] Competition that arises because of anomalies in the regulated pricing structure may be uneconomic in the sense that the competitors are able to offer services at prices above the regulated firm's actual cost of providing the service but below regulated rates. Such a situation may result in what has been called "creamskimming" behavior in which competitors target the most profitable market segments. For more on creamskimming see Alfred E. Kahn, *The Economics of Regulation, Volume II: Institutional Issues* (New York: John Wiley, 1971), 221-246 and William A. Brock and David S. Evans, "Creamskimming," in David S. Evans, ed., *Breaking Up Bell: Essays in Industrial Organization and Regulation* (Amsterdam: North

However, LECs should not be given leeway to foreclose entry or maintain their traditional dominance. Commissions must walk a fine line: they must encourage competitive entry and allow competition to become established, but they must not unduly restrict the LECs' ability to respond.

Thus, LECs would be given sufficient leeway to respond to competition as it emerges or becomes more fully developed. Criteria are proposed in tables 6-1 and 6-2 for telecommunications services sold in partially and fully competitive markets. Because most segments of the telephone business have traditionally been monopolized, proposed criteria are relatively conservative in the sense that markets classified as highly competitive in telecommunications would be considered highly concentrated in other industries.[19] Even for highly competitive services, other classification schemes, such as the Department of Justice merger criteria would indicate that the market was "concentrated" rather than competitive. These criteria are sensitive to the peculiar history and concentration levels found in the telecommunications markets. They should be useful to state regulators as they provide a more robust and realistic set of indicators of the competitive nature of given telecommunications markets.

A service that is not either competitive or partially competitive will have a score that is less than the minimum criteria identified in table 6-1. A partially competitive service will meet the criteria in table 6-1 and competitive services will meet or exceed the criteria in table 6-2. Such a market does not have a viable nonLEC provider and the incumbent LEC has over 70 percent of the market. This classification scheme is workable and should offer protection to nascent competition.

Holland, 1983), 61-94.

[19] The Herfindahl-Hirschman Index (HHI) used in the table is a measure that is used to determine the size structure of a market by looking at the market shares of all firms.

TABLE 6-1

PROPOSED CRITERIA FOR PARTIALLY COMPETITIVE CLASSIFICATION

Partially Competitive Services		
Number of Competitors	*Market Share Criteria*	*Other Criteria*
There must be at least one (but preferably more) viable competitor, unaffiliated with the LEC, which (a) is capable of providing an adequate alternative service and (b) is actively soliciting business throughout the relevant geographic area. The entire geographic area need not be served by the same viable competitor, but each subpart of the area must have at least one such viable competitor.	No service for which the LEC has as much as 70 percent of the market can be considered to be partially competitive. If the HHI can be calculated, it should be no greater than 5,200 for the market to be considered partially competitive.	There must be sufficient evidence, based on acceptance of one or more competitors by the market, that existing customers are choosing the competitors' services over the LEC's offerings and that new customers consider the alternative providers to be effective competitors. Such evidence can be a reduction in the LEC's market share, a reduction in the number of (or percentage of) customers served by the LEC or a reduction in usage of the LEC's service offerings.

Source: Author's construct.

TABLE 6-2

PROPOSED CRITERIA FOR FULLY COMPETITIVE CLASSIFICATION

Fully Competitive Services		
Number of Competitors	*Market Share Criteria*	*Other Criteria*
There must be at least three viable competitors, unaffiliated with the LEC, each of which (1) is capable of providing an adequate alternative service and (2) is actively soliciting business *throughout* the relevant geographic area. The same three competitors need not serve the area uniformly, but the entire area must have at least three active competitors.	No service should be classified as highly competitive so long as the LEC has over 45 percent of the market for that service. Moreover, at least two identifiable competitors must, individually, have at least 10 percent of the market, or one identifiable competitor must have at least 25 percent of the market. If the HHI can be calculated, it should be no greater than 3,000 for the market to be considered highly competitive.	There must be market confirmation that existing customers are choosing the competitors' service over the LEC's offering and that new customers consider the alternative providers to be effective sources of service. Confirmation includes evidence of a reduction in the LEC's market share, number of customers served, or usage of the LEC's services.

Source: Author's construct.

If a commission expects competition to develop to its fullest capability, it should provide the room for it to do so. Classifying services as competitive too quickly may keep competition from developing fully. For instance, under price-cap regulation, the ability of a LEC to practice limited price discrimination through the use of selective price cuts for competitive services is enhanced relative to the situation under traditional regulation.[20]

These criteria are reasonable. Other jurisdictions may have stricter standards, especially with regard to the HHI. The District of Columbia Public Service Commission adopted a test for classifying a service as competitive and, thus, allowing flexible pricing, when the HHI for that service is less than 1,800. Other tests used by the District of Columbia Commission include: (1) whether the service can be duplicated by CPE or some other technology, (2) whether the service is nonessential, and (3) whether the own-price elasticity of demand is high. A service is more likely to be classified as competitive if it passes more of the tests. If (1) or (3) are satisfied, pricing flexibility may be granted without further analysis. Satisfaction of (2) or the HHI test by themselves would not be sufficient.[21]

In implementing such criteria, a commission should maintain flexibility and exercise common sense and judgment in cases where a service does not meet all of the individual criteria. Also, the quantitative measures (HHI and CR) may be difficult to calculate for some services; in that case, it may be sufficient to require a demonstration that there are competitors and that they are successful in attracting a significant customer base and are viewed by customers as providing genuinely competitive options. What is important is that reasonable

---

[20] If price-cap regulation is to be effective, services within a basket should be relatively homogeneous because the more similar—in terms of elasticity of demand—the services are within a basket, the less likely it will be that the regulated firm will practice a form of quasi-Ramsey pricing. That is, if demand elasticities for services in a basket are similar, the firm will not be tempted to raise the prices of some services by the maximum amount allowed under the price-cap plan and lower others by the maximum amount allowed in order to fend off competitors.

[21] See District of Columbia Public Service Commission, Order No. 10147, January 15, 1993, 15-42. Interestingly, the company (Chesapeake and Potomac Telephone Co., a Bell Atlantic subsidiary) argued that satisfaction of *any* of the criteria should be sufficient evidence of a service's competitiveness, and the Office of People's Counsel argued that *all* of the criteria must be met to provide sufficient evidence of competitiveness.

criteria be adopted and that the LEC has a significant burden of proof in demonstrating that a service should be assigned to given a competitive classification.

Implementing these or other criteria is not without problems. There will be some difficulty encountered when attempting to draw clear lines between monopoly, partially competitive, and fully competitive services. The difficulty encountered in drawing clear lines is compounded by the fact that a service may be classified differently depending on geographic point of delivery. Moreover, it appears that the trend toward increasing and expanding competition in various telecommunications segments will continue. Having more-or-less uniform treatment for all services that are not both basic and noncompetitive would greatly simplify administration and decisionmaking.

Plan for Periodic Review of the Status of Competition.

Competition is a dynamic process. Commissions should periodically undertake a thorough review of the nature, strength, and progress of competition in various segments. Competition is not fully developed in most telecommunications segments. Therefore, the commission should continue to view LECs as dominant firms in most of segments and should monitor the progress of competition. If a segment becomes less competitive, the commission should be prepared to change its classification. In addition, the cutoff points in terms of market share, CR, and HHI for competitive classifications should be gradually lowered as competitors increase in number and gain market share.

Allow Geographic Deaveraging but Reserve the Power To Tie Partially Competitive Prices To Fully Competitive Prices for the Same Services.

One rationale for adopting alternative regulation plans is to create incentives that more effectively mimic the operation of a competitive market. If the prices of competitive services fall, other prices should also fall. Therefore, if a service is highly competitive in some geographic areas and partially competitive in other geographic areas, and this condition persists for some time, a commission could tie partially competitive prices to the lowest price

(adjusted for cost differentials) for the same service in highly competitive areas. The price of a service in competitive markets would, thus, be used as a benchmark for the price of the same service in less competitive markets. Price cuts for customers in more competitive areas would result in similar reductions for customers in less competitive areas; this would allow customers in markets that do not attract significant competitive entry to share the benefits of competition.

Enforce Nondiscriminatory Access and Pricing Policies.

Competitors often use LEC facilities and purchase LEC service elements which are then used to supply their customers. Workable competition requires that competitors have access to needed service elements on acceptable terms. Moreover, such access should be on terms equivalent to those given the LEC's affiliated users of the same facilities and service elements.

Restructuring plans, such as those proposed by Ameritech and by Rochester Telephone, may result in giving all users access to the local switch and network services on equal terms. Equal access and nondiscriminatory pricing of such facilities and service elements is an important consideration in promoting competition. Equal access and nondiscriminatory pricing policies (including unbundled pricing of service elements) can be put into place in the context of implementing open network architecture rules. Commissions can act proactively to ensure that competitors have equal access and can enforce and oversee equal access provisions. A promise of equal access is not enough.[22]

Nondiscriminatory access should also be defined to include number portability and the inclusion of nonLEC access customers (including wireless customers if they wish such listings) in the white pages listings. Without number portability and white pages listings,

---

[22] For more information on equal access policies, see John D. Borrows and Robert J. Graniere, *An Open Network Architecture Primer for State Regulators* (Columbus, OH: The National Regulatory Research Institute, November 1991).

customers obtaining access from nonLEC providers face additional costs that may create unnecessary barriers to entry or to broadbased competition.

Conclusion

Regulation of telecommunications utilities has long been justified by the need to have countervailing regulatory power to offset the market power of the incumbent monopolist. At the heart of this proposition is the principle of congruence: noncompetitive markets are to be regulated and competitive markets are not to be regulated. Technological changes have made it possible for smaller and differently-structured telecommunications providers to offer telecommunications services that directly compete with those offered by the incumbent regulated LEC. State and federal pricing and market structure reforms were instituted that allow the incumbent LEC and its challengers to offer partially and fully deregulated telecommunications services. The overall intent was to ensure universal service and to promote the development of competitive telecommunications markets.

Congruence remains an important principle of regulation, especially for services or markets perceived to be or actually in transition to competitive markets. Premature deregulation can strengthen the incumbent LEC and ensure that only fringe competition will develop. Equally, inflexible regulation can also have anticompetitive outcomes where markets are clearly evolving.

Judicious application of the competition measures and criteria presented in this chapter can help state commissions better understand and determine the kind of markets they actually face. In each of the alternative forms of regulation examined or implemented by state commissions a key issue was the classification of services into competitive, partially competitive, and noncompetitive services. Commissions largely apply the principle of congruent regulation once the classifications are established. The tools presented in this chapter allow for an accurate and reliable classification of telecommunications services. It also enables commissioners and senior commission staff to better evaluate the various competition measures and reports submitted by contending parties in regulatory reform proceedings.

CHAPTER 7

DESIGNING A REGULATORY REFORM EVALUATION INDICATOR SYSTEM

Introduction

In chapter 3 basic program evaluation concepts and procedures were identified along with an explanation of how program evaluation fits into the policymaking process. In this chapter the rationale for a program evaluation or a regulatory reform indicator system is presented and an illustrative indicator system is introduced. A regulatory evaluation indicator system (REIS) can be used in conjunction with the evaluation designs discussed in chapter 3.

A REIS is defined as a *method to produce data needed by decisionmakers about the operation and impact of a particular program.* It is not synonymous with a government-wide information system, such as census or tax or other economic data collected and utilized by government. Instead the focus is on a particular program and generally for a particular time period. As previously mentioned, regulators do not always use the term "program," but regulatory reform policy initiatives, such as price caps, flexible pricing, setting QOS standards, eliminating entry barriers, and encouraging the deployment of advanced infrastructure and sophisticated telecommunications services are all examples of regulatory programs designed to achieve certain ends or goals.

An ongoing REIS provides diagnostic information to regulatory decisionmakers about key features of the program. It is not comprehensive or encyclopedic. Instead, the rule is: collect only information known to be useable to the program's decisionmakers. It can, of course, be difficult to completely know ahead of time what information will be useful to decisionmakers and many similar nonregulatory information systems have suffered accordingly when well-intentioned technicians designed information monitoring systems without adequate up-front identification of the information needs of decisionmakers. In these situations the decisionmakers tend to feel overwhelmed and frustrated by what they see as useless data and the program managers and researchers can feel betrayed because their hard-

won data were not used. The key to avoiding this problem is to involve policymakers in the design of a REIS.

Unlike other areas where it is difficult to design, manage, or even understand massive governmental programs, state regulatory pricing reforms are especially well-suited to an evaluation effort. In part this is because state commissions issue official orders or other actions that specify start dates, who is effected, and what actions are to take place. Given the definitive and authoritative nature of the involvement of commissioners in issuing orders, it is appropriate to have the commissioners identify their information needs. Three main ways available to accomplish this are discussed below.

(1) The proceeding can be modified to include a consideration and identification of evaluation information needs. Sign-off by the commissioners in an order or other action would indicate that a consensus existed on the measures to be used to evaluate a pricing reform. An important advantage is that commissioners sign the order or action and in doing so indicate their agreement (within certain constraints) on the measures or indicators chosen.

(2) Included in the order can be a process established, such as a workshop, or hearing, or task force report, that identifies the information and measures needed. This approach allows the commission to focus more fully on the measurement issues without being distracted by the hurly-burly associated with promulgating a regulatory pricing reform. This approach has the added advantage that all features of the program are known. The main disadvantage is that commissioners may not always be kept in the process long enough, given the pressing nature of their other time demands.

This workshop approach, however, is especially well-suited to allow the affected parties to participate in selecting the measures to be used. This is a significant advantage because (if properly carried out) it can result in a REIS data set that commissioners, staff, the utility, other providers, and consumer groups are

comfortable with. Of course, no party would be precluded from gathering other data. The gain is that a consensus is reached on the core data to be gathered and this is done before the adversarial nature of any subsequent review complicates or colors the evaluation process.

(3) A hybrid of the two approaches may actually be the most useful. Here commissioners can indicate in their order the measures they feel are important and other parties can further expand on this set of measures in a workshop-type proceeding after an order is signed.

Selecting a Measure

Four approaches to identifying and selecting the "numbers" to be used as measures for a REIS are presented and analyzed below. Each has the advantage of being significantly superior to ad hoc data, adversarial data, and unsolicited data. Ad hoc data are data gathered after the fact and represent a commonly used approach. Typically its use implies that available data were selectively mined in order to obtain some information. This approach tends to argue that decisionmakers must make use of "the best available data," even when the data are not very good nor on target. Adversarial data have many of the same characteristics and are submitted by one of the parties in order to prove something. The data are suspect because only data and assumptions supporting the party's position are advanced. The regulatory challenge then becomes one of disaggregating and unraveling the data sufficiently so that biases can be discerned. This is time consuming and eats up resources that could be more productively employed elsewhere. Unsolicited data may include features of both preceding approaches. Its core characteristic is that the data submitted may not measure items of interest to commissioners and may define issues because of the availability of the data rather than because important things are being measured.

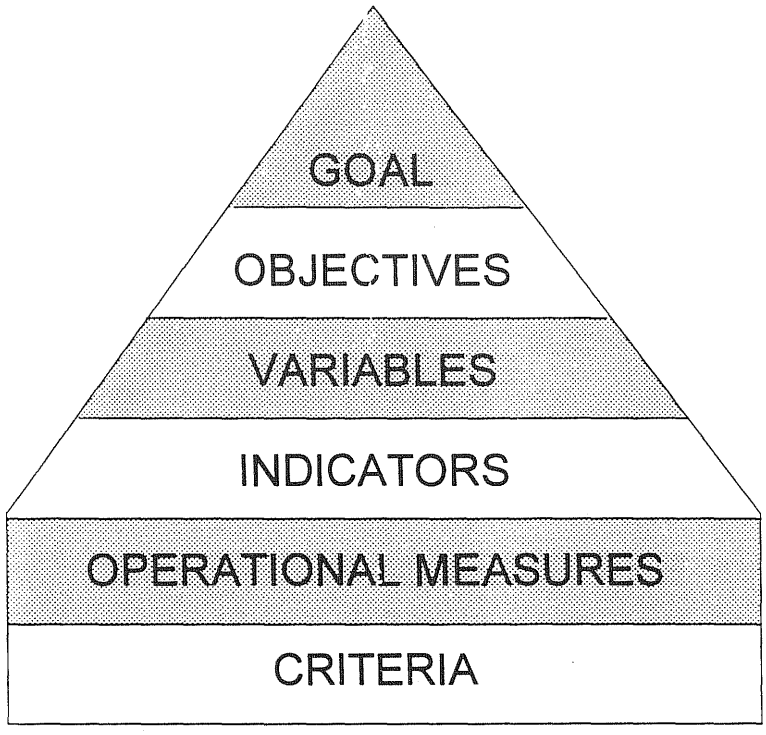
Goal-Defined Measures

A standard way to derive measures for any management information system or evaluation is to use the program goals. As depicted below this is done through a six-step process.

	Steps	Example
Step one :	Identify goals	To ensure affordable universal service
Step two:	Identify objectives	To provide affordable service in rural areas of state
Step three:	Identify variable(s)	Rural single-party service
Step four:	Select indicator(s)	Rural demand for single-party service
Step five:	Select operational measure	Ratio of requests for rural single-party service to all rural residential service requests
Step six:	Define criteria to evaluate measure	Ratio will increase by 10 percent

In the above example a broad goal is operationalized by disaggregating the goal and specifying in ever increasing detail how to measure goal achievement. The universal service goal in practice would have more than one objective and each objective would have multiple variables and indicators. As shown in Figure 7-1, a goal can best be envisioned in this six-step approach as being on the top of a pyramid, supported by a larger number of objectives, variables, and measures. Measuring a goal solely by reliance on a single measure would be inappropriate. Without over extending the analogy, the pyramid has great strength because it is supported by multiple measures, each having a posited relationship to the goal. A goal measured by one indicator would not have the same stability and reliability.

The advantage of this approach is that an up front agreement is reached on the dimensions of the goal and the way to measure progress. The sixth step establishes criteria so that the numbers produced can be evaluated. Without this last step, all parties will have to debate what degree of change in a measure indicates success.



Source: Authors construct.

Fig. 7-1. Depiction of relationship of goals and operational measures.

The main problem with the approach is that regulatory decisionmakers do not always explicitly identify the complete set of regulatory goals affected by their actions. Further, it is time consuming to engage in the six-step process. However, furtherance of regulatory goals is a lode stone against which state commissioners routinely formulate policies and make decisions. Lack of clear goal statements is more an artifact of the minimalistic nature of the formal documents produced rather than to any lack of concern. Often goal statements in legislation, or prior orders, or proceedings are implicitly included or assumed. Use of a workshop-type approach is an effective way to overcome any initial lack of goal specificity. It also lets other parties participate in specifying the goals and the operational measures.

Linkage-Defined Measures

Some of the same advantages occurring under the goal approach can be accomplished in a linkage approach. Here the decisionmakers and the affected parties agree in advance to link specific features of a reform to specific operational measures. The goal-based approach, by way of comparison, is silent on the actual actions taken to achieve a goal. The linkage approach takes a very pragmatic view and (ideally) would get the parties to match their rhetoric to predicted outcomes. Causality is less important here than having a party say that "pricing flexibility will increase the number of service offerings." Accountability for ones' rhetoric is an underlying thesis. This is particularly important because this approach can help expose rosey projections and unrealistic imagery at the policy formulation stage and can help to define standards that will be used to evaluate the success of a reform.

A de facto and ad hoc linkage is what can often happen when linkages are specified for the first time in an adversarial evaluation proceeding. The parties at this stage are frustrated and imprisoned by the need to rely on available data. Specifying linkages in the beginning allows for data collection routines to be designed. While it would be ideal for the parties to specify criteria, such as "how many new service offerings" may be expected to occur with increased pricing flexibility, it may be unrealistic to expect sophisticated practitioners to limit their freedom by setting explicit tests. A commission may elect subsequently to set criteria itself, or to let "the data speak for itself" in a hearing process.

The linkage process can occur at the order-writing stage or in a workshop. Multiple measures can be used or the parties can agree to rely on a single measure. As with the goal approach, it is important that the measures be expressed operationally. This will minimize subsequent debates over alternative ways to measure variables.

Impact-Specified Measures

The same line of reasoning as above is applied here, except the focus is on impact. The distinction is a fine one, which attempts to distinguish between actions taken and results observed. In the above example "the number of new services introduced" is important, but it

is an action, perhaps even an intermediate or linked action. It shows that the utility acted by introducing new services. It does not show the impact, that is, how many customers used it, or whether the delivery of health care and other services was improved, or whether the economy improved, or jobs were created. Impact measures attempt to answer the "so what question," rather than "what did we do?"

Specifying impacts is more difficult than either the linkage or goal approaches. The measurement issues are not particularly more difficult. Instead, the difficulty occurs in trying to chain or model actions and to predict impacts. Transactions costs go up accordingly, and this approach may turn out to be more time consuming.

Red Flag Approach

This approach could also be called a criteria approach. It sets criteria that guard against failure and send up red flags when a prespecified level is passed for a particular REIS measure. Examples could include household penetration rates below 93 percent, call blocking greater than .1 percent, initial service deployments in rural areas that are not within three years of initial urban deployments, and a decrease greater than 10 percent in actual versus planned infrastructure deployment.

The intent here is not to be comprehensive but rather to set action zones with "indicators as trip wires." When a wire is tripped a red flag goes up and commission oversight routines are activated. This mode is appropriate in a hands-off minimal regulation approach. Goal attainment is less important than failure avoidance. A competitive market (ideally) is allowed to develop and regulatory intervention is called for only when a red flag goes up. In this instance, if household telephone penetration fell below 93 percent, the commission would investigate and take any appropriate action, including the option of saying that an investigation revealed that no further action was required.

The main advantage of this approach is that the REIS measures only what is important. The major problem is that it produces little information about positive trends and may place a commission in a crisis response mode.

Of the four approaches examined, the goal and linkage approaches are the two approaches most easily adaptable to commission needs. The impact and red flag approaches require more and better information than may be reasonably available. That is, they need a lot of work to design causal models, something that may require significant research resources. The goal approach follows a business logic that may be familiar to many commission staff and, and accordingly, be much easier to implement. The linkage approach requires a little more daring, but it may have higher payoff because it focuses exclusively on important indicators and how these indicators are explicitly linked back to a particular regulatory reform, say flexible pricing. In the rest of this chapter and in the sample REIS developed, a goal-based approach will be used. With a moderate effort a REIS could be modified to handle a linkage approach.

Some reluctance to build a REIS can be traced to a "Don't they kill the bearer of bad news?" syndrome. That problem needs to be straight forwardly discussed as some have a pragmatic concern that systems that produce information are inherently dangerous because information on any failures can be used to hurt the program managers and any associated decisionmakers.

The underlying beliefs here are that mistakes are best swept under the rug or should be dealt with by "smoke and mirrors." By not having an information system, it could be argued, no shortcomings need be acknowledged.

It would be easy from an ivory tower, academic perspective, to refute these ideas by urging that all who work in the service of the public should welcome evaluative data. In practice, however, real people often have their professional careers tied in many ways both to the initiation and implementation of the reform. For some, the fear of criticism may be enough to reject an evaluative approach.

Commissioners and their senior staffs have public stewardship obligations and oversight responsibilities that require feedback information. Traditionally, advocacy, investigatory, and various report filing approaches have been used by state commissions to carry out these obligations. More recently commissions have used alternative dispute resolution techniques and workshops to gather the information they need to responsibly carry

out their duties. The thesis of this report is that the program evaluation approach can produce more accurate information and do it less expensively than other approaches.

Further, inherent in the program evaluation approach is a diagnostic perspective that envisions correcting potential problems before they reach unmanageable proportions. Ideally, a REIS would allow the commission to recognize any potential problems early on and allow more lead time for additional monitoring, research, or ameorilative action. In an adversarial proceeding (especially one with time limits) usually enough time does not exist to adequately understand or test submitted data because the process is "end-loaded" by the natural way all parties deal with deadlines. In contrast, an evaluative approach supported by a REIS allows constant monitoring by the commission (and any one the commission elects to share data with) early on in a "pilot" or test period. Like good wine, data do improve over time as the richness and complexities hiding behind the numbers become revealed and understood by the parties and the commission. The thesis here is that a good REIS can keep a commission from being blindsided or being the "last one to know" about a potential problem. Stated positively, with REIS a commission can be the first to know, announce, or celebrate successes achieved under the regulatory pricing reform. Either way, the public is well-served by the ability of the commission to monitor developments early, effectively, and efficiently.

Data

Another positive benefit of this approach is that a well-designed REIS can eliminate or greatly reduce some of the contentious debate that often surrounds data or studies presented to the commission. If the design of a REIS and the evaluation process takes place in an open and partipatory setting, a greater legitimacy and familiarity and confidence can be had with the REIS data. Contending parties can then debate about the meaning of the data rather than about the reliability or validity of the data, or the techniques used to generate the data.

Illustrative Reform Evaluation Information System (REIS)

In this example a goal-based approach will be used for a price-cap state. The same six steps identified above will be followed, and will indicate how to move from a goal statement to a REIS number that can be readily and confidently used by commissioners and senior commission staff for decisionmaking purposes.

Step One: Goals

Examination of the price-caps orders and associated documents reveals at least twelve goals that are frequently mentioned. Few price-caps orders actually describe and delineate the exact goals to be achieved by the price-caps reform or its specific features. This kind of information is often found in the supporting legislation or in the proceedings of workshops, or rule setting procedures. Goal statements discovered may not conform to academic or business school standards in terms of format or clarity. However, they do provide meaningful guidance in terms of the intentions of the legislators or commission policymakers. Generally commissions have many of the following goals in common in their price-caps reforms.

- to improve quality of service;
- to cap prices by a formula;
- to have increased consumer savings;
- to allow the utility to earn a fair profit;
- to promote deployment of an advanced infrastructure;
- to establish conditions favorable to competition;
- to allow pricing flexibility;
- to establish conditions for consideration of exogenous adjustments;
- to promote economic development;

- to improve the effectiveness of the delivery of governmental, health, educational, business, and public safety services; and

- to ensure that disabled citizens receive telecommunications services.

These goal statements are necessarily quite broad and reflect the diverse nature of the needs to which commissions respond. Nonetheless, these goal statements are helpful because they do provide guidance about the intent of the commission which can be used to set up a REIS. If a state chose only six of the above goals, it would not be necessary to design a REIS to produce data in all twelve goal areas. In practice, most states do not have clearly expressed goals that cover all twelve goal areas. Further, a state might have an emphasis, say, on economic development and have a separate rural economic development goal in addition to a more general state-wide economic development goal.

It is a common failing to have technicians attempt to design a comprehensive evaluation system: that is, to deal with all twelve goals when the commission is concerned with just six main goals. In addition to consuming a disproportionate amount of resources, comprehensive REIS can collapse under its own weight: they require such massive data gathering on items not always important to the commission and die from overexertion and nonuse. This type of failure can largely be avoided by a discipline that only uses goals identified by the commission or legislature.

The reform process is usually characterized by consensus and coalition building. This can produce a laundry list of goals that require affirmation, and create a situation where every goal is not of the same importance or capable of credible linkage to any particular feature of the price-cap reform. Most of this problem can be resolved by priority setting or agreeing that only a subset of goals will be included in REIS. The most important test is what goal performance information will commissioners need to make a decision on continuing, modifying, or dropping the price-cap reform. The exact set of goals does not need to be completely known in advance, rather, this method is effective in allowing a commission to avoid collecting data it is unlikely to use by deliberately excluding information known not to be needed for decisionmaking purposes.

Step Two: Objectives

In the above list of goals, the QOS goal will be the one goal followed up on in detail for the purpose of designing this sample REIS. Recall the pyramid notion that has each goal being supported by one or more objectives. Here the goal to "increase quality of service" is disaggregated and made more specific. This step is necessary because the goal statement is broad and could include disparate items, such as conformance to official standards, technology diffusion, repair, service to the disabled, interconnection, and complaint handling. One or more objectives could be prepared for each of these items. In the present example, the "conformance to official standards" aspect is translated into the following objective: to have a QOS higher than the previous year.

Step Three: Variables

A number of ways exist to think about improving the level of service quality. Complaints, certification by commission staff, surveys of clients, and engineering models could be used to conceptualize and begin to measure service quality. In this example, because price caps is a reform and the goal is to increase service quality, a reasonable variable to use is " price-adjusted quality of service". The thought here is that the customers should be better off under the price-caps reform than they were under, say, ratebase, rate-of-reform regulation. One way to do this is to say that the concern is that the amount of quality received for a given price should be greater than that received the previous year. A good part of the rationale for this is that the efficiency-increasing infrastructure investments that must be made by an economically rational utility under price caps will automatically increase the QOS because of the increased capabilities and reliability of the new technologies. Common sense and the history of innovation confirm that a new technology will not replace an existing technology, if the new technology is less reliable than the existing technology. Accordingly, a variable defined as "price-adjusted quality-of-service" is appropriate.

Step Four: Indicators

Selection of a variable narrows the focus considerably but does not quite say exactly how to measure the quality. Again, several measures could be used, but the one selected here is "the ratio of service quality to price." This measure affirms interest in service quality and price and focuses on a ratio measure.

Step Five: Operational Measures

The above ratio needs to be defined by the actual data that will be used to measure or operationalize the indicator. At this step it is necessary to say exactly what piece of data will be used in the numerator and what will be used in the denominator of the ratio. Specifying this in advance of any evaluative analysis helps avoid having the analyst "peek at the data" and adjust the data to achieve particular ends. By making the operational measurement decision upfront, a more dispassionate and analytical decision can be made. If the process is participatory the measurement chosen could have additional validity directly proportional to the number and variety of participants involved.

The operational definition could be as follows:

quality-of-service	the percent of utility surveyed respondents indicating that service quality was B+ or better (when using a standard A-F grading approach)
price	the price-cap inflation index

The ratio would show that quality would have to improve at least as much as the price-cap index in order for the ratio to stay the same between any two years. To "improve" quality would have to increases greater that the change in the price-cap index. Thus, year one might have a QOS of 89 percent and a base-price index of 100, and result in a score of 89 (89/100=89). In year two if the score was greater than 89, then the quality could be said to have been improved. A score of 91 for QOS satisfaction and a price index of 101 in year

two would produce a score of 90 (91/101= 90) and would indicate an increase in the price-adjusted QOS. Scores of 85/105 wold become a net 84 and would show a decrease in price-adjusted quality.

Step Six: Criteria

The final step is optional, but very useful, and involves saying in advance what level of change is necessary to score or indicate as an increase. This step is especially helpful if the reform advocates paint a picture that has a significant increase (or decrease) envisioned. Of course, a decision can also be made that "any increase is an increase" and this is fairly typical. In the example here the criteria is a "three percent increase in the price-adjusted quality ratio." If the increase is 3 percent or greater, then the threshold is passed and the REIS measure indicates success. If the score is less than the agreed upon 3 percent, then corrective action or more careful monitoring may be required.

Conclusion

As illustrated in this chapter, a regulatory reform evaluation indicator system can be an effective and efficient way for state commissions to exercise their oversight responsibilities. Regulatory pricing reforms are especially appropriate subjects for a regulatory evaluation system because they represent discrete breaks with past pricing practices and specify the new pricing policies in detail. As a reform, the intent is that better results, new and positive impacts should occur that represent an improvement over the form of traditional regulation replaced by the reform. Existing commission data collection and reporting routines may be insufficient to monitor these hoped for positive improvements. Furthermore, data submitted by interested parties is inherently suspect and requires some effort to decode and analyze. A consensually defined REIS is superior to other approaches in providing a commission with needed, low cost, objective, and timely information on the impact of a pricing reform on state regulatory goals.

CHAPTER 8

SUMMARY AND CONCLUSIONS

Summary

It was the intent of this report to analyze and present different approaches to measuring the impact of alternative regulatory pricing reforms. In particular three major approaches were examined—program evaluation, qualitative approaches, and economic analysis—for their appropriateness and usefulness. The strengths and weaknesses of each approach were identified along with an appraisal of the regulatory context.

This report responds to an expressed need for an objective appraisal of the suitability of various approaches without having the appraisal being embedded in the context of a specific regulatory proceeding. Often by the time that regulatory staff become uneasy about a report, measure, or data set submitted in the course of a pricing reform proceeding it is too late to do anything about it. Also decisions at certain stages of a proceeding can be perceived as advantaging one party at the expense of another.

Further, a number of state commissions have a regulatory pricing reform in effect and must make a decision about renewing, revising, or discarding the pricing innovation. Other states may not necessarily have a formal appraisal scheduled but may still feel the need for an evaluation either as a part of their general oversight duties or because sufficient time has passed to produce meaningful data. Additionally, nearly all states with pricing reforms have expressed the view that the reforms are intended for the transitional period between full regulation and workable or full competition.

The report also presents a description and analysis of the current pricing reforms examined or implemented by state commissions. The wide range of reforms affirms the role of states as change agents searching for the right pricing programs for transitional telecommunications markets.

Program Evaluation

The above circumstances are particularly appropriate for an approach, such as program evaluation, which is somewhat curiously not as widely used by state commissions as one might expect. Program evaluation borrows from the classic scientific experimental design model and has been pragmatically modified by its practitioners to meet the real world needs of policymakers. Commissions and commission staff seem to have a decisionmaking perspective that fits well with the use of program evaluation. That is, they have a concern about impact, cost, and the ability of the pricing reform to do what its advocates claim. Dogmatic defense of particular regulatory tools, processes, or investigatory techniques is not a style commonly in evidence at state commissions.

The program evaluation tools and concepts presented in this report are nearly able to be used "right off the shelf." They build on business school-type goal analyses and other research techniques that commission staff are familiar with. Most commission staff have not used these techniques in the integrative manner made possible by program evaluation. Because commission orders specify dates, actions to be undertaken, geographical area, and the affected providers, the regulatory arena is nearly an ideal candidate for programs evaluation. Unlike some large scale government programs—such as welfare, health, crime prevention, job training, or education—a commission can say clearly and authoritatively what program (e.g., a pricing reform) is to be adopted, by who (the regulated utility), for what time period, and the characteristics of the reform. While utility management cannot be viewed as robotically carrying out commission policies, it is fair to say that they implement policies in a uniform and consistent manner not always found in other sectors. This allows program evaluation efforts to focus more on results, rather than on implementation and goal clarification analyses that so often plague health, job training, or economic development reforms.

Qualitative Analysis

Regulatory pricing reform proceedings, and other commission proceedings, have a wide range of information submitted in various forms by the affected parties. The basic training of most regulatory analysis is generally in quantitative research methods. Paradoxically, a large amount of the information that is submitted to commission is not quantitative. Most analysts have not received formal training in qualitative methods and find that commissioners are often very comfortable with qualitative analyses. In this environment, commission analysts learn to use and accept qualitative information.

This report lays out the qualitative analysis approach in a systematic manner and presents specific examples of qualitative analytical tools that can be used by regulators. This approach puts testimonials, verbal scenarios, and many nonquantitative assertions or propositions in an analytical context useful to regulators.

Systematic qualitative analysis allow the researcher to have a better insight into the nature of the pricing reform or its impacts. Quantitative analyses are superior in that the explicit methods allow unbiased researchers to replicate and assess any given quantitative study. The pinpoint and explicit results possible from a quantitative analysis, however, may not permit the analyst to fully understand the larger context. It may be the case, for example, that it is not practical (or even feasible) to empirically model and forecast the job growth possible when price caps are adopted. Yet, testimonials from firms in a state with price caps about job growth—however suspect using quantitative standards—can be analyzed qualitatively to mine for insights and understanding about the dynamics of job creation. Qualitative analysis does not mean that "soft data" are accepted uncritically; instead rigorous techniques are available through which qualitative studies can be evaluated and appraised. Analyses of actual or potential regulatory reforms can benefit from having a full tool kit that includes both quantitative and qualitative tools.

Economic Perspective

Over the past twenty-five years economic analysis has been the lynch pin of regulatory analysis at state and federal commissions. Applied economic analysis of pricing alternatives, costing approaches, competition, and market structure have had a profound and prevailing impact on the regulatory policies adopted by state and federal regulatory commissions.

This report presents an economic perspective on how to analyze regulatory telecommunications pricing reforms. It does not attempt to duplicate or even summarize the massive (and relevant) economic literature available. Instead, certain techniques and concepts were selected and analyzed in terms of their suitability in a regulatory context. These economic techniques and concepts are discussed generically and are not directly linked to specific utilities or states. This helps avoid partisan interpretations of the critiques and focuses discussion where it should be: on the ability to analyze regulatory pricing reforms. Additionally, it is thought that basic, robust and simple analytical benchmarks are more useful than complex, situation-dependent, and sophisticated benchmarks. It is more important to know, for instance, that key (and standard) economic variables are not included in one party's study, than to know how to deal with third-order derivatives and multicollinearity adjustments.

An integral part of most regulatory reforms has been the conceptualization and measurement of actual or potential competition in telecommunications markets. This report contains a framework in which to place analyses about competition, as well as providing specific measures that can be used in assessing the extent of competition. In order to deal with the structural peculiarities of current telecommunications markets—namely the presence of a large incumbent provider—some standard economic criteria have been modified in ways that recognize the realities of the current market structure, yet permit early classification of services as partially or workably competitive.

Regulatory Evaluation Indicator System

A sample REIS that state commissions can use to evaluate regulatory pricing reforms was presented. This REIS allows a commission to carry out an evaluation when resources are limited and when the use of control groups is not always feasible. The REIS, of course, can also be profitably used when a full-scale program evaluation effort is planned.

Considering Alternative Regulation

Although the criticisms of traditional regulation by an influential group of economists are well-known, it is not at all certain that regulatory reform would have arrived without the push of technological forces and the competitive entry that they made possible. The forces of competition cannot be ignored. Few regulatory or technical barriers are insurmountable. If technology makes competitive entry possible, individuals will find a way to compete. Indeed, some analysts now believe that technological advancement and the competition that it makes possible eventually will lead to complete or nearly complete elimination of telephone regulation. Under such scenarios, competition, entering telephone markets from all directions, would obviate the need for much, if not all, of traditional regulatory oversight. Once technological forces altered the cost structure of telephone services and made possible both new applications of telephony and competitive entry, regulators and legislators began to critically review both current regulatory theories and methods and the various alternative models of regulation (including deregulation). Several caveats are in order that regulatory analysts should follow in evaluating alternative pricing plans. Each of these caveats is a "caution light" or an "analytical red flag" that analysts should recognize when critiquing a study or carrying out their own evaluation.

Avoid Unrealistic Comparisons.

When considering alternative regulation plans legislators and regulators should simultaneously avoid applying unrealistic or unfair standards of comparison and avoid being complacent. Unrealistic or unfair standards of comparison should be avoided because it is not constructive to declare traditional regulation to have failed due to the evidence that the results obtained diverge from an envisioned standard of perfection. It is unlikely that any real-world regulatory method simultaneously will satisfy all the economic and social goals imposed on regulation, especially given the inherent conflict among some of the goals and the lack of a clear or consistent ordinal ranking of their importance.

Moreover, unless regulated markets are highly contestable, an unregulated market process will not produce optimal outcomes either. Whatever the form of regulation, the players (regulated firms, unregulated firms, customers) will adapt to the constraints imposed and optimize their individual positions, given the "rules of the game." Therefore, it is inappropriate to compare imperfect regulatory outcomes with outcomes of perfect, but unattainable, market processes. What is needed is evaluation that compares admittedly imperfect methods of regulation with each other and with the imperfect market mechanisms, which are the real alternatives to traditional regulation. For example, in various forms price-cap regulation is one of the dominant forms of alternative regulation.

Price-cap regulation is reputed to have many positive aspects but it may be somewhat oversold. It is claimed that price caps stabilize prices, promote competitive entry, and promote economic development. The validity of these claims has yet to be proved. Price caps may keep some prices from falling due to normal forces, including productivity increases and technological advances. Price caps may hinder competition if the baskets include competitive and monopoly services and allow flexibility for the utility to engage in anticompetitive pricing. In addition, price caps may not be required to promote economic development (as discussed below). Price-cap regulation (or any incentive regulation) and economic development are separable issues—although infrastructure deployment plans, which may aid economic development, are sometimes included in alternative regulation plans.

Avoid Complacency.

Complacency should likewise be avoided. Acceptance of the view that no method of regulation simultaneously can satisfy all the goals of regulation should not lead to the conclusion that traditional regulation should be discontinued or continued. Traditional regulation should not be maintained merely because regulators and utilities understand it or because of the "it's better to deal with the devil you know" phenomenon. Regulators and legislators should strive to improve traditional regulation whenever possible. They should also be willing to consider alternative methods when necessary, provided that there is credible evidence or a belief that, on balance, the alternative methods are preferable.

Avoid Confusing Alternative Regulation with Optimal Regulation.

Regulators need to be mindful that none of the various alternative regulation schemes that have been implemented can be described as "optimal regulation" in the economists' jargon. Although the alternatives suggested purport to deal with one or more of the shortcomings of traditional regulation, all such schemes involve departures from optimality.[1] Moreover, the effects of the departures have not been analyzed as thoroughly as the effects of traditional regulation.

Avoid Confusing Alternative Regulation with Deregulation.

Although some analysts and LECs may view alternative regulation as an interim step or way station on the journey toward total deregulation of telephony; alternative regulation is not deregulation. Rather, it is an attempt to modify existing forms of regulation or create new forms of regulation in order to meet changing needs. Unless and until competitive forces

---

[1] For discussions of optimal regulation see Kenneth E. Train, *Optimal Regulation: The Economic Theory of Natural Monopoly* (Cambridge, Massachusetts: The MIT Press, 1991) and Jean-Jacques Laffont and Jean Tirole, *A Theory of Incentives in Procurement and Regulation* (Cambridge, Massachusetts: The MIT Press, 1993).

eliminate the need for regulatory oversight to control potential abuse of market power, some form of social control is necessary. So long as telecommunications markets are not competitive or contestable, regulatory oversight is needed, though the form of that oversight may change.

Recognize the Cost of Fairness.

Regulators must also realize that, though equity and fairness are important, there is a cost to being fair. Being fair to favored users will create distortions and pressure to ameliorate them. Moreover, distortions that can be maintained in a static, natural monopoly environment may not be sustainable in potentially competitive markets. Alternative regulation almost certainly means that regulators will lose some, if not all, of their ability to create internal subsidies. Costs will be disaggregated and deaveraged. Some consumers will gain, others will lose. Low-cost competitors may take over some segments. Regulators and legislators may have to become more explicit in devising subsidy schemes and inducing new entrants to participate in them, since the protected monopoly will no longer be able to collect and administer such subsidies.[2]

Consider Alternative Regulation and Infrastructure Investment Separately.

Regulators should consider the issues of alternative regulation and advanced infrastructure deployment as largely separable, rather than as linked in some form of *quid pro quo*. Alternative regulation should be adopted if and when it makes sense and is more efficacious in meeting public goals than traditional regulation. Similarly, advanced infrastructure deployment should make business sense for the telephone companies and other firms currently or potentially involved (cable television systems, electric utilities, CAPs, and others).

---

[2] A discussion of the role of regulation in creating and administering a system of internal taxes and subsidies is contained in Richard A. Posner, "Taxation by Regulation," *The Bell Journal of Economics and Management Science* 2, no. 1 (Spring 1971): 22-50.

Regulators should consider carefully the true value of the educational and medical services and hookups that are being offered as sweeteners in the LECs alternative regulation proposals. Although it is difficult to be "against" upgrading the capabilities of the educational and healthcare systems, regulators and legislators would be well-advised to avoid striking a deal while blinded by the sparkling promise of telemedicine and distance learning.

There are three potential problems with making implicit linkages between adoption of alternative regulation and advanced infrastructure deployment. First, such linkages could result in granting a reward (alternative regulation) to utilities to induce them to make investments that their own business interests would have them do anyway, with or without adoption of alternative regulation. Second, such linkages could lead utilities to make investments that are not needed or provide services for which there is not sufficient demand. Third, such linkages could result in government choosing the technology and the winners in the competitive race to deploy advanced services. This would run counter to the goal of creating more effective competition, and it is not clear that government is able to or should choose the best technology.

Evolve Alternative Regulation and Plan for Its Evaluation.

Alternative regulation does have potential benefits as compared with traditional regulation, and regulators should be willing to adopt plans that are likely to improve on traditional regulation and capture those benefits. However, the net effect of individual alternative regulation plans is uncertain. Therefore, regulators should treat the implementation of such plans as experiments and plan to evaluate the outcomes after sufficient time has elapsed to allow for reasoned analysis. Moreover, finding the "optimal" plan for a given state may require an iterative process, so regulators should monitor the results of alternative regulation experiments in other jurisdictions and be willing to revise alternative regulation plans to better meet public policy goals.

Be Circumspect.

Regulators and legislators should not allow themselves to be overwhelmed by the rhetorical din created by those who never viewed regulation as having much merit and who now use arguments based on infrastructure investment, economic development, international competitiveness, and the benefits of competition to call for deregulation or alternative regulation.[3] Regulators and legislators should circumspectly consider proposals for alternative regulation. Such consideration should include clear examination of questions such as the rationale for or goals of regulation (Why regulate at all?), the appropriate scope of regulation (What to regulate?), and the extent and method of regulation (How to regulate?). Ultimately, the appropriate regulatory regime is the one that best meets the chosen goals.[4]

Why Should Telephony Be Regulated At All?

It is difficult to design or evaluate regulatory regimes unless the objectives or goals of regulation are reasonably well-stated. Regulation generally has several (possibly overlapping and/or conflicting) broad goals or missions including the following:

- to promote progress toward social goals, such as universal service and/or rapid and ubiquitous deployment of advanced capabilities;

---

[3] The potential for alternative regulation to promote economic development by speeding infrastructure investment may not be as great as the advocates believe. Although each state may believe that it can pull itself past its neighbors using telecommunications as a bootstrap, it is not likely that such a result is sustainable. All states are likely to have relatively similar infrastructure when the dust settles. Not everyone can be in "first place."

[4] A recent multi-country comparative study of these issues may be found in Michael Tyler and Susan Bednarczyk, "Regulatory Institutions and Processes in Telecommunications: An International Study of Alternatives," *Telecommunications Policy* 17, no. 9 (December 1993): 650-676. Some discussion of the various roles of regulators may be found in Edwin A. Rosenberg, John D. Borrows, Carl E. Hunt, Rohan Samarajiva, and William P. Pollard, *Regional Telephone Holding Companies: Structures, Affiliate Transactions, and Regulatory Options* (Columbus, Ohio: The National Regulatory Research Institute, March 1993), 109-42.

- to protect the interests of consumers and/or competitors by limiting the ability of the dominant provider to exploit its market power; and
- to foster innovation and competitive entry into the marketplace by creating a level playing field.

What Parts of Telephony Should Be Regulated?

Given the goals of regulation, regulators must choose what they will and what they will not control. In some jurisdictions regulators may have considerable freedom of choice; in other jurisdictions they may be constrained by legislative mandates that either preclude or require regulation of certain activities. Items under regulatory control may include the following:

- market entry and exit through franchises or licensing;
- technical standards for equipment, the terms, conditions, and quality of service, and interconnection and interoperability among networks of competing carriers;
- accounting, record keeping, and reporting of operational and financial data;
- investment and planning decisions of producers;
- new service offerings and old service deletions; and
- prices, either specifically or generally.

How Should Telephony Be Regulated?

Given the choices of what to regulate, the mode or method of regulation must still be determined. Regulators may choose to use direct or indirect methods of control. For example, price-cap regulation is an indirect method of control. In addition, regulators may set general or overall goals or standards for utility performance or QOS, or they may set detailed performance and QOS requirements.

Regulators may choose to exercise their oversight continuously, periodically, or only as needed. Regulators may choose to let market forces determine outcomes, stepping in only to investigate complaints made customers or by competitors.[5] Regulators may also elect to apply different degrees of regulation to different firms. Examples include the distinction made between regulation of dominant and nondominant firms and some jurisdictions' streamlined regulation of smaller LECs.[6] In addition, combinations of methods may be used depending upon circumstance.

Conclusions

Much of impetus for alternative regulation is a response to competitive pressures that have been building for some time. Therefore, some of the impact of alternative regulation will reflect the elimination or reduction of numerous pricing distortions that have developed during the previous decades.

The combination of greater pricing flexibility for the LECs and opening markets to competitive entry will result in a rebalancing of prices. Prices will tend to become more reflective of costs, with the prices of more competitive services being driven relatively closer to costs than less competitive services—because the effect of alternative regulation plans depends on the conditions in the market at the time of implementation. Customers in markets that are more competitive are likely to benefit relatively more from the implementation of alternative regulation. However, since the degree of competition will vary geographically and

---

[5] This mode is similar to the role of the courts in antitrust enforcement.

[6] Dominant firms (incumbents) may be subject to stricter rules of behavior and/or more stringent record keeping and reporting requirements than their nondominant rivals. The dominant firm may be more limited than its rivals in its ability to offer new services or change prices of existing services. Dominant firms may also be burdened with "provider of last resort" and "common carriage" obligations that are not imposed on rivals. In addition, a dominant firm may also be required to provide its rivals with access to its facilities without receiving reciprocity. See John R. Haring and Dennis L. Weisman, "Dominance, Non-Dominance and the Public Interest in Telecommunications Regulation," *Telecommunications Policy* 17, no. 2 (March 1993): 98-106.

by product, no single alternative regulation plan is likely to be correct in all jurisdictions or for all market segments.

The process of bringing prices more in line with costs will include considerable pressure to deaverage rates geographically and to unbundle access and usage charges, even for residential customers. As local access competition emerges, the differential between business and residential access charges will decline. Furthermore, there will be pressure on local access charges to reduce or eliminate the use of value-of-service pricing in which customers in urban exchanges pay more for access even though their average and marginal costs may be lower than those in rural exchanges. The system of internal subsidies created to promote universal service will come under increasing pressure as intraLATA toll competition develops. Local access rates will tend to rise relative to toll rates. Moreover, toll rates will be deaveraged and toll message-minute mile rates will be inversely related to route density. Expanded extended area service service can be seen as an attempt to make intraLATA toll competition less entry less desirable.

In a simple analysis of the relative distributional consequences of these effects, some preliminary statements may be made. Business customers will likely benefit relatively more than residential customers; urban customers will benefit relatively more than rural customers; and toll usage will benefit relatively more than local usage. The consequences for the LECs, their rivals, and for regulators is less certain. Creative and aggressive LECs may prosper; others may lose business to new entrants. Entrants that can provide service efficiently will prosper; those that exploited arbitrage and pricing anomalies to provide uneconomic bypass services will find the LECs to be tough competitors.

Regulators' focus will change as well; they may become more like antitrust referees, policing access and interconnection rules and pricing rather than setting the prices of individual services. As regulators' focus changes, they may find it useful to examine the existing organizational structure and the skills and duties of the commission staff to determine whether they are well-taylored to evaluate outcomes under the alternative regulation. Alternative regulation may not necessarily lead to less regulation. Instead, as the form and focus of regulation changes, staff resources may need to be redeployed, and some additional training may be necessary.

Although in a static analysis, several of these likely effects may bode ill for residential access and usage customers, especially those in rural exchanges, the outcome need not be unpleasant. An optimistic view is that as LECs respond to the opportunities created by the greater incentives to modernize plant and cut costs and to the pressure of competition, even rural residential customers may benefit in a dynamic sense as technological advances, profit incentives, and competitive pressure put downward pressure on costs. Furthermore, the LEC or competitive local access providers (cable TV or wireless, for example) may be able to serve these customers more efficiently and cheaply using new technologies rather than traditional technologies. A telephony analogue to the airlines' hub and spoke concept may evolve, with low-cost, local access providers (the telephony analogue of commuter airlines) entering those markets.

Other potentially positive effects of alternative regulation include: lower regulatory costs, higher profits for efficient LECs, greater innovation and faster introduction of new services, more rapid deployment of advanced infrastructure, higher service quality, and the evolution of more competitive markets, and more rapid economic growth and development. Potentially negative effects of alternative regulation include greater difficulty in obtaining information on costs and profits of LECs and their affiliates (exacerbating existing information asymmetry problems); possible enhancement of LECs' ability to use monopoly power, including their ability to adopt Ramsey pricing, if competitive and noncompetitive services are included in the same price-cap baskets; loss of subsidies to favored uses and customers; possible failure of inefficient LECs; universal service problems; and possible division of markets into information haves and have-nots.

Each of the potentially negative outcomes, if thought to be significant, may require protective measures, which may be as simple as monitoring behavior or as complex as traditional regulation, though with a different focus. However, each such protective measure has a cost. If regulators attempt to create a regime that simultaneously promotes competition, provides incentives for efficiency and innovation, and protects favored customers, the likely result is that they may not accomplish any of these goals.

It is always tempting for analysts or policymakers to believe that the current period is the most tumultuous transitional period ever faced. In current times regulators are faced with rapid and constant changes. But the next decade will have even more rapid and widespread sets of changes. The expressed concern by regulators for an analysis of approaches to assessing the impact of regulatory pricing reforms attests to the need to measure and analyze these changes. Commissions are change agents that seek to establish optimal and equitable conditions in which universal service and competition goals can be reached successfully. Part of the design of favorable conditions rests on commissions receiving reliable and valid information about the outcomes and impacts of pricing reforms so that subsequent reforms can be better designed to accomplish universal service and competition goals.

The program evaluation and qualitative approaches can include economic analyses, although most economic analyses are solidly and emphatically quantitative. Overlaps can occur and each approach is undoubtedly strengthened when overlaps occur. Program evaluations that do not consider economic analysis should be considered incomplete. Likewise, economic analyses of pricing reforms are strengthened to the extent that the program evaluation linkage of goals and impacts is used. Qualitative analyses are important because standard program evaluation analyses and economic research designs draw their strength from using time-tested approaches. As new relationships, factors, and dimensions emerge in a transition to future workably competitive information-age telecommunications markets, time-tested research tools may have their results enhanced through qualitative research. As regulators are on a rapidly changing frontier, the use of multiple analytical approaches can only strengthen the reliability, validity, and usefulness of a commission's evaluation of the impact of its pricing reform.

QUALITATIVE RESEARCH BIBLIOGRAPHY

Cook, T. D. & D. T. Campbell. *Quasi-experimentation: Design and Analysis Issues for Field Settings.* Chicago: Rand McNally, 1979.

Cronbach, L. J. et al. *Toward Reform of Program Evaluation: Aims, Methods and Institutional Arrangements.* San Francisco: Jossey-Bass, 1980.

Denzin, Norman K. (1989). *The Research Act.* Englewood Cliffs, NJ: Prentice-Hall.

Hoaglin, D. C., Light, R. J., McPeek, B., Mosteller, F., & M.A. Stoto. (1982). *Data for Decisions: Information Strategies for Policymakers.* Cambridge, MA: Abt Books.

Kirk, J. & M. Miller. (1986). *Reliability, Validity and Qualitative Research.* Beverly Hills, CA: Sage.

Lincoln, Yvonna S. & Egon G. Guba. (1985). *Naturalistic Inquiry.* Newbury Park, CA: Sage.

Lofland, J. (1971). *Analyzing Social Settings: A Guide to Qualitative Observation and Analysis.* Belmont, CA: Wadsworth.

Miles, M. B. & Huberman, A. M. (1984). *Analyzing Qualitative Data: A Source Book for New Methods.* Beverly Hills, CA: Sage.

Murphy, J. T. (1980). *Getting the Facts: A Fieldwork Guide for Evaluators and Policy Analysts.* Santa Monica, CA: Goodyear.

Patton, Michael Q. (1990). *Qualitative Evaluation and Research Methods.* Newbury Park, CA: Sage.

Strauss, Anselm & Juliet Corbin. (1990). *Basics of Qualitative Research.* Newbury Park, CA: Sage.

Tesch, R. (1989). *Qualitative Research: Analysis Types and Softward Tools.* Philadelphia, PA: Taylor and Francisc.

Van Maanen, J., Dabbs, J. M., Jr., & R. R. Faulkner. (1982). *Varieties of Qualitative Research.* Beverly Hills, CA: Sage.

Yin, Robert K. (1989). *Case Study Research: Design and Methods.* Newbury Park, CA: Sage.