Final Report on

REGULATION AS A SYSTEM OF INCENTIVES

Daniel Z. Czamanski
Institute Fellow and Associate Professor of City and Regional Planning

J. Stephen Henderson
Institute Associate and Assistant Professor of Economics

Curtis J. Odle
Graduate Research Associate

Vivian Witkind
Graduate Research Associate

THE NATIONAL REGULATORY RESEARCH INSTITUTE
2130 Neil Avenue
Columbus, Ohio 43210

prepared for the

U.S. Department of Energy
Economic Regulatory Administration
Division of Regulatory Assistance

Grant No. DE-FG01-80RG10268

December 1980

EXECUTIVE SUMMARY

Regulation is a set of practices that forms rules intended to frame the behavior of utility companies. In this report, the behavior of utility companies is viewed in terms of the self-interest of decision-making managers. The extent to which the observed behavior differs from that which regulators attempt to elicit depends upon the nature of information exchanges between regulators and utility managers and the existence of enforcement practices in cases of broken rules.

Incentives are viewed in this report as mechanisms that motivate behavior. Thus, the current set of regulatory practices is viewed as an incentive mechanism. Other mechanisms are possible, and several are outlined. All are studied in terms of information flows and reward patterns. An extensive analysis is performed of current rate-of-return regulation, and empirical studies are examined for evidence bearing on the hypotheses of this report. These hypotheses are directed toward potential efforts by regulators to institute a program of utility cost control.

Among the hypotheses are several concerned with managerial reaction to rate of return regulation and monitoring of inefficient, or wasteful, behavior. If monitoring can detect all waste, the regulator can achieve behavior similar to that under competitive conditions by reducing the allowed rate of return. If not, positive profits, beyond those required to keep factors of production employed, are needed to avoid waste that can be as large as unregulated monopoly profits. Thus, beyond a certain point, attempts to eliminate profits and to reduce rates are counterproductive.

There may be barriers to introducing new incentive mechanisms. An attempt is made to identify problems affecting the feasibility of alternative regulatory practices with particular reference to state governments.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

COST CONTROL AND INCENTIVES

Cost Control and Regulation

Criticism of rate-of-return regulation as the cause, or as a contributing cause, of the "cost control problem" is typically framed in very general terms, such as the absence of a ceiling on increases in utility bills. Such criticism obscures the fundamental problem and serves to focus public attention on cost control issues during periods of inflation only. The cost control problem with which this report is concerned does not stem from inflation. It is a direct manifestation of the incentives that the current practice of regulation provides utilities.

While it is true that economic regulation has been introduced sometimes in response to an industry's desire to protect itself from competitive pressures, public interest advocates justify regulation as a mechanism for the protection of consumers in cases where the existence of monopoly is inevitable or desirable. The inevitability, or desirability, of monopoly is attributed to the technological, or natural, circumstances that for some industries create subadditive costs, that is, a situation such that the costs incurred by a single producer of any quantity of output are less than the costs that would be incurred by two or more producers. Thus, a natural monopoly is justified by reference to cost reductions. Yet, once a firm establishes itself as a monopoly, there is a need for regulation to control it from exercising excessive power in the absence of disciplinary forces associated with competition.

For the purpose of this report, the cost control problem should be viewed as the socially undesirable propensity of regulated utilities to do the following:

(1) underproduce their output, that is, display technical inefficiency
(2) combine factors of production in a manner that does not lead to the minimum attainable costs in light of the current technology and prices, that is, display allocative inefficiency
(3) pay for factors of production in excess of amounts that are needed to secure the factors' services
(4) underinvest in new technology and organizational change so that future cost-minimizing combinations of factors can be assured

It is important to begin consideration of this cost control problem by recognizing that the extent to which it occurs and the cost that society bears as its consequence are not fully known. Equally important is the recognition that changes in regulatory practice may lead to other, and perhaps more serious, manifestations of this problem. Finally, for purposes of this report, utilities are viewed as heterogenous complex organizations that do not harbor institutional motives, but whose behavior can be understood in large part and directed by reference to the self-interest of the individuals who manage them.

Managerial Self-Interest and the Discipline of Markets

The need to view utilities in terms of the self-interest of their managers is associated with the character of the environment within which they operate. The single motive of profit maximization that is generally ascribed to firms may not be appropriate for describing utilities' behavior.

The basic reason for the use of profit motive as the sole driving mechanism in attempts to explain the behavior of unregulated firms is the presumption that the environment within which such firms must operate provides a number of forces that discipline them to behave as if they were maximizing profits. These forces arise from a variety of directions. Several have been identified very early in the history of economics by Adam Smith. For example, product market competition prevents firms from controlling the prices that are charged for products. Furthermore, entrepreneurial firms that succeed to bring down the price at which they can offer products are rewarded by increased sales and are soon emulated by other firms in the industry.

If product market competition were the only disciplining force, it would be unreasonable to expect that profit maximization is a sufficient description of the behavior of modern enterprises. Dissatisfaction with the profit motive as the full explanation of firm behavior arises out of recognition that in modern enterprises the interest of managers may be different from those of security holders. Management's interest in perquisites and shirking is satisfied and occurs at the expense of profits that otherwise accrue to holders of the residual claims on firms.

In addition to product market competition, three forces are presumed to discipline managers from making decisions that deviate from those consistent with profit maximization. One is the current prominence of the market for corporate control and the frequency with which the managements of industrial concerns are replaced by outside managers, suggests that non-profit-maximizing behavior by management leads to deviation of the book value of corporate assets from their market value. Such deviations are sufficient to invite takeover bids.

A somewhat different, and yet related, disciplining force arises out of the market for managerial labor. Managerial mobility is circumscribed by the fact that a manager who has permitted several successful takeover bids in his lifetime will experience a decrease in the present value of his human capital. There is no reason to assume that managers are not self-interested and not rational and therefore that they do not attempt to maximize the value of their human capital.

Still another disciplining force is associated with a market for financial capital. The basic cost of capital is determined through the interaction of the demand for, and the supply of, investable funds. The cost of capital to specific firms, however, is also a function of the past and current performance of those firms in terms of profits. Inasmuch as management perquisites are bought out of profits, ability to raise capital in the capital market is also in the interest of management and serves as another disciplining force.

The extent to which these various forces discipline managements and permit the retention of the profit motive as the sole driving characteristic of models of firms depends crucially on the information that is available in the various markets. Such information is, to a large extent, supplied by managements. It is a basic assumption of this report that there are strong incentives for managements to withold information and to provide misinformation. Particularly in the case of regulated utilities, some of the external forces that are typically presumed to discipline the management of firms are altogether absent. The introduction of regulation as a control mechanism introduces a new set of incentives that requires examination.

Regulation and the Principal-Agent Relationship

To posit the existence of a cost control problem, as defined above, is not equivalent to a criticism of the effectiveness of the current practice of regulation. It is plausible that little can be done to reduce the effects of the cost control problem. Nevertheless, examination of the incentives that the current practice of regulation provides utility managers and probing into the potential repercussions of alternate practices constitutes a productivity assessment of regulation that is rarely undertaken.

Incentives, or forces that motivate individuals to action, will be viewed in this report in the context of a highly stylized model of the relationship between a principal and the agent who is hired to perform actions on behalf of the principal. As in all such relationships, it is presumed that the agent is predisposed to limit his activity on behalf of the principal and that he is self-interested. All contracts that specify principal-agent relationships are faced with the problem of ensuring that the agent does in fact perform in the principal's interest. The problem arises due to the informational asymmetry that characterizes all such relationships. In particular, the principal is not able to observe the activity of the agent in its entirety. All attempts to monitor the activity of the agent are costly. At the same time, it is an assumption of this model that it is in the interest of the agent to obscure the information that the principal's monitoring is intended to uncover.

The current regulatory practice represents a contract between society and the utility. As society's agent, the utility is expected to provide its services to all who demand it at the least possible cost. The regulatory contract specifies that in return for its services the utility will be allowed to earn with mimimum risk a certain level of revenues that is consistent with the earnings of other industries with similar risk factors. To ensure that the contract's requirements are met, society through its representatives, the regulatory commission, "controls" the activity of the agent. There are two instruments that are typically employed by the principal: (1) the principal engages in monitoring to ensure that the utility does not pass onto the principal costs that should not have been experienced in the process of producing the utility's services, and (2) the principal sets an upper limit on the profits that the utility can earn.

Alternate contracts between principals and agents, or alternate regulatory practices, have the potential of generating different behavior on the part of the agent or utility. It is a major purpose of this report to examine the implications of current regulatory practices for the behavior of utilities. In particular, the implications of such practices will be studied in terms of the cost behavior of utilities. Furthermore, an attempt will be made to describe and to examine other types of contracts that should be considered in the context of regulation. These contracts have been developed to deal with general principal-agent problems in other frameworks. Their suitability depends on the particular circumstances that arise in the context of regulation. An additional objective of this report is to explore the feasibility of introducing new regulatory practices.

A cautionary note is required at the outset. The analysis of regulation as a system of incentives and the discussion of associated policy options is in its infancy, and so is the application of the principal-agent model here applied to the regulatory process. The current state of knowledge does not permit firm policy recommendations. The major suggestions that emerge from this report concern the need to formulate empirical generalizations and to test selected hypotheses.

Chapters 2 and 3 treat the effects of various regulatory practices upon utility behavior. The two chapters provide competing hypotheses in need of further empirical research. Chapter 4 provides a review of existing knowledge concerning the actual effects of regulatory practices. Although the current practice of regulation can be viewed in incentive terms, the current practice includes many different signals-giving practices. Chapter 5 offers a typology of these incentive mechanisms. Chapters 6 and 7 are concerned with the feasibility of introducing new incentive mechanisms. Chapter 8 contains a summary of findings and suggestions for future research directions.

CHAPTER 2

OVERCAPITALIZATION AND EXPENSE PREFERENCE

Introduction

There is no general theory of firm behavior under rate-of-return regulation. Despite the long history of this type of regulation, the modern literature dates to the early 1960s [2]. This conspicuous gap in our knowledge is a reflection of a more serious problem in the general theory of the firm. While profit maximization is a sufficient characterization of modern enterprises for some purposes, in the case of regulated utilities it has led to a few useful research findings.

By far the vast majority of studies to date have focused on two hypotheses. This chapter reviews the Averch and Johnson theory of the regulated firm [2], the associated A-J effect, and Crew and Kleindorfer's theory of the expense-preference-regulated firm [17]. Both models explore the potential misallocation of resources, wasted capital, and operating inefficiencies that can occur if regulation permits the firm to earn a rate of return on capital above the cost of capital. In the final section of this chapter, an attempt is made to reconcile the results arising from these two models. Specifically, it is shown how an expense-preference firm's allocation of resources is related to overcapitalization induced by a regulatory authority that allows a firm to earn a return on capital in excess of the firm's true cost of capital.

The A-J effect is the result of a firm's failure to minimize costs while attempting to maximize profits. The traditional firm encountered in neoclassical microeconomic theory will achieve both profit maximization and cost minimization in an unregulated environment. Crew and Kleindorfer's firm maximizes utility with both profits and staff expenditures included in the objective function. This implies that the firm makes a trade-off between profits and staff expenditures in its attempt to maximize the welfare of its security holders and management team. Expenditures on staff may exceed some minimum level necessary to produce the required level of output. If this is the case, resources are being wasted on nonpecuniary benefits to managers of regulated firms. Like the A-J effect, appropriating unnecessary expenditures on staff is wasteful. The regulatory authority must employ some sort of monitoring mechanism if it hopes to detect wasted revenues and to eliminate this source of waste.

In the recent literature, attention has focused on the use of a utility function to represent the preferences of a firm. Criticisms of such an approach have come mainly from advocates of perfect capital markets. They contend that using a single utility function to represent the firm's preferences implies that both security holders and owner-managers have identical preferences and tastes; this is an unnecessary assumption and a misleading one. In chapter 4, the use of such a function does not justify the above assumption.

The Averch-Johnson Model

The following notation is used in the subsequent analysis:

L = physical units of labor,

K = physical units of capital,

w = wage rate per physical unit of labor (constant),

c = acquisition cost per physical unit of capital,

r = the market cost of borrowing funds (constant),

s = the fair return on investment set by the regulator,

$R(L,K)$ = revenues as determined by the firm's demand and production functions,

$g(L,K)$ = a continuous and twice differentiable production function,

$\pi(L,K)$ = profits.

Bailey [4] presents the Averch-Johnson (A-J) model in a form slightly different from the original model. Bailey permits operation off the production frontier and includes an explicit parameter for the acquisition cost of capital. The model explored below is of the Bailey variety.

The A-J model assumes a profit-maximizing firm subject to a regulatory constraint. The constraint imposes a ceiling on the firm's return-on-capital investment. The objective function can be stated as,

$$\underset{L,K}{\text{Maximize}}\ \pi(L,K) = R(L,K) - wL - rcK \qquad (2\text{-}1)$$

$$\text{Subject to: } \frac{R(L,K) - wL}{cK} \leq s, \quad s > r. \qquad (2\text{-}2)$$

The model assumes that the regulator determines the maximum allowable rate of return on investments, s, and the firm then adjusts its price and other decision variables in accord with s. In this manner, the price adjustment is implicit rather than an explicit component of the model.

Labor input, L, is total man-hours worked and the average cost per man-hour, w, is assumed constant. Capital input, K, is units of capital and cK is the cost of tying up the assets required for production. Presumably, this capital cost includes plant and equipment (physical capital) as well as financial assets, such as working capital. Thus, cK is the capital stock expressed in dollars. Furthermore, since c is constant, the firm does not influence the price it pays for capital goods. The cost of capital, r, is the minimum return the firm must earn in order to continue to be able to raise funds over the long run. The cost of capital is assumed to be known and independent of the mix of financial instruments employed.

The rate-of-return constraint, equation (2-2), can be translated into a profit constraint by multiplying both sides of (2-2) by cK, which yields,

$$R(L,K) - wL \leq scK. \tag{2-3}$$

Subtracting rcK from each side yields,

$$\pi \leq (s - r)cK. \tag{2-4}$$

Bailey points out that (2-4) suggests that rate-of-return regulation is a type of profit ceiling regulation, under which the profit ceiling increases with increased capital usage. However, the ceiling is independent of increases in labor usage. This suggests that the firm may adjust its use of inputs other than capital without limiting its profit potential. The choice of capital, however, is an explicit determinant of the profit ceiling.

If profit is considered a random variable, the ceiling imposed on profits serves as a truncation point. In this way, profits can be written as,

$$\pi = \begin{cases} (s - r)cK, \text{ if } \pi \geq (s - r)cK \\ \tilde{\pi} \text{ otherwise.} \end{cases} \tag{2-5}$$

where $\tilde{\pi}$ is random profits (the tilde denotes a random variable).

If unconstrained profits are larger than those permitted in the regulatory framework, then equality holds in the constraint and the firm will maximize profit if it earns the maximum return permitted by the regulator. In this case, the objective is to

$$\text{Maximize } (s - r)cK \tag{2-6}$$

$$\text{Subject to: } \frac{R(L,K) - wL}{cK} = s, \; s > r. \tag{2-7}$$

Figure 2-1 illustrates the profit function of a hypothetical firm operating under the regulatory constraint. Point M on the profit hill in figure 2-1 represents the combination of capital and labor that yields maximum profits to the firm. The rate-of-return constraint precludes the firm from earning M profits, however. The constraint is $\pi = (s - r)cK$ that is a ray passing through the origin with slope $(s - r)c$. The point E gives the solution to the constrained maximization problem. Figure 2-1 shows that the firm will attain profits of E by using K_E units of capital. In the absence of a constraint, the firm will earn profits of M by using K_M units of capital. Since $K_M < K_E$, the regulated firm employs more units of capital than it would if the constraint was eliminated. The implications of the results depicted in figure 2-1 are now formally presented.

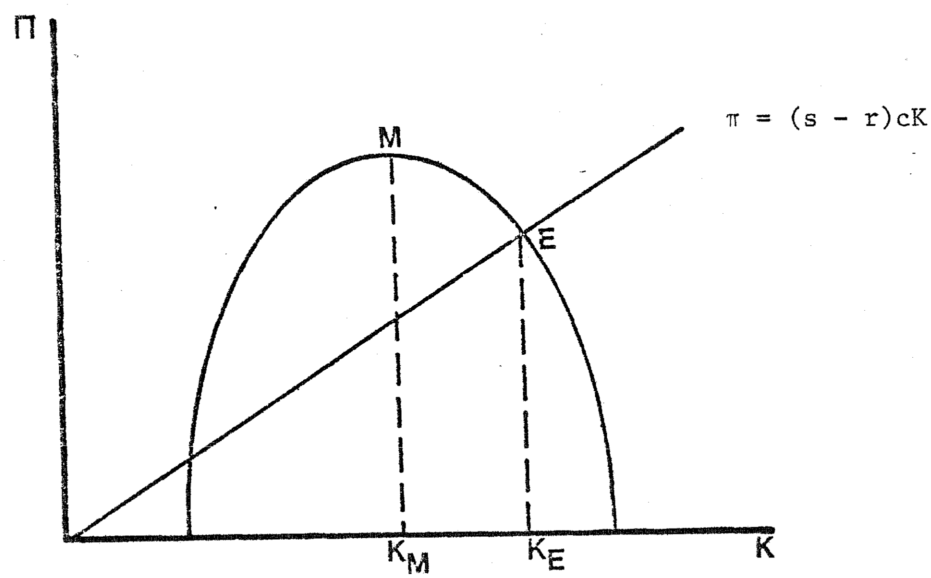


Figure 2-1 The profit function of a firm subject to a rate-of-return regulatory constraint

Source: Bailey, E.E. [4, p. 71]

Consistent with equations (2-1) and (2-2) the objective of the firm is stated as follows:

$$\text{Maximize} \quad (L,K,L^*,K^*) = R(L,K) - w(L + L^*) - rc(K + K^*) \qquad (2\text{-}8)$$

$$\text{Subject to:} \quad (L,K,L^*,K^*) \leq (s - r)c(K + K^*), \; s > r, \qquad (2\text{-}9)$$

where L* = number of units of labor wasted,

K* = number of units of capital wasted.

The associated Lagrangian can be written as follows:

$$\Phi(L,K,L,K^*) = (1 - \lambda)[R(L,K) - w(L + L^*) - rc(K + K^*)] + \lambda\, [(s - r)c(K + K^*)]. \qquad (2\text{-}10)$$

Maximizing (2-10) with respect to L, L^*, K, K^* and λ yields the following first-order conditions necessary for firm profit maximization:

$$\Phi_L: \quad (1 - \lambda)(R_L - w) = 0; \qquad (2\text{-}11a)$$

$$\Phi_K: \quad (1 - \lambda)R_K = (1 - \lambda)rc - \lambda(s - r)c; \qquad (2\text{-}11b)$$

$$\Phi_L{}^*: \quad \lambda < 1, \quad L^*(1 - \lambda)(-w) = 0; \qquad (2\text{-}11c)$$

$$\Phi_K{}^*: \quad \lambda < \frac{r}{s}, \; K^*c(-r + \lambda s) = 0; \qquad (2\text{-}11d)$$

$$\Phi_\lambda: \quad R(L,k) - wL - scK = 0; \qquad (2\text{-}11e)$$

$$L>0, \quad K>0, \quad \lambda \geq 0, \quad L^* \geq 0, \quad K^* \geq 0. \qquad (2\text{-}11f)$$

The subscripts on variables indicate the partial derivative of the variable with respect to the subscripted variable, for example, Φ_L: $\partial\phi/\partial L$. The implications of the result shown in figure 2-1 are formally derived in equation (2-11b). Cost minimization requires that the firm operate where the marginal revenue of capital is equal to the unit cost of capital: $R_K = rc$. The A-J model, however, shows that the regulated profit-maximizing firm operates where marginal revenue is less than rc. From equation (2-11b), we find that

$$R_K = rc - \frac{\lambda}{1-\lambda} (s - r). \qquad (2\text{-}12)$$

Marginal revenue is smaller by $(\lambda/(1 - \lambda)(s - r) > 0$. Moreover, dividing equation (2-11b) by equation (2-11a) and rearranging terms yields

$$\frac{R_K}{R_L} = \frac{R\ q_K}{R\ q_L} = \frac{q_K}{q_L} = \frac{rc}{w} - \frac{\lambda}{1-\lambda} \frac{(s - r)\ c.}{w} \qquad (2\text{-}13)$$

Equation (2-13) differs from the familiar cost-minimizing solution found in the traditional theory of the firm: efficient operation requires that

$$- \frac{dL}{dk} = \frac{q_K}{q_L} < \frac{rc}{w} \qquad (2\text{-}14)$$

Comparing (2-13) and (2-14), we see that for the firm regulated by a rate-of-return constraint,

$$- \frac{dL}{dk} = \frac{q_K}{q_L} < \frac{rc}{w} \qquad (2\text{-}15)$$

This clearly shows that the regulated firm is using relatively too much capital and too little labor in its production process. Bailey [4, p.76] states that "the profit maximizing firm constrained to earn at most a fair return on investment selects a production technique that uses more capital and less labor than are consistent with minimum-cost operations."

Wasteful Expenditures

The model presented in equations (2-8) and (2-9) allows the firm to operate off the production frontier by wasteful expenditures either on staff or on capital. In practice, capital waste is allocated to the rate base. Waste, x, can be written as,

$$x = wL^* + rcK^*. \qquad (2\text{-}16)$$

Cost increases if the firm operates off the production frontier; revenue, however, is not altered in this model. Revenue is a function of L and K only, that is, labor and capital used productively.

An alternative formulation including the possibility that the firm will operate off the production frontier in an A-J model was suggested by Shapiro. (See [4, p. 73 note f].) In Shapiro's model, the firm's objective is

$$\underset{L,K,q}{\text{Maximize}} \ \pi = pq - wL - rck \qquad (2\text{-}17)$$

$$\text{Subject to: } \pi \leq (s - r)cK, \ s > r, \qquad (2\text{-}18)$$
$$q \leq q(L,K),$$

where

$q(L,K)$ = potential production,

q = actual production.

Shapiro's model, like Bailey's formulation, posits that revenues are determined by actual rather than by potential production. Shapiro was able to derive essentially the same set of results as Bailey.

Using Bailey's model, it is possible to prove that the regulated firm has no incentive to operate off the production frontier if the marginal physical product of capital is positive. Such a result suggests that waste, x, is zero for the regulated profit-maximizing firm. Thus, equation (2-16) is not applicable to the firm under study.

The A-J model is a useful paradigm for examining what would happen if commissions and utility managements did not do what they were supposed to do; that is, set the allowed rate of return equal to the market determined required return on investment.

The A-J effect implies that capital waste comes about only if the allowed rate of return exceeds the cost of capital. Thus, overcapitalization will occur when the regulatory authority is guilty of misidentifying the true cost of capital. The A-J thesis assumes that earnings above the cost of capital will lead to inefficiency and higher costs and that the utility will be rewarded via higher rates. Inefficient operation will continue as long as the utility is allowed to earn more than its cost of capital.

Morton [51], however, contends that commissions sometimes seek to encourage efficiency by allowing the utility to earn a rate of return exceeding its cost of capital if the utility achieves this rate through efficient operation. However, a commission dedicated to efficiency and eliminating misallocations of resources will take away the excess earnings even if the utility earned a return above the cost of capital due to its superior efficiency.

It is a generally accepted rule that utility ratemaking should attempt to identify the opportunity cost of capital and to set the allowed rate of return equal to this opportunity cost over a period of time. This is, needless to say, a difficult task. Efficient management of utilities presumes that the firm will attempt to identify the least-cost combination of resources. This calculation depends on present and projected changes in factor prices, on the elasticity of demand, on scale of output, on changing technology, on relative prices, and on other forecasted economic factors. Thus, the least-cost combination of factor prices is not precisely measured. It is better thought of as an estimate of the least-cost combination that will continuously change as existing plants, processes, and relative prices change.

Morton contends that the problems of production, rising costs, and capital expansion are real and earnest. The A-J effect, however, is unsupported by evidence. Economic conditions and changing circumstances constantly keep utilities under pressure to control costs in order to maintain or to increase their rate of return without applying for rate increases that are both costly and bring about customer resistance. Utilities aim to maximize their return on capital and are aware that the economic environment is capable of penalizing them for inefficient operation.

The existence of a ceiling on profits may create a disincentive for efficient operation. Any ceiling on earnings that does not consider the source of earnings may create a situation where managers become less cost conscious. If management feels that lower costs will lead to a reduction in rates and profits, it may not attempt to initiate or to monitor cost-reducing programs. Commissions, it seems, must therefore assume that managers of utility companies are subservient to the public's benefit and welfare. Morton [51, p. 21] states that commissions must "...rely on management morale together with surveillance to stimulate efficiency and to avoid indifference and lethargy."

Because of the moral hazard problem, it is not a desirable policy to strip the utility of all incentives to cost reduction and service improvements, even if it were possible to limit the utility's earnings to a fixed amount. This may cause the utility to become reckless in its efforts to control expenses. Since a ceiling on profits does exist in rate-of-return regulation, however, management may have an incentive to appropriate nonpecuniary benefits if such expenditures, or shirking, will not be detected by the commission's monitoring mechanisms. The implications of this sort of behavior are examined further in the following section.

The Crew and Kleindorfer Model

Crew and Kleindorfer [17] address the question of the significance and control of the A-J effect if public utilities are not profit maximizers. They examine the interaction between managerial discretion and regulation in order to examine the A-J effect and rate-of-return regulation.

Crew and Kleindorfer (CK) use the concept of an "expense-preference" firm and represent the utility function of a representative firm as $U(S,\pi)$, where S is staff expenditures and π is discretionary profit. The inclusion of S in the firm's utility function is based on the notion that managers conceivably obtain more utility from expenditures on staff than they do from physical assets or additional labor. It is assumed, therefore, that $\partial U/\partial S > 0$ and $\partial U/\partial \pi > 0$.

Profit, π, is defined in the CK model as follows:

$$\pi = xP(x,A) - wL - \beta K - A - S, \tag{2-19}$$

where

x = the quantity of output,

$P(x,A)$ = the inverse demand function,

w = the constant price of labor,

L = the quantity of labor input,

π = the price of a physical unit of capital,

K = the physical quantity of capital,

A = advertising expenditure,

S = staff expenditures.

It is assumed that $x \leq F(K,L)$ and $S \geq \underline{S}(x,A)$, where F is a quasi-concave, neoclassical production function, and $\underline{S}(x,A)$ represents minimal expenditure on staff sufficient to sustain an output and advertising expenditure A.

The objective of the expense-preference regulated firm may be stated as follows:

$$\text{Maximize } U(S,\pi) \tag{2-20a}$$

$$\text{Subject to: } S \geq \underline{S}(x,A) \tag{2-20b}$$

$$x \leq F(K,L) \tag{2-20c}$$

$$\pi = xP(x,A) = wL - K - S - A \leq (s - \beta) K \tag{2-20d}$$

$$x,K,L,A,S \geq 0, \tag{2-20e}$$

where s $(s > \beta)$ represents the allowed rate of return.
The Lagrangian may be written as:

$$\mathscr{L} = U(S,\pi) - \mu_1(S - \underline{S}(x,A) + \mu_2(F(K,L) - x) + \mu_3(s - \beta)K - \pi), \tag{2-21}$$

where μ_1, μ_2, and μ_3 are the Lagrangian multipliers.

Maximizing (2-21) with respect to x, A, K, L, S, the following first-order conditions are necessary for a maximum solution:

$$x\text{:}\quad (U_2 - \mu_3)\,[P + xP_1] - \mu_1\underline{S}_1 - \mu_2 = 0; \tag{2-22a}$$

$$A\text{:}\quad (U_2 - \mu_3)\,[xP_2 - 1] - \mu_1\underline{S}_2 = 0; \tag{2-22b}$$

$$K\text{:}\quad (U_2\beta - \mu_3 s) - \mu_2 F_1 = 0; \tag{2-22c}$$

$$L: \quad (U_2 - \mu_3)w - \mu_2 F_2 = 0; \tag{2-22d}$$

$$S: \quad (U_1 - U_2 + \mu_1 + \mu_3 = 0; \tag{2-22e}$$

$$\mu_i \geq 0; \; \mu_1(S - \underline{S}) = \mu_2(F - x) = \mu_3((s - \beta)K - \pi) = 0. \tag{2-22f}$$

The numerical subscript, i, to a function indicates partial differentiation of the function with respect to its i-th arguments.

In order to examine the significance of the A-J effect, assume $F_1 > 0$, $F_2 > 0$. From (2-22c) and (2-22d) we get:

$$\frac{(U_2\beta - \mu_3 s)}{F_1} = \frac{(U_2 - \mu_3)w}{F_2} = (\mu_2). \tag{2-23}$$

Rearranging terms yields:

$$\frac{\beta}{F_1} - \frac{(s - \beta)\mu_3}{(U_2 - \mu_3)F_1} = \frac{w}{F_2}. \tag{2-24}$$

Using Bailey's notation, (2-24) may be written as:

$$\frac{rc}{q_k} - \frac{\mu_3(s - r)c}{(U_2 - \mu_3)q_k} = \frac{w}{q_L} \tag{2-25}$$

or

$$\frac{q_k}{q_L} = \frac{rc}{w} - \frac{\mu_3}{(U_2 - \mu_3)} \frac{(s - r)c}{w}. \tag{2-26}$$

Comparing (2-26) with (2-13) shows that the term $\lambda/(1 - \lambda)$ in (2-13) is now replaced with the term $\mu_3/(U_2 - \mu_3)$ in (2-26). Thus, the CK model is consistent with Bailey's model in that the firm is not cost minimizing, that is, the firm employs too much capital and too little labor. The overcapitalization effect persists for expense-preference firms unless $s = \beta$ or $\mu_3 = 0$.

The Effect of Changing s on the Regulatory Constraint

Bailey shows that as the fair rate of return is set closer to the cost of capital in an A-J model, the level of constrained profit declines. This proposition is graphically presented in figure 2-2. The constraints are drawn with $s_1 > s_2$. At point A, the firm will use K_A units of capital. A decrease to s_2 causes the profits to decline. The firm responds by employing K_E units of capital corresponding to point B on the profit hill. Bailey shows that $d\pi/ds = (s - r)cdk/ds = (R_K - rc)ck/R_K - sc) > 0$. The change in profits, caused by a change in s, will induce a response similar to that illustrated in figure 2-2.

Crew and Kleindorfer note that for the regulated profit-maximizing firm the magnitude of the A-J effect is the strongest when s is near β and decreases monotonically to zero when s equals the pure monopoly rate of return. This is shown in figure 2-2 by the fact that as s approaches r, the magnitude of the A-J effect increases. Note, however, that as s approaches r, each incremental decrease in $\bar{s}$ causes the firm to respond by employing smaller and smaller increments of capital. That is, the marginal increase in K, induced by a small decrease in s, declines as s approaches r. This is obviously a consequence of the shape of the profit hill.

For the expense-preference firm, the A-J effect is zero when s equals β or when $\mu_3 = 0$. Changing s in the constraint has the effect of changing the effect profit has on the value of the objective function, $U(S,\pi)$. As s approaches β, a dollar invested in an additional unit of capital (beyond the cost minimizing level, point M) has decreasing attractiveness relative to spending the same dollar on staff. The convexity of the objective function requires that staff expenditures be raised as profit declines. Thus, as the allowed rate of return decreases and K increases, staff expenditures must also increase. Regulation has the effect of substituting inefficiency in the use of staff for the A-J type of inefficiency in the use of capital.

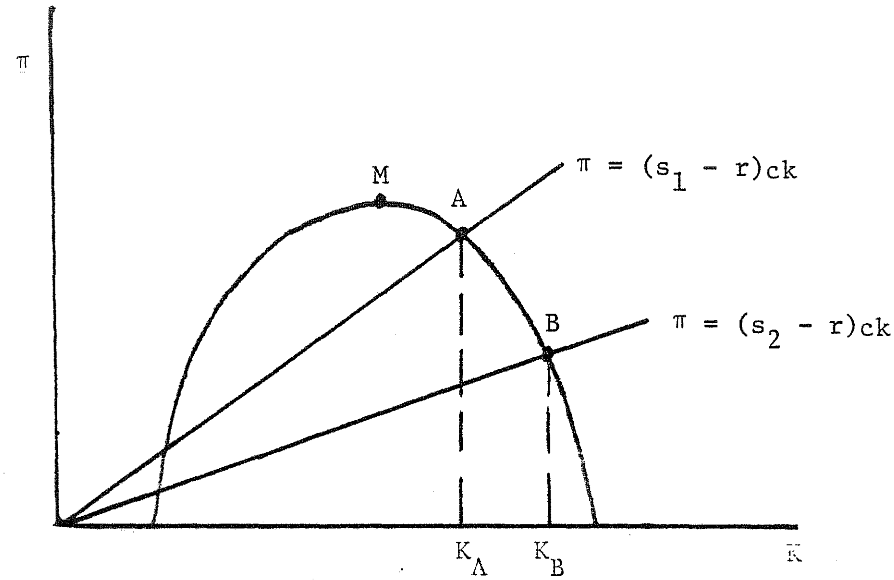


Figure 2-2 Capital and profit response to changes in fair return ($s_1 > s_2$)

Source: Bailey, E.E. [4, p. 90]

Some additional insight into the CK analysis can be gained by totally differentiating the objective function and setting the resulting equation equal to zero:

$$dU = U_S dS + U_\pi d\pi = 0. \qquad (2\text{-}27)$$

Rearranging, we obtain:

$$\frac{U_S}{U_\pi} = -\frac{d\pi}{ds} \qquad (2\text{-}28)$$

Equation (2-28) is graphically presented in figure 2-3.

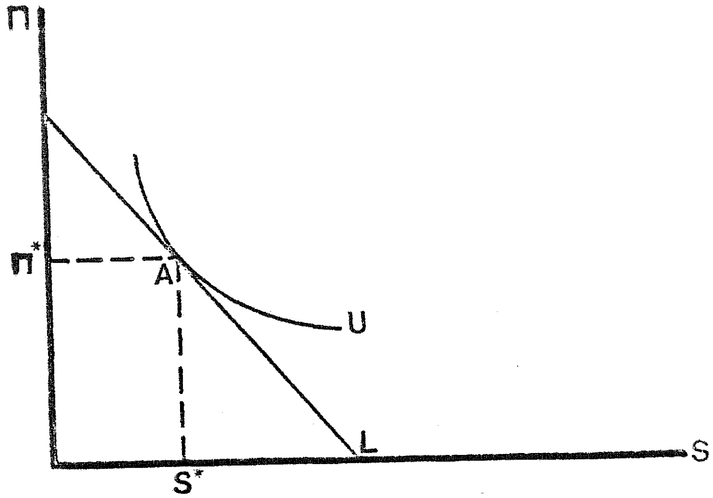


Figure 2-3 Indifference curve for the expense-preference firm

Equations (2-27) and (2-28) and figure 2-3 tell us that the utility maximizing expense-preference firm will trade off profits for staff expenditures, represented by curve U in figure 2-3, at a rate equal to the slope of line L in figure 2-3. Technically, the firm will equate the ratio of the marginal utility of staff to the marginal utility of profits to $-\frac{d\pi}{ds} = -1$. This corresponds to point A in figure 2-3.

Combining figures 2-2 and 2-3, a representation of the relationship between staff and capital can be formulated; this is done in figure 2-4.

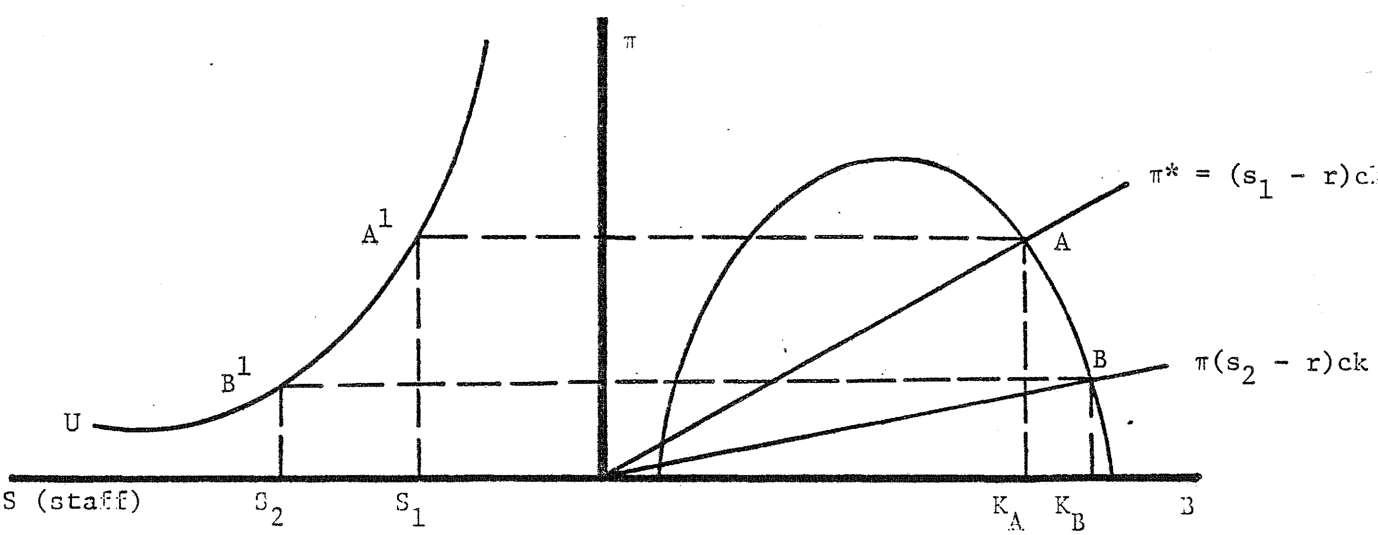


Figure 2-4 The relationship between capital (K) and staff expenditures (S)

Curve U in figure 2-4 is an indifference curve like that in figure 2-3, showing the firm's willingness to trade off profits for capital while maintaining a constant level of utility. For every point on the profit hill, there is a corresponding point on curve U; as $s \rightarrow r$, $S \rightarrow \infty$. The firm is willing to give up profits for an increase in staff. Therefore, as the allowed rate of return decreases, for example, from s_1 to s_2, the firm will desire an increase in staff expenditures equal to $S_2 - S_1$, thereby maintaining a constant level of utility.

A decrease in the allowed rate of return also causes the firm to invest in additional capital, for example, by the amount $K_B - K_A$. Note, however, that as s become smaller and smaller, the marginal increments to capital become smaller and smaller while the marginal increments to staff become larger. As profits decline, the firm responds by investing smaller amounts in capital and larger amounts in staff. It appears as though the firm desires to appropriate more dollars on staff and fewer dollars on capital as the regulatory constraint forces profits to decline. On the other hand, it appears that as the allowed rate of return increases and profits are allowed to move up the profit hill, the firm desires to reduce expenditures on staff at a faster rate than it desires to reduce capital expenditures.

Combining the results of the A-J model and the CK model shows that the firm's allocation of expenditures to staff has a definite relationship to the overcapitalization generated by a rate of return that exceeds the cost of capital. Figure 2-4 is a very general picture of this relationship and is a first approximation to reconciling the similarities and differences between the A-J effect and the CK results. Before further propositions can be stated, additional effort must be directed toward analyzing the comparative statics relevant to the scenario presented here and illustrated in figure 2-4.

CHAPTER 3

THE BEHAVIOR OF PUBLIC UTILITY MANAGERS UNDER RATE-OF-RETURN REGULATION

The behavior of managers who have a preference for particular kinds of expenditures has been examined previously in two ways. Williamson [88], Rees [67], and Crew and Kleindorfer [17] assumed that certain expenditures, in addition to being preferred in their own right, also enhance the firm's revenue either directly or because the staff expenditures increase advertising that in turn raises revenue. In contrast, Alchian and Kessel [1], Bailey [4], and Jensen and Meckling [42] discussed aspects of purely wasteful activity such as a manager's consumption of perquisites or simply leisure. Traditional rate-of-return regulation of each managerial type has distinct implications. Crew and Kleindorfer [17] recently studied the effects of regulating managers whose expense preferences lead to enhanced revenue. While the Crew and Kleindorfer model was examined in the previous chapter, this chapter examines the reaction of the purely wasteful manager to regulating pressure and then derives optimal regulation strategies.

The chapter has three sections. In the first, the manager's behavior is predicted from a utility maximization model. The manager has a taste for nonproductive activity and also benefits from the firm's profits. The regulator has two instruments with which to influence the manager--the regulated rate of return and monitoring of perquisites. In general, there are conflicting substitution and income-type effects that prevent the signs of the reaction paths from being determined. If output and perquisites are normal goods, however, tightening the regulatory pressure by reducing the allowed rate of return will induce the manager to consume more perquisites, while additional monitoring encourages more output. The response of output to the rate of return and that of perquisites to monitoring are ambiguous.

There are, in addition, two extreme expansion paths with clear regulatory implications. First, if the regulator can monitor perfectly, he can prevent managerial waste altogether, and the optimal policy would be to reduce the rate of return to eliminate profits above the competitive norm. Second, if the manager believes the regulator is limiting profits at a constant level (instead of taking the rate of return and monitoring activity as given), the manager would reduce profits to the acceptable level by consuming more perquisites. In particular, output is not increased in response to such a constraint. Clearly, this policy would never be in the public interest.

By contrast, rate-of-return regulation generally has favorable output effects over some range of the instrument variables. The second section of this chapter develops optimal fair rate-of-return and monitoring rules. An important issue is whether it is socially beneficial to eliminate profits. This is shown to depend on the manager's reaction path. In particular, if the expansion path in response to the rate of return has an output elasticity larger than the perquisites' elasticity, profits should be eliminated. This has a straightforward graphical interpretation. The optimal regulatory rules that account for managerial emoluments are quite different from the results of Klevorick [44], Sheshinski [75], and Bailey [4] in which returns to scale are most important.

Throughout the chapter we assume that other forces that may discipline managers are imperfect. Any perfect monitoring of perquisites and enforcement mechanisms, such as Fama's perfect labor market for managers, would make regulatory monitoring unnecessary and zero profits optimal. Wage contracts in regulated industries do not typically provide penalties for imprudent behavior. Also, managerial mobility among regulated firms would not appear to provide the same disciplinary force as is suggested by Fama's competitive example, since the profit motive that drives stockholders to reward or punish new managers for past behavior is diluted by the regulation itself. For these reasons, the issue of optimal regulation does indeed seem substantive. Some concluding remarks are offered in the third section.

Rate-of-Return Regulation and Managerial Behavior

This section examines the behavior of managers who value nonpecuniary benefits in addition to their share of profits. The firm's profits are limited by rate-of-return regulation. In addition, the regulator can monitor the manager's consumption of nonpecuniary benefits and exclude these from allowed revenue. Several authors, including Williamson [88], Rees [67], and Crew and Kleindorfer [17], have hypothesized that managers prefer some type of expenses more than others. Staff expenditures and well-furnished offices are examples. Such expenses are assumed to affect revenue in these theories. In contrast, we examine expenses that have no productive component whatsoever. The manager's "quiet life" is an often used illustration that captures the notion of nonproductive expenses. Bailey [4] briefly, and in a very minor way, addressed the problem of regulating such a manager, an idea we extend in order to develop the policy implications of monitoring and setting the fair rate of return.

Nonpecuniary benefits are denoted as S. The manager chooses productive activities and S to maximize his utility $U(S,B(\pi))$, where π is profits and $B(\pi)$ is a profit-sharing formula such as $\alpha\pi + b(\pi)$, where α is the manager's share of the firm's stock and $b(\pi)$ is a bonus. To simplify matters, assume the firm combines labor and capital in fixed proportions. This specification prevents an analysis of the Averch-Johnson type of overcapitalization. By an appropriate selection of measurement units, the constant returns to scale, fixed proportions technology can be represented as constant costs with profits written as

$$\pi = P(x)x - cx - S \tag{3-1}$$

where $P(x)$ is the inverse demand function, x is output, and c is marginal cost. In practice, the rate-of-return constraint allows a fair rate of return on capital, but in this model it can be specified as a rate of return on output:

$$P(x)x = R(x) \leq rx + S - D(S,M) \tag{3-2}$$

where $R(x)$ is revenue, r is the regulated rate of return per unit of output, and $D(S,M)$ are those nonpecuniary benefits or wasteful expenditures that have been discovered by the regulator through monitoring activity. Assume that zero monitoring discovers nothing, $(D(S,0) = 0)$, and that only a fraction of waste is indeed discovered, $(0 \leq D_s \leq 1)$, while more careful monitoring uncovers more total waste, $(D_M > 0)$. Note that discovered waste is excluded from allowed revenue but is included in profits. That is, the regulator is able to prevent such costs from being passed onto consumers but despite this, the manager continues the wasteful activity that must be deducted from profits. This implies that the stockholders are unable to discipline their managers even though the regulator has discovered some slack. At least some auditing by regulators has this character. Alternatively, suppose managers did indeed desist from consuming any discovered waste. Then only net waste, $N(S,M) = S - D(S,M)$, would be subtracted from profits. Also, utility is derived only from waste actually consumed, $U(S,B(\pi))$. Clearly, the manager's choice of N or S are equivalent problems; hence, for a given level of monitoring, the manager can calculate his preferred net waste consumption. If the regulator discovers 50 percent of all waste, the manager may engage in twice the desired activity knowing that half will be discovered. In these circumstances, a rational regulator would never monitor. The discussion of the $M = 0$ case, below, represents such a net waste model.

Given the regulated return, r, and monitoring activity, M, the expense-preference manager maximizes utility subject to (3-2). The Lagrangian expression for the problem is

$$L = U[S,B(\pi)] - \lambda[R(x) - rx - S + D(S,M)]$$

where λ is the regulatory shadow price. The first-order conditions require

$$L_x = B'U_2(R' - c) - \lambda(R' - r) = 0 \qquad (3\text{-}3a)$$

$$L_s = U_1 - B'U_2 + \lambda(1-D_s) = 0 \qquad (3\text{-}3b)$$

$$L_\lambda = [R - rx - S + D] = 0 \qquad (3\text{-}3c)$$

Without the restriction of regulation, the monopolist would select the output that equates marginal revenue and cost and consume perquisites until $U_1/U_2 = B'$. The latter is the Jensen-Meckling insight [42] that partial ownership changes the relative price of nonpecuniary benefits, since outside owners pay a portion, 1-B', of the manager's quiet life. Condition (3-3b) shows that regulation drives a wedge between the marginal utility of perquisites and money profits. Since $B'U_2 - U_1 = \lambda(1-D_s) > 0$, the manager responds to regulatory pressure by consuming more fringe benefits, thereby driving their marginal utility lower.

The manager's equilibrium is illustrated in figure 3-1. Point A represents the unregulated monopolist's choice of productive and non-productive activity, where utility is maximum. The regulation constraint for some r > c lies to the northeast of A with the feasible region shaded

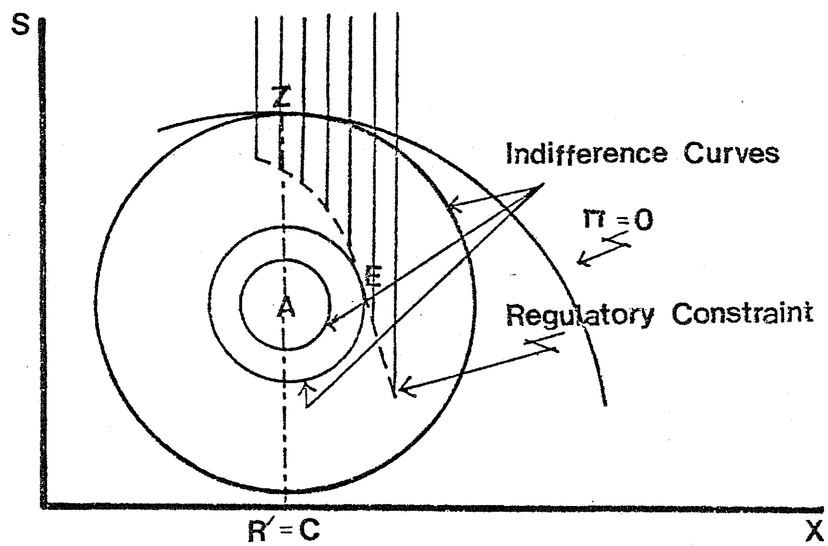


Figure 3-1 Manager's Equilibrium

in the diagram. If r is sufficiently high, the constraint will not be binding and A will be feasible. The regulated manager's best choice is point E, where his indifference curve is tangent to the nonlinear regulatory constraint. The indifference curve has a slope of

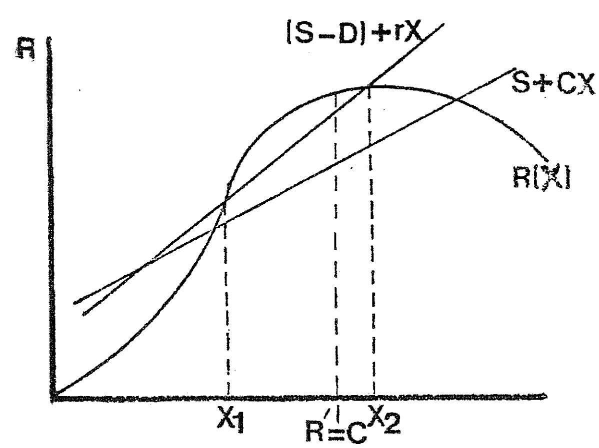
$$\frac{ds}{dx} = \frac{B'U_2(R' - c)}{B'U_2 - U_1} \qquad (3\text{-}4a)$$

where the slope of the regulatory constraint is

$$\frac{ds}{dx} = \frac{R' - r}{1 - D_s}. \qquad (3\text{-}4b)$$

Equating these two yields the first-order conditions (3-3a) and (3-3b). Notice that indifference curves have zero slope at points above the monopoly output, ($R' = c$). Ƶ is such a point in figure (3-1) and in particular lies on the zero profit locus that is also horizontal at $R' = c$. It must be true that the indifference curve through Ƶ lies within the zero profit locus because utility is clearly decreased by holding profit constant while perquisites are reduced. Hence, an indifference locus is more tightly curved than a constant profit locus. In turn, the regulation constraint has the same curvature as a constant profit line, both of which depend on the second derivative of the revenue function. Consequently, point E is indeed a constrained maximum, since the regulation constraint is flatter than the indifference curve. We shall assume it is also the global solution.[1]

[1]An informal but plausible argument that E is the global solution is to imagine the contrary, specifically that E lies to the left of monopoly output. If true, the regulation constraint cuts the indifference curve on the left of $R' = c$. It would then be possible to find a pair of (S,X) points with the same S such that the smaller output yielded more utility. The accompanying diagram shows this can never happen in a quasi-concave portion of the revenue function. At $R' = c$, profit is maximum for the fixed S. The regulation constraint has a smalle intercept to the extent of any discovered waste. The regulation yields two positive equilibriums; however, X_2 that exceeds monopoly output always yields more profit.



The example in figure 3-1 shows both output and perquisites increasing from the unregulated level in response to the rate-of-return regulation. This need not be the case. As usual in these problems, there are conflicting income and substitution effects. The familiar comparative statics exercise confirms that which is easier to illustrate with the diagram. Namely, a lump sum reduction in allowed revenue moves the constraint away from the utility maximum. If output and perquisites are normal goods, such an income reduction results in more consumption of each. The income effect component of reducing the regulated rate of return, then, normally induces both more output and more waste. The substitution effect is represented by a reduction in the absolute value of the slope of the regulation constraint as r is reduced. From a point such as E, holding utility constant, tightening r leads to less output and more waste. Hence, normally the effect of lowering the rate of return is to encourage more waste, since both the income and substitution effects are negative.

Output, however, is subject to conflicting income and substitution forces. For now, examine the manager's response to an increasingly tighter rate of return by supposing there is no monitoring. Figure 3-2 shows the expansion path for this case. As the rate of return is lowered from the monopoly level to marginal cost, the manager chooses points from A to Z.

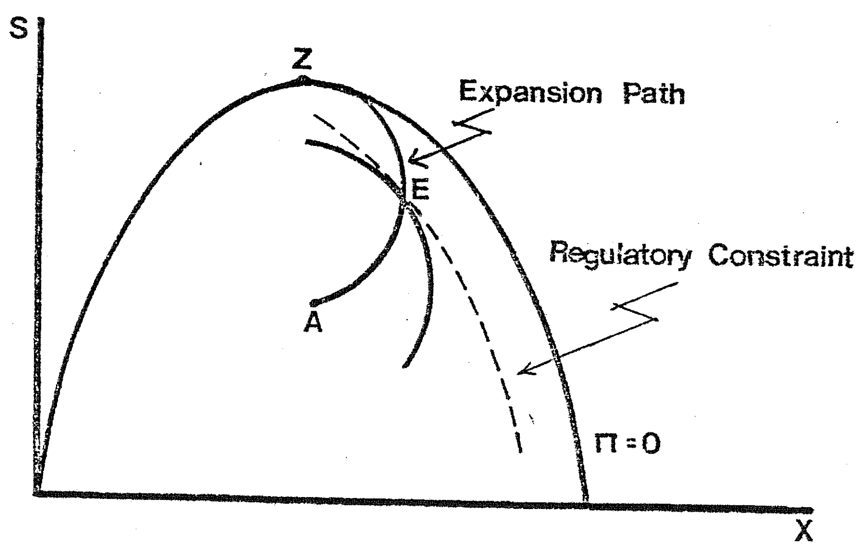


Figure 3-2 Expansion path with respect to r for M = 0

In particular, point Z satisfies two conditions. Since $r = c$ when regulation is most strict, Z must be on the zero profit locus. Second, setting r equal to c in equation (3-3a) shows that the manager chooses a point that satisfies $R' = c$, or the monopoly output level. In effect, by selecting $r = c$, the regulator establishes the zero profit locus as the constraint, and since constant profit lines are always tangent to indifference curves above monopoly output, the result is to induce the manager to behave as a monopolist in the output decision and to engage in more wasteful activity. Indeed, the added waste is equal to all of the regulated monopoly profits. Denote the unregulated monopolist's choices with an M subscript and those of the totally regulated ($r = c$) firm with a T. Then $x = x_M = s_T$ that means $R_M = R_T = R$. Let r_M denote the regulated return that is sufficiently high as to make λ zero. Then $R_M = r_M x_M + S_M$ and $R_T = cx_T + S_T$. The implication is that $(r_M - c)x = S_T - S_M$, and that the completely regulated manager consumes additional nonpecuniary benefits equal to the entire difference between monopoly return and marginal cost. Stating the conclusion more sharply, exerting maximum regulatory pressure is successful only in transferring the monopoly rents from the stockholders to the managers in the form of perquisites. The consumer does not benefit at all.[2]

---

[2]The stockholders have no way of converting managerial waste into profits. The Jensen-Meckling capital market disciplines the manager by forcing down the stock price of new shares offered as the owner reduces his share of the firm. No similar mechanism exists as the regulator reduces the allowed rate of return. In the limit, for example, if stockholders could extract profits at management's expense, these would be subject to the regulation and eliminated. Hence, stockholders have no incentive to discipline managers in response to regulatory action. The capital market, however, can protect itself against a reduction in management's ownership share, α , in the same fashion as Jensen-Meckling discussed for unregulated firms. The only difference for the regulated case is that the market must anticipate that output is likely to fall as α is reduced, whereas it remains constant in the absence of regulation. The comparative statics of the regulated expense-preference manager show that as ∂ is reduced, output declines and perquisites increase, if and only if the absolute value of the elasticity of the marginal utility of profits with respect to α is less than one.

In figure 3-2, the income effect dominates along the positively sloped segment of the expansion path, while the substitution effect becomes stronger and eventually turns the manager's behavior back toward Z. With no monitoring, the substitution effect must have this characteristic of ultimately inducing what appears to society as perverse behavior--less output and more waste. Monitoring, however, offers the regulator an opportunity to avoid such grim circumstances. This second instrument of the regulator, like the first, also has income and substitution components. The income effect (due to D_M) is ambiguous but if output and waste are normal, additional monitoring induces more of each. If, in addition to detecting more total slack, monitoring also raises the fraction of waste that is discovered, ($D_{SM} > 0$), the manager is encouraged to substitute output for waste. The slope of the regulation constraint becomes absolutely larger. Hence, monitoring normally induces more output, since both effects are favorable. Waste, however, is discouraged only by the substitution portion, since tightening the feasible region, whether by monitoring or reducing the allowed return, normally has the unfavorable result of more waste. Figure 3-3 shows several expansion paths for various fixed values of M. As M increases, the path must be to the right of previous paths, since monitoring induces output. The terminal point on the zero profit locus moves away from Z and towards C with monitoring.

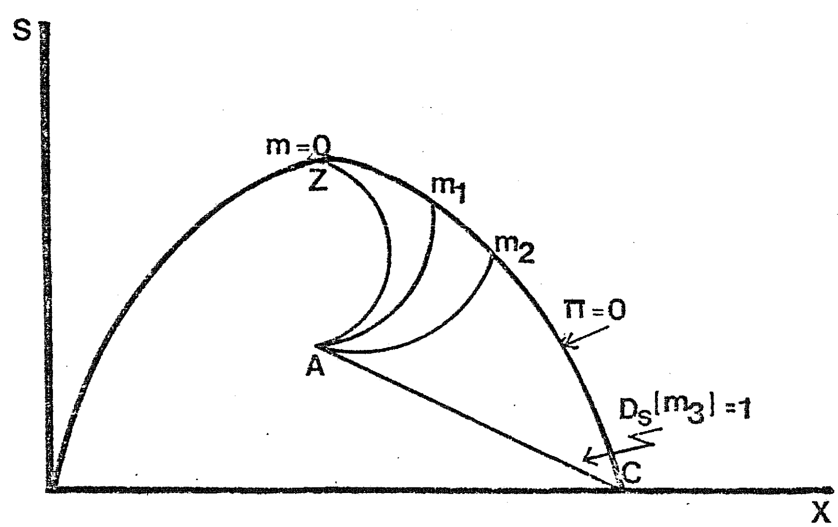


Figure 3-3 Expansion paths with respect to r for M > 0

The ultimate imaginable, but perhaps not feasible, effect of monitoring is to make the manager believe that all attempts to increase nonpecuniary benefits will be discovered. If the regulator possesses such a perfect monitor, then $D_S = 1$, and the regulation constraint is vertical. As the rate of return is reduced, the vertical constraint moves rightward from A and traces out an expansion path that in figure 3-3 is shown between points A and C. The details of this important case are shown in figure 3-4. Suppose, without any loss of generality, that waste is measured relative to that of a competitive firm that may or may not be zero. The regulated manager's equilibrium is shown as point E in figure 3-4. The expansion path is the locus of points such that indifference curves are vertical. Referring to equation (3-4a), the curves are vertical at points that satisfy $B'U_2 - U_1 = 0$. The slope of this locus is

$$\frac{ds}{dx} = \frac{(B'U_2 + (B')^2\, U_{22} - U_{12})\ (R' - C)}{U_{11} + B'U_{22} + B'U_2}$$

and its sign is negative if (1) the utility function is concave; (2) the marginal utility of perquisites is independent of profits ($U_{12} = 0$); and (3) the manager's bonus is linear in profits ($B' = 0$). These are sufficient but not necessary conditions for the unregulated monopolist utility maximization. Hence, it is plausible that a regulator with perfect information can induce more output and less waste by reducing the rate of return.

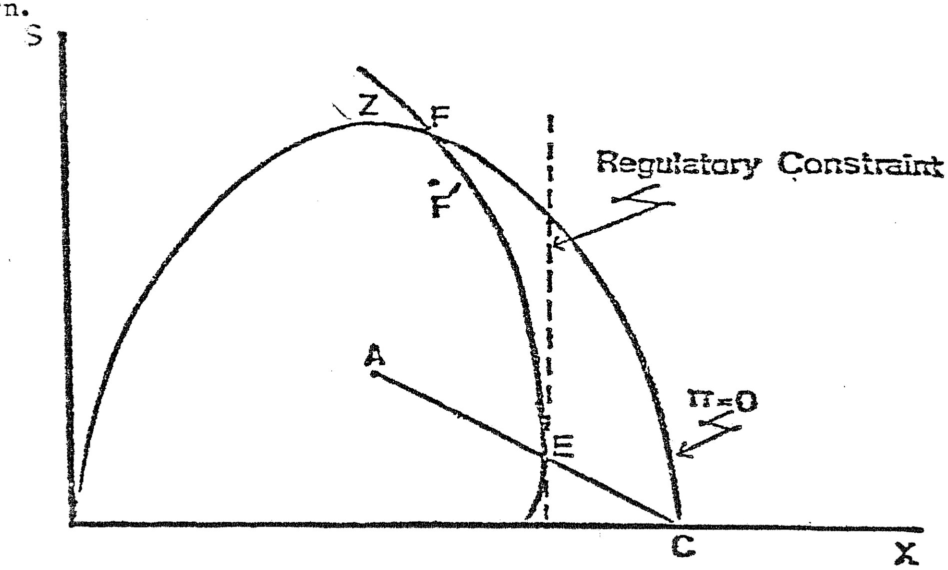


Figure 3-4 Expansion path with respect to r for extreme monitoring ($D_s = 1$)

The terminal point C in figure 3-4, in addition to representing maximum regulatory pressure, can also be interpreted as the competitive equilibrium. Competition eliminates profits; therefore, the competitive point must lie on the zero profit locus. At a point such as F, a new competitor can reduce waste to F' and thereby increase both profits and utility. Hence, the Nash equilibrium must occur at a point where profits are zero and the indifference curve is vertical. Point C satisfies these two conditions by construction. Hence, maximum rate-of-return pressure exerted by a perfect monitor forces the monopoly manager to behave as a competitor.

In practice, the regulator cannot observe the manager perfectly and consequently the expansion path terminates between Z, perverse regulation, and C, perfect regulation. The regulator's problem is to find the best mix of instruments given his limited observation powers.

Optimal Regulation

Suppose the public utility regulator wishes to maximize the public interest as traditionally measured by the sum of consumer and producer surplus plus the cost of monitoring, dM, where d is marginal cost of monitoring. The policy instruments are r and M, the fair rate of return and monitoring of waste. Letting x(r,M) and S(r,M) be the monopolist's reaction functions from the previous section, the problem is to maximize

$$\int_0^{x(r,M)} P(h)dh - cx(r,M) - S(r,M) - dM$$

while never allowing negative profits

$$R - cx - S \geq 0.$$

The Lagrangian is

$$L = \int_0^{x(r,m)} P(h)dh - x(r,M) - S(r,M) - dM$$

$$+ \Theta(R(x(r,M)) - cx(r,M - S(R,M)),$$

where Θ is the shadow price that is positive if the regulator's best choice is to eliminate profits. The first-order conditions are

$$L_r = Px_r - cx_r S_r + \ (R'x_r - cx_r - S_r) = 0 \qquad (3\text{-}5)$$

$$L_S = Px_M - cx_M - S_M - d + \ (R'x_M - cx_M - S_M) = 0. \qquad (3\text{-}6)$$

We shall assume that social welfare at the optimal r and M point in (3-5) and (3-6) is the global optimum. In particular, it must be better than the alternative--no regulation. Assuming some regulation is needed essentially means that the path from A to Z has at least a minimum amount of bulge. If the path is very close to the vertical line, regulation may bring forth insufficient output to justify the added waste. If managers have strong preferences toward quiet living, the usual discipline of monopoly profits may be superior to that of regulation. We assume it is not. If the social optimum allows positive profits, then equations (3-5) and (3-6) show that price exceeds marginal costs according to

$$P - c = \frac{S_r}{x_r} \qquad (3\text{-}7)$$

and

$$P - c = \frac{S_M + d}{x_M}. \qquad (3\text{-}8)$$

Since profits are positive and the left-hand side of (3-7) is positive, S_r/x_r must be positive and the slopes of the reaction functions are the same. Differentiate the manager's constraint [equation (3-2)] with respect to the policy instruments to find

$$(R' - r)x_r - S_r(1 - D_S) = x \qquad (3\text{-}9)$$

and

$$(R' - r)x_M - S_M(1 - D_S) = D_M. \qquad (3\text{-}10)$$

Equation (3-9) implies that both x_r and S_r cannot be positive. Hence, if there is an internal solution to the regulator's problem, equations (3-7) and (3-9) show that it occurs along a positive sloped segment of the manager's reaction curve. An example of an internal social optimum is shown in figure 3-5. The regulator chooses the highest social indifference curve from the feasible points on the manager's reaction curve. Figure 3-5 shows the tangency condition for the rate-of-return choice while holding M constant. A similar diagram could be drawn for the monitoring decision.

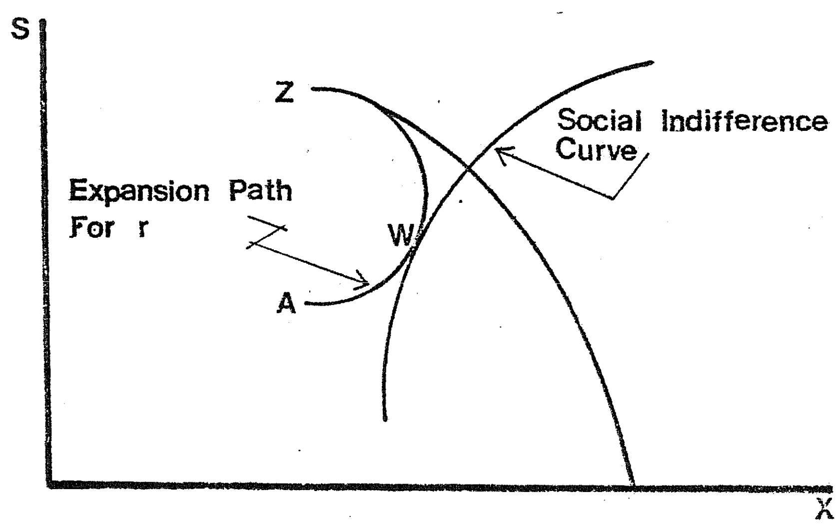


Figure 3-5 Social welfare internal solution

The optimum point, W, can be characterized in several ways. Since profits are positive, $Px > cx + S$ or $P - c > S/x$ that combined with (3-7) shows that $|\varepsilon_{Sr}| > |\varepsilon_{xr}|$, or that the elasticity of waste at the optimal rate of return exceeds that of output. The regulator sets r so as to increase output, and in doing so, forces the monopolist to a point where the output response is small. Also, equation (3-7) can be written in elasticity form as

$$\frac{(P - c)x}{S} = \frac{\varepsilon_{Sr}}{\varepsilon_{xr}}$$

This suggests that the regulator adjusts the rate of return until the ratio of two types of unproductive expenses are equal to the corresponding ratio of elasticities. The numerator is unproductive payments for output in excess of marginal cost; the denominator is waste. Monitoring continues until the same ratio of unproductive expenses is equal to the similar ratio of monitoring elasticities except that these must be corrected for the real resources absorbed by the regulator's auditing activity.

If the regulator possesses no monitoring tools at all, then the manager's expansion path ultimately returns to the monopoly output level as shown in figure 3-2. The optimal fair rate of return is clearly interior in such circumstances. The regulator never chooses to reduce r to a point where output is decreasing and waste is increasing. As monitoring becomes more and more effective, however, eventually it may be efficient to eliminate profits. In figure 3-3, perfect monitoring should clearly be combined with a regulated return that achieves the competitive point c. For intermediate monitoring technologies, $0 < D_S < 1$, whether the social optimum is at a zero profit corner or not is easily illustrated with the aid of figure 3-6. In (S,x) space, the slope of the social indifference curve is the difference between price and marginal cost. If profits are zero, $P - c = S/x$, and hence the social indifference curve is tangent to a straight line from the origin at each point along the zero profit locus. In figure 3-6, point N is not socially optimal because the expansion path intersects the zero profit boundary at a slope greater than that of the ray from the origin. There is an interior tangency that dominates point N. Point W, however, is indeed a corner solution. No other point along the expansion path can yield higher social welfare. If the expansion path is tangent to the ray from the origin, the manager's output elasticity with respect to r and waste elasticity are equal. Hence, profits are eliminated if the output elasticity exceeds that of waste at the zero profit boundary. It is clearly beneficial to continue to reduce the rate of return as long as the manager responds with more output than waste. The manager's expansion paths generally may have almost any slope. One possibility is that the paths are sloped as depicted in figure 3-3. That is, at the zero profit boundary, the path is convex if positively sloped and concave if negatively sloped. Since the zero monitoring path has a strong tendency to display such a curvature, call such examples normal. For normally sloped reaction functions, the simple geometry described above provides a final insight. If it is optimal to eliminate profits, a normal manager is forced

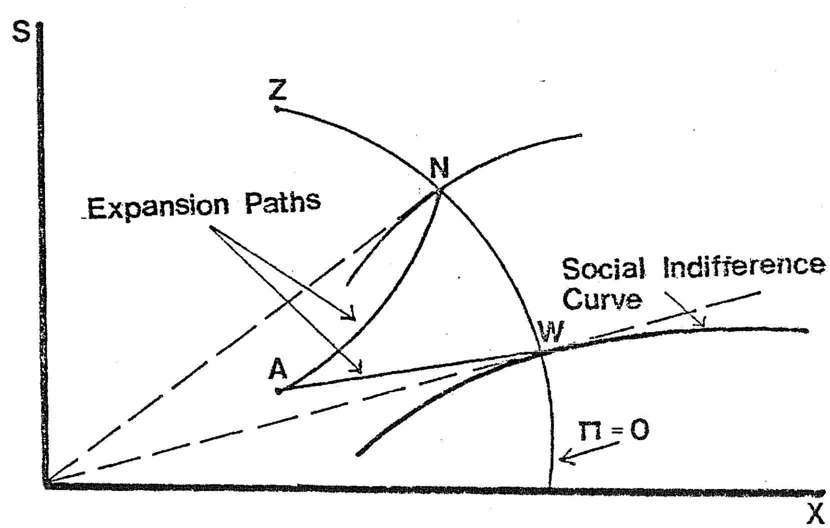


Figure 3-6 Social welfare corner solution

to consume a smaller ratio of waste to output than he would if unregulated. Figure 3-7 illustrates this. Point N is not socially efficient because the expansion path has a steeper slope than the ray from the origin. That is true for any normal path terminating between points Z and N. Between N and C, however, point W is an example of a corner solution to the regulator's problem. Since point W must lie between N and C, the manager must be consuming fewer perquisites relative to output, since all such rays lie below point A. Hence, under plausible circumstances, called normal here, in order for efficient regulation to eliminate profits, monitoring must be sufficiently accurate so that the manager can be forced to reduce waste relative to output.

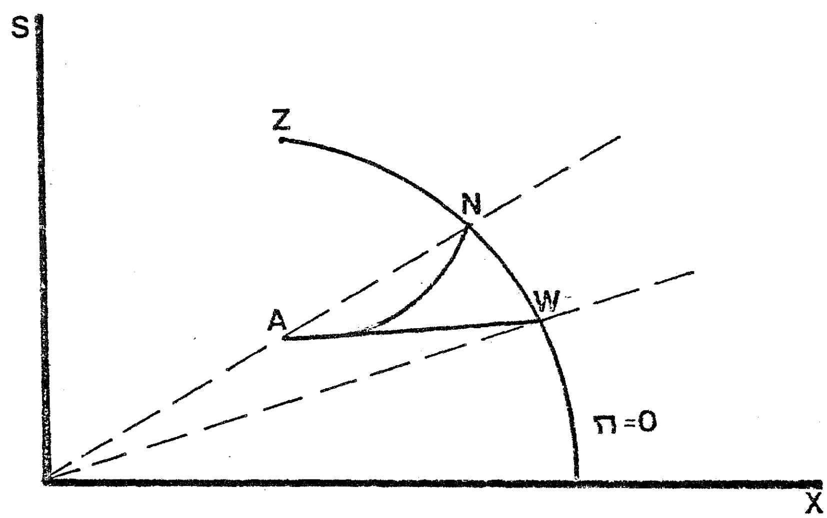


Figure 3-7 Social welfare and normal expansion paths

Summary

Historically, public utility regulation has resulted in fair rates of return in excess of the cost of capital. Theories based on profit maximization have not adequately explained this regularity. Klevorick [44] and Sheshinski [75], and in particular Bailey's [4] discussion of these two treatments of Averch-Johnson cost distortion, concluded profits should be eliminated if the utility has constant or increasing returns to scale. Since such conditions are widely believed to prevail for public utilities, it would seem that zero profits are optimal. This chapter suggests that managerial preference for an easy life may result in so much waste that the socially efficient rate of return may indeed exceed marginal cost, even for constant returns. Profit elimination is not optimal unless the regulator can carefully observe and disallow managerial waste.

Even though the technical apparatus is not of the same genre, this expense-preference model is closely related to the theory of agency. The intuition is the same as with agency problems; it is not typically possible to fashion contracts that are wholly compatible with the principal's interest. The principal usually must share a part of the profits with the agent to induce diligent behavior. Implicitly, fringe benefits, or shirking behavior, are not measured perfectly, else they would be removed from the agent's compensation. The expense-preference model is in the tradition of the optimal taxation literature where measurement problems are not explicitly considered. Clearly, the regulator lacks perfect information here, since if costs and the demand function were known, the rate-of-return form of discipline is plainly inferior to a contract that specified a large penalty if output deviated from the competitive level. The relation of Williamson's expense-preference theory to the agency literature and the costs of metering nonproductive perquisites is an interesting but currently open question.

In that context, this chapter has examined managerial reaction to rate-of-return regulation and monitoring of waste. If monitoring can detect all incremental waste, the regulator can achieve competitive conditions by reducing the allowed rate of return so as to eliminate profits.

If not, whether profits are optimally zero depends on the fraction of detected managerial waste, D_S. If D_S is zero, positive profits must be allowed to avoid waste that can be as large as unregulated monopoly profits. There is some intermediate D_S such that profits are just eliminated, ($\pi = 0$ and $\beta = 0$). For normally shaped expansion paths, the regulator's fraction of successful waste detection must be sufficiently large so that the manager consumes a smaller ratio of waste to output than an unregulated monopolist. Given the manager's substantial information advantage over the regulator, it is not surprising that the regulator allows some positive profit in order to avoid an excessive encouragement to waste. By contrast, Crew and Kleindorfer [17] showed when managerial perquisites generate revenue, and hence are at least partially productive, that output responds favorably to tighter regulation. Whether eliminating profits is socially desirable or not depends on the social value of advertising, which may be small if excessive amounts are encouraged by regulation. That issue has been avoided here, allowing a clear distinction between output, a "good", and waste, a "bad." In practice, the regulator's problem is complicated indeed, since pure waste must be detected and socially desirable output must be distinguished from that associated with excessive advertisement.

CHAPTER 4

SOME EMPIRICAL STUDIES OF EXPENSE PREFERENCE BY UTILITY MANAGERS

Chapters 2 and 3 presented several alternate hypotheses concerning utility behavior in response to various regulatory actions. These hypotheses are rooted in different assertions concerning the driving forces, the motives, and the perceptions of "environmental" pressures of, and by, utility decision makers. In general, there is an absence of empirical tests of these hypotheses. Despite a long history of regulation in the U.S., there is uncertainty about utility behavior, and this uncertainty makes it difficult to design mechanisms intended to motivate utilities to control costs.

In an effort to reduce this uncertainty, we next review five studies that test empirically the A-J thesis. Each study reviewed uses a combination of statistical and econometric techniques designed to facilitate a test of the hypothesis that regulated electric utilities, subject to the regulatory constraint, overcapitalize. The chapter highlights the theoretical foundations on which the tests were designed and discusses the methods used in testing the A-J thesis. In addition, the empirical results of estimated equations are presented, together with a statement of the investigator's major conclusion.

The results are mixed. Courville [16], Peterson [64], and Spann [76] present findings that confirm the existence of an A-J bias, while Boyes [10] and Fox [30] reject the A-J thesis. The conclusions of the five studies reviewed in this paper are representative of the mixed results found in other studies not included here. Clearly, there appears to be no consensus among economists regarding the actual existence of overcapitalization on the part of regulated electric utilities.

William T. Boyes Study

Boyes [10] examined factor demand functions using individual plant data from the electric power industry. Boyes was concerned with testing the A-J effect in plants subject to rate-of-return regulation.

Boyes asserted that the first-order conditions that result from maximizing a regulated profit-maximizing firm's objective function yield

$$K = g_K[(r - \Theta s)/R'(1 - \Theta),\ Y], \tag{4-1}$$

and

$$L = g_L(C_L/R',\ Y). \tag{4-2}$$

Equations (4-1) and (4-2) are derived using the implicit function theorem. The terms in equations (4-1) and (4-2) are defined as:

K = capital employed in the production process,
L = labor employed in the production process,
g_K = marginal product of capital,
g_L = marginal product of labor,
R' = marginal revenue,
Y = output,
C_L = the cost of a unit of labor,
Θ = the Lagrangian multiplier,
r = the cost of capital,
s = the allowed rate of return on capital.
F = fuel,
M = Maintenance

Boyes employed a four input constant ratio of elasticity of substitution (CRES) production function of the following form:

$$Y = \left\{ \alpha_K K^{-\rho_K} + \alpha_L L^{-\rho_L} + \alpha_M M^{-\rho_M} + \alpha_F F^{-\rho_F} \right\}^{-1/\rho} e^{\mu} \qquad (4\text{-}3)$$

the terms α_i, ρ_i, ρ are parameters describing the CRES production function and μ is the stochastic term.

Using equations (4-1), (4-2), and (4-3) Boyes derived expressions for ln K, ln F, ln L, and ln M. All four of these structural equations derived by Boyes allow a unique determination of the parameters. The variable Θ, in the structural equations, is a measure of the effectiveness of rate-of-return regulation. If $\Theta = 0$, the model collapses to the traditional nonregulated profit-maximizing model. The question, then, is whether or not Θ is zero.

The structural equations are estimated by a three-stage maximum likelihood technique. The data employed by Boyes consist of a sample of annual observations on 60 steam-electric-generating plants as reported in the plant reports of the Federal Power Commission. The plants were placed in operation between 1957-64; all plants used in the sample were privately owned.

Boye's results are shown in tables 4-1 and 4-2. The search for a maximum likelihood estimate of Θ produced the results, $\Theta = .02$. Table 4-1 presents selected values of Θ and the corresponding coefficients of the capital equation and calculated log-likelihoods. The likelihood ratio test does not lead to rejection of the hypothesis $\Theta = 0.0$ at even the 0.10 level.

TABLE 4-1

MAXIMUM LIKELIHOOD ESTIMATES OF THE RATE-OF-RETURN PARAMETER

			Coefficient of			
Equation		Constant	Output Term	Relative Price Term	R^2	L(Θ)
1	0.00	3.869	.6926	.2748	.9287	-32.927
2	0.01	3.872	.6927	.2743	.9291	-32.683
3	0.02	3.872	.6927	.2741	.9297	-32.101
4	0.03	3.871	.6927	.2736	.9276	-33.991
5	0.04	3.870	.6928	.2729	.9269	-34.611

Source: Boyes [10]

Given the estimated value of Θ, a straightforward application of two stage least squares gave estimates of the entire system of input demand functions. These results are shown in table 4-2. Table 4-2 shows that the fuel and capital equations exhibit a high R^2. Boyes pointed out that since fuel and capital represent about 95 percent of input costs to the electric generating industry, it is comforting that these equations fit the sample data well. Boyes observed that the empirical results imply that the elasticity of substitution between capital and fuel is much larger than between capital and labor. Furthermore, the elasticity of substitution between capital and maintenance is larger than that between labor and maintenance.

Leon Courville Study

Courville [16] investigated a cross-section of electric power plants and found evidence supportive of the A-J effect.

TABLE 4-2

TWO-STAGE LEAST SQUARES ESTIMATE OF THE INPUT DEMAND EQUATIONS AT $\Theta = 0.02$

Input	Constant	Coefficient of Output	Coefficient of Relative Price	R^2
Capital	3.872 (6.32)	.6927 (8.37)	.2741 (5.77)	.929
Labor	-.2901 (-3.90)	.3611 (0.23)	.1427 (0.33)	.241
Fuel	3.511 (10.87)	.8945 (16.71)	.3552 (6.21)	.992
Maintenance	-.219 (-.286)	.5119 (4.41)	.1791 (0.82)	.638

Note: t-values in parentheses.

Source: Boyes [10]

Courville examined the following Cobb-Douglas production function:

$$Q_i = AK_i^{\alpha} F_i^{\beta} L_i^{\lambda} e^{v_i} e^{\delta u_i} e^{bc_i} \qquad (4\text{-}4)$$

where

Q_i = the output of plant i,

K_i = the capital of plant i,

F_i = fuel consumption of plant i,

L_i = labor consumption of plant i,

v_i = a random variable distributed normally with mean 0 and variance π^2, such that $E(u_i\ u_j)_{i \neq j} = 0$.

δu_i = the capacity utilization of plant i,

c_i = the capacity of plant i.

The structural equation fitted for a cross-section of plants of a given vintage can be written as:

$$\log Q_i = \log A + \alpha \log K_i + \beta \log F_i + \delta u_i + bc_i + v_i. \qquad (4\text{-}5)$$

Labor was dropped from the estimated equation. When labor was included as an independent variable, its coefficient had the wrong sign in one case and was not significantly different from zero at the 10 percent level in any of the vintage groups.

Using electric power plant data, equation (4-5) was estimated for three vintage groups. The results are presented in table 4-3.

On the basis of the empirical evidence presented in table 4-3, Courville concluded that the existence of substitutability between capital and fuel in electricity generation is confirmed. Furthermore, the results confirm the alleged existence of returns to scale in electricity generation and indicate that there might be returns to scale to fuel alone. In all cases, the coefficient u_i was significant and had the proper sign.

Courville's test of the A-J effect may be stated as a null hypothesis in the following manner:

$$H_o = \frac{MP_K}{MP_F} = \frac{P_K}{P_F}, \qquad (4\text{-}6)$$

where MP_K and MP_F are the marginal products of capital and labor, respectively, and P_K and P_F are the unit prices of capital and labor, respectively. Equation (4-6) is the necessary condition for cost minimization. To test the null hypothesis, some manipulation of the cost-minimizing condition is necessary. Differentiating (4-4) with respect to K and F and rearranging terms yields:

$$MP_K = \alpha \frac{Q}{K}, \qquad (4\text{-}7)$$

and

$$MP_L = \beta \frac{Q}{F}$$

TABLE 4-3

REGRESSION ESTIMATES FOR EQUATION (4-5)

Vintage Group		LOG A	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\delta}$	b	R^2	D.F.
1948-50	U	-1.6255 (2.49)	0.1498 (2.54)	1.0473 (17.10)	0.2625 (1.91)	-0.0009 (2.37)	0.995	23
	D	-0.8715 (2.16)	0.1392 (2.27)	1.0682 (18.11)	0.2281 (1.58)	-0.0010 (2.53)	0.994	23
1951-55	U	-1.5657 (3.58)	0.1357 (3.31)	1.0042 (22.27)	0.2713 (2.74)	-0.0002 (1.43)	0.994	37
	D	-0.7786 (2.62)	0.1101 (2.43)	1.0269 (21.32)	0.2383 (2.27)	-0.0001 (1.18)	0.994	37
1960-66	U	-1.2602 (3.40)	0.1036 (3.10)	0.9705 (17.36)	0.3361 (3.04)	0.00002 (.96)	0.994	35
	D	0.7347 (3.46)	0.1044 (3.08)	0.9711 (17.36)	0.3372 (3.04)	0.00012 (0.13)	0.994	35

In the first column of this table, the vintage groups indicated at"U" in the second column indicates that the measure of capital was undeflated, while at "D" indicates it was deflated.

Source: Courville [16]

The ratio of the two marginal products is,

$$\frac{MP_K}{MP_F} = \frac{\alpha Q/K,}{\beta Q/F} \qquad (4\text{-}8)$$

or

$$\frac{\alpha F}{\beta K} = \frac{P_K}{P_F} \qquad (4\text{-}9)$$

To test if (4-9) holds, we can now restate the null hypothesis as:

$$\text{Ho:} \quad \hat{\alpha}\,\frac{F}{K} - \hat{\beta}\,\frac{P_K}{P_F} = 0. \qquad (4\text{-}10)$$

Estimates of α and β, $\hat{\alpha}$ and $\hat{\beta}$, can be used to test (4-10). These estimates are determined by the estimation of (4-5).

The appropriate t-statistic for use in the accept-reject decision of the null hypothesis is:

$$t = \frac{\hat{\alpha}\,\frac{F}{K} - \hat{\beta}\,\frac{P_K}{P_F} - 0}{\sigma} \qquad (4\text{-}11)$$

where

$$\sigma^2 = \frac{F_i}{K_i}^{2} \sigma_\alpha^2 + \frac{P_K}{P_F}^{2} \sigma_\beta^2 - 2\,\frac{F}{K}\,\frac{P_K}{P_F}\,\sigma_{\alpha\beta} \qquad (4\text{-}12)$$

In (4-12), σ_α^2 σ_β^2, and $\sigma_{\alpha\beta}$ are the variances of α, β, and the covariance of and β, respectively.

It is a straightforward procedure to plug in the appropriate values for the variables in (4-10), (4-11), and (4-12).

For each of 110 electric power plants used in Courville's sample, the null hypothesis is tested on the basis of the t-statistic. Using a one-tailed test, Courville rejected the null hypothesis in 105 cases, using a 95 percent confidence level.

Based upon this evidence, it is clear that Courville's results differ from Boyes's results. Courville's results strongly suggest that electric power plants were overcapitalized in the periods under examination, providing empirical evidence that the A-J effect is more than just a theoretical construct.

John Fox Study

Fox [30] proposed a test based on the coefficients of the derived demand functions for the inputs used in electric generation. An empirically testable model was developed. It distinguishes among Averch-Johnson behavior, unconstrained profit maximization, and revenue maximization. Fox used cross-section data for a selected sample of electric utility firms.

The results of Fox's test suggest there is no A-J effect because the firms used in his sample appear to behave like revenue maximizers who are constrained by their cost of capital.

In order to distinguish among the Averch-Johnson model, the unconstrained profit maximization model, and the revenue maximization model, the following notation is used:

K = capital input,
L = labor input,
F = fuel input,
s = the allowed rate of return,
Q = Q(K,L,F) = output,
= profits,
P = price of output = $P(Q,Z_i)$; Z_i demand shift parameters for Q,
R = PQ,
t = the profit tax rate,
r = cost of capital,
w = cost of labor,
f = cost of fuel.

The objective functions for each model are as follows:

The Averch-Johnson model:

$$\text{MAX } \pi = (1 - t)\ (R - wL - rK - fF) \qquad (4\text{-}13)$$

$$\text{Subject to:} \quad (1 - t)\ (R - wL - fF) = sK = 0$$

The unconstrained profit maximization model:

$$\text{MAX } \pi = (1 - t)(R - wL - rK - fF). \qquad (4\text{-}14)$$

The revenue maximization model:

$$\text{MAX } (R) \qquad (4\text{-}15)$$

$$\text{Subject to:} \quad (1 - t)(R - wL - fF) - rK = 0,$$

$$\text{or} \quad (1 - t)(R - wL - fF) - sK = 0.$$

The comparative static properties of each of the models are found by totally differentiating the first-order conditions with respect to L,K,F, and one of the parameters r,s,w,f,t, or Z_i. Cramer's rule can then be used to solve the system of equations.

To test the three models presented above empirically, the comparative static properties were used as the coefficients of the derived demand functions for the inputs used in electric generation. Assuming the derived functions are linear, they may be written as:

$$K = a_1 + \frac{dK}{ds} s + \frac{dK}{dr} r + \frac{dK}{dw} w + \frac{dK}{df} f + \frac{dK}{dt} t + \frac{dK}{dZ_i} Z_i + e_1, \qquad (4\text{-}16)$$

$$L = a_2 + \frac{dL}{ds} s + \frac{dL}{dr} r + \frac{dL}{dw} w + \frac{dL}{df} f + \frac{dL}{dt} t + \frac{dL}{dZ_i} Z_i + e_2, \qquad (4\text{-}17)$$

$$F = a_3 + \frac{dF}{ds} s + \frac{dF}{dr} r + \frac{dF}{dw} w + \frac{dF}{df} f + \frac{dF}{dt} t + \frac{dF}{dZ_i} Z_i + e_3, \qquad (4\text{-}18)$$

where a_1, a_2 and a_3 are constants and e_1, e_2, and e_3 are random error terms. Equations (4-16), (4-17), and (4-18) are the equations that Fox estimated in an attempt to determine which of the models shown in equations (4-13), (4-14), and (4-15) are consistent with the sample data.

In addition to the independent variables shown in equations (4-16), (4-17), and (4-18), the following demand shift parameters were used in the regressions:

NRES = number of residential customers,
AVJT = the average July temperature,
PPHU = the average household size in the state,
PGAS = the residential rate for 25 therms of natural gas,
VAMH = the value added per manhour in manufacturing in the state,
UPOP = the percentage of urban population in the state.

The derived demand functions were estimated by using data from steam-electric plants for 28 firms in 1967.

The objective of Fox's study was to determine which of the three models shown in equations (4-13), (4-14), and (4-15) is consistent with the data used in the test. In other words, the objective is to find which set of a priori constraints associated with each form of behavior is consistent with the data.

The estimation procedure used was an iterative Zellner method with and without the constraints on the coefficients implied by the various forms of behavior. The estimated coefficients of the unconstrained regressions are the same as coefficients estimated by ordinary least squares. The regression results are presented in table 4-4.

TABLE 4-4

LINEAR REGRESSION RESULTS

Year	Assumed Behavior	Dependent Variable	Constant	s	r	w
				Coefficients and t-Statistics of Independent Variables		
1967	None	K	2184.35 (1.015)	18.6104 (0.252)	73.1490 (0.905)	-245.325 (-1.891)
1967	None	L	-8028.11 (-1.097)	-272.649 (-1.086)	463.475 (1.685)	61.1506 (0.139)
1967	None	F	808736 (2.592)	-15167.1 (-1.417)	-29959.8 (-2.555)	12354.5 (0.657)
1967	Profit maximization	K	2045.77 (1.361)	*	-11.8214 (-0.175)	-80.3858 (-0.857)
1967	Profit maximization	L	2892.56 (0.554)	*	-80.3858 (-0.857)	103.676 (0.258)
1967	Profit maximization	F	58459.0 (0.244)	*	-13.6862 (-1.727)	-54.6735 (-1.726)
1967	Averch-Johnson	K	2657.02 (1.320)	-3.83401 (-0.056)	*	-214.795 (-1.897)
1967	Averch-Johnson	L	-4860.11 (-0.666)	-413.941 (-1.678)	*	227.860 (0.533)
1967	Averch-Johnson	F	614364 (1.952)	-5978.76 (-0.520)	*	-29.4622 (-0.957)

TABLE 4-4 (continued)

Coefficients and t Statistics of Independent Variables								
f	t	NRES	AVJT	PGAS	PPHU	UPOP	VAMH	R
-12.7148 (-1.459)	-12.5506 (-0.690)	1.35411 (17.066)	1.50655 (0.133)	50.6144 (1.119)	-303.367 (-1.000)	0.831782 (0.182)	-51.3292 (-1.491)	0.962
-33.4361 (-1.128)	59.3069 (0.958)	8.17053 (30.278)	51.3680 (1.332)	71.0041 (0.461)	-148.638 (-0.144)	4.96909 (0.319)	23.8296 (0.204)	0.988
-45.4786 (-0.035)	-3845.87 (-1.457)	158.138 (13.744)	-3358.63 (-2.042)	-2318.81 (-0.353)	22124.6 (0.503)	-1349.81 (-2.033)	-7616.65 (-1.526)	0.948
-13.6862 (-1.727)	*	1.33111 (17.244)	-1.31414 (-0.122)	45.3587 (1.181)	-354.182 (-1.231)	-0.83118 (-0.206)	-34.3869 (-1.160)	0.957
-54.6735 (-1.726)	*	8.09803 (26.474)	-8.44297 (-0.228)	221.520 (1.454)	-463.284 (-0.402)	-11.4093 (-0.702)	43.4707 (0.386)	0.981
184.650 (0.127)	*	162.423 (11.362)	-924.592 (-0.558)	-5740.75 (-0.808)	63193.5 (1.209)	-561.915 (-0.850)	-12353.9 (-2.351)	0.901
-12.1188 (-1.418)	-7.70123 (-0.456)	1.34207 (17.417)	-3.32066 (-0.344)	46.6332 (1.052)	-305.518 (-1.024)	0.836998 (0.187)	-37.0698 (-1.227)	0.960
-29.4622 (-0.957)	88.1813 (1.435)	8.09615 (29.174)	20.1022 (0.576)	46.8362 (0.293)	-154.025 (-0.143)	5.25009 (0.325)	113.626 (1.045)	0.986
-290.396 (-0.203)	-5823.70 (-2.320)	163.059 (12.671)	-1378.52 (-0.891)	-693.165 (-0.094)	22969.1 (0.462)	-1353.06 (-1.850)	-13454.4 (-2.660)	0.927

TABLE 4-4 (continued)

Year	Behavior Assumed	Dependent Variable	Constant	Coefficients and t-Statistics of Independent Variables: s	r	w
1967	"r" constrained revenue maximization	K	2225.66 (1.067)	*	66.3675 (0.895)	-241.246 (-1.928)
1967	"r" constrained revenue maximization	L	-8633.10 (-1.177)	*	562.814 (2.158)	1.39733 (0.032)
1967	"r" constrained revenue maximization	F	775075 (2.421)	*	-24433.6 (-2.146)	9030.63 (0.470)
1967	Unconstrained revenue maximization	K	1909.06 (1.377)	*	*	*
1967	Unconstrained revenue maximization	L	2036.28 (0.403)	*	*	*
1967	Unconstrained revenue maximization	F	58076.2 (0.255)	*	*	*
1967	"s" constrained revenue maximization	K	2704.13 (1.311)	-3.58006 (-0.0517)	*	-222.065 (-1.7560)
1967	"s" constrained revenue maximization	L	-4734.73 (-0.638)	-413.253 (-1.658)	*	208.527 (0.458)
1967	"s" constrained revenue maximization	F	595845 (1.722)	-6078.39 (-0.523)	*	2828.14 (0.133)

TABLE 4-4 (continued)

f	t	NRES	AVJT	PGAS	PPHU	UPOP	VAMH	R
-11.9590	-11.4971	1.35301	1.55055	46.4198	-307.426	1.11816	-47.6835	0.962
(-1.503)	(-0.668)	(17.569)	(0.141)	(1.135)	(-1.044)	(0.259)	(-1.570)	
-44.5080	43.8726	8.18661	50.7205	132.455	-89.1531	0.77343	-29.5836	0.987
(-1.590)	(0.724)	(30.223)	(1.308)	(0.921)	(-0.086)	(0.051)	(-0.277)	
-661.398	-4704.46	159.033	-3394.60	1099.65	25433.5	-1583.20	-10588.0	0.942
(-0.541)	(-1.779)	(13.449)	(-2.005)	(0.175)	(0.562)	(-2.392)	(-2.270)	
*	*	1.28368	-1.84419	13.9796	-382.710	-3.63452	-39.2419	0.944
		(16.265)	(-0.195)	(0.384)	(-1.269)	(-0.967)	(-1.293)	
*	*	7.96185	-13.5814	93.6258	-276.156	-13.2629	23.1246	0.980
		(27.683)	(-0.395)	(0.705)	(-0.251)	(-0.968)	(0.209)	
*	*	162.959	-886.116	-5323.22	62774.2	-547.596	-12321.0	0.901
		(12.545)	(-0.571)	(-0.888)	(1.265)	(-0.885)	(-2.467)	
-12.0664	-8.20554	1.34259	-3.50559	46.9249	-303.278	0.90445	-37.2209	0.960
(-1.397)	(-0.470)	(17.236)	(-0.356)	(1.047)	(-1.005)	(0.199)	(-1.219)	
29.32802	86.8385	8.09752	19.6110	47.6264	-148.084	5.42955	113.223	0.986
(-0.943)	(1.382)	(28.879)	(0.553)	(0.295)	(-0.136)	(0.331)	(1.030)	
-311.036	-5625.51	162.858	-1305.83	-807.724	22088.8	-1379.57	-13395.1	0.927
(-0.214)	(-1.920)	(12.457)	(-0.790)	(-1.074)	(0.436)	(-1.805)	(-2.615)	

Coefficient constrained equal to zero by omitting the variable from the regression.
Source: Fox [30].

The constraints on the coefficients implied by the various forms of the model are

	Restriction
(1) Averch-Johnson Model:	$\frac{dK}{dr} = \frac{dL}{dr} = \frac{dF}{dr} = 0, \frac{dL}{df} = \frac{dF}{dw};$
(2) Profit maximization:	$\frac{dK}{ds} = \frac{dL}{ds} = \frac{dF}{ds} = 0, \frac{dL}{dr} = \frac{dK}{dw}, \frac{dF}{dr} = \frac{dK}{df};$
(3) Revenue maximization:	
(a) If r is constrained:	$\frac{dK}{ds} = \frac{dL}{ds} = \frac{dF}{ds} = 0;$
(b) If s is constrained:	$\frac{dK}{dr} = \frac{dL}{dr} = \frac{dF}{dr} = 0;$
(c) No constraints:	All derivatives w.r.t. s,r,w,f, and t are equal to zero.

The consistency of a particular set of constraints with the data was tested using a chi-square test. If the ratio of the determinant of the covariance matrix of residuals of the unconstrained equation system to the determinant of the covariance matrix of residuals of the constrained equation system is less than one, then the constraints are not consistent with the data.

Using this test procedure, Fox concluded that the only type of behavior consistent with the data is revenue maximization shown in equation (4-15). Moreover, it appears that the revenue maximizing form is constrained by the cost of capital that is consistent with the constraint

$$\frac{dK}{dr} = \frac{dL}{dr} = \frac{dF}{dr} = 0.$$

The chi-square test, therefore, indicates that there is no A-J bias because the firms in the sample appear to behave like revenue maximizers constrained by their cost of capital.

H. Craig Peterson Study

Peterson's [64] findings are consistent with the A-J hypothesis. Peterson showed that as the allowed return approaches the cost of capital, costs increase and the percentage of total costs paid to capital also increases.

Peterson argued that the implicit functional form of production costs can be written as:

$$c^* = c(Q,\ P_L,\ P_F,\ P_K,\ s), \qquad (4\text{-}19)$$

where Q is a fixed level of output; P_L, P_F, and P_K are the prices of labor, fuel, and capital, respectively; and s is the allowed rate of return. Furthermore, differentiating the familiar Lagrangian expression of the A-J model yields,

$$\frac{\alpha \mathcal{L}}{\partial s} = \lambda K \qquad (4\text{-}20)$$

The first-order conditions for a guarantee that the value of the Lagrangian, $\mathcal{L}$, at the optimum values of K, L, F, and s, is always equal to the value of the objective function, -c, (for all s) are,

$$\frac{\partial \mathcal{L}}{\partial s} = \frac{\partial (-c)}{\partial s} = \lambda K \qquad (4\text{-}21)$$

and

$$\frac{\partial c}{\partial s} = - K \qquad (4\text{-}22)$$

Since $\lambda > 0$ and $K > 0$, it follows that $\frac{\partial c}{\partial s} < 0$.

Peterson contends that these results form a testable hypothesis of the A-J model. Specifically, if quantity and input prices are held constant, then the cost of production increases as regulation becomes tighter, that is, as s approaches P_K.

A general functional form is adopted for empirical analysis. In contrast to Fox, Peterson did not put a priori restrictions on the coefficients of the cost function. The sample consists of steam-generating plants that experienced a large addition to capacity just before the sample period, 1966 to 1968. The method of ordinary least squares was used in estimation.

The regression results are presented in table 4-5, together with variable definitions. The following implications are drawn:

Regression A:

In regression A, REG is significant at .01; this suggests that unit costs in states with commissions, are on average, 7 percent higher.

Signs of other coefficients are consistent with a priori expectations.

Regression B:

Regression B utilizes the generalized cost function but without the interaction terms of SS with LQ, LF, LL, and LK. The F-statistic is 0.759,

TABLE 4-5

DEFINITION OF VARIABLES USED IN PETERSON'S STUDY

REG = { 1, in states with a commission,
0, in states with no commission.

FAIR = { 1, in states determining the rate base on an "original cost" basis,
0, in states determining the rate base on a "fair value" basis.

SS = $S_e - P_e$,

where

S_e = the allowed rate of return on equity capital, and

P_e = the cost of equity capital.

(Variable SS is adjusted for capitalization; the data used in determining SS are those for the firm owning the plant under observation.)

LCOST = log of unit cost of production,

PC = percent of total unit cost going to capital,

Y = year of observation,

TC = index of technological change,

LQ = log of quantity produced,

LL = log of annual wage rate,

LK = log of capital rental rate,

LF = log of fuel price,

NH = dummay variable: Value equals one if firm is part of a holding company,

LQLQ = log of quantity times log of quantity,

LQLL = log of quantity times log of wage rate,

LQLF = log of quantity times log of fuel price,

LQLK = log of quantity times log of capital rental rate.

Regression Results:

Regression A

$$LCOST' = -\ 0.1717 + 0.0880NH - 0.1008LQ' + 0.0590LQ'LQ'$$

(3.85) (-8.11) (8.42)

$$+\ 0.6237LF' + 0.1784LL' + 0.2680LK' - 0.0189TC - 0.0006Y$$

(15.26) (2.75) (2.18) (-4.45) (-0.06)

$$+\ 0.0701REG + 0.0013FAIR$$

(3.04) (0.08)

Number of observations: 168
$R^2 = 0.82$.

Regression B

$$LCOST' = -0.0185 - 0.1160LQ' + 0.0409LQ'LQ' - 0.0127TC$$

(-7.81) (4.19) (-3.20)

$$-0.0020Y + 0.5299LF' + 0.3260LL' + 0.4088LK' - 4.763LK'LK'$$

(-0.23) (9.15) (4.79) (2.44) (-3.29)

$$-0.8516LF'LF' + 0.6055LL'LL' + 1.704LF'LK' - 1.277LF'LL'$$

(4.04) (1.36) (2.76) (-4.69)

$$-0.0846LF'LQ' - 5.874LK'LL' + 0.4102LK'LQ' + 0.1943LL'LQ'$$

(1.06) (-5.88) (2.00) (1.90)

$$-3.625SS' - 275.1SS'SS'$$

(-3.73) (-4.93)

Number of observations: 141
$R^2 = 0.87$.

Regression C

$$PC = -0.0502 - 0.0114LQ' + 0.0132LQ'LQ' - 0.1510LF'$$

(-2.40) (4.89) (-9.63)

$$+0.0675LL' - 0.0037LK' - .0003TC - .0083Y$$

(2.71) (-0.29) (-0.20) (-2.30)

$$+0.0223NH + .0267REG + .0026FAIR$$

(2.53) (3.02) (0.38)

Number of observations: 168
$R^2 = 0.49$.

Regression D

$$PC = -0.0036 - 0.0095LQ' + 0.0132LQ'LQ' - 0.1878LF'$$

(-1.80) (4.62) (-9.09)

$$+\ 0.1023LL' - 0.0253LK' + 0.0013TC - 0.0090Y - 1.336SS'$$

(3.53) (-0.41) (0.70) (-2.20) (-2.96)

Number of observations: 141
$R^2 = 0.50$.

implying that the joint hypothesis that the coefficients of these terms are zero cannot be rejected. All linear terms in input prices are positive and significant, however.

Regression C:

In regression C, Y and TC are included to allow for year-to-year and technology effects, while NH adjusts for differences attributed to holding companies. REG and FAIR are both positive, and REG is significant at the .01 level. This implies that tighter regulation is associated with a greater proportion of cost going to capital. Other coefficients, both the linear and quadratic terms in quantity, are significant. Peterson states that this suggests there are savings on the use of capital in comparison to other inputs up to some point. The fuel price coefficient is negative and significant. The coefficient of the wage rate is positive and significant. The capital price coefficient is negative but not significant. Peterson suggests that this is due to the fact that high-rental rates on capital increase the cost of using a given amount of capital, but this is offset by substitution of other inputs (primarily fuel) as capital becomes more expensive. Furthermore, regression C suggests that technology has been neutral among inputs, since newer technology in a plant seems to have no effect on the share of cost going to capital.

Regression D:

When the variable SS is included in the equation, as a determinant of the effect of regulation on the proportion paid to capital, the findings are similar to those of regression C. In essence, as regulation becomes more stringent, the firm spends a greater fraction of total cost on capital, evidenced by the fact that SS is negative and significant at the .01 level.

Regressions C and D support the A-J hypothesis. Peterson concludes that as regulation tightens, by reducing the allowed rate of return, unit costs increase. This result is substantiated by the coefficient associated with REG and the coefficient associated with SS. Peterson also found that the percentage of cost going to capital increases with more stringent regulation.

The conclusions here are consistent with those reported by Courville.

Robert M. Spann Study

Spann [76] tested the A-J hypothesis using the trans-log production function and data from regulated electric utilities. Spann's test, like those of Courville and Peterson, confirms the A-J hypothesis.

The trans-log (transcendental logarithmic) production function is essentially a Taylor series expansion of any arbitrary production function[1] that may be written as:

$$\log Q = \alpha_0 + \beta_1 \log K + \beta_2 \log L + \beta_3 \log F + \beta_4 [\log K]^2 + \beta_5 [\log L]^2 + \beta_6 [\log F]^2 + \beta_7 \log K \log F + \beta_8 \log \mu \log L + \beta_9 \log F \log L, \quad (4\text{-}23)$$

where

Q = output,

K = capital,

L = labor,

F = fuel,

α = the constant term,

$\beta_i, i = i, \ldots 9$, are the function's coefficients.

[1]The trans-log production function was derived by Christensen, Jorgenson, and Lau; see "Conjugate Quality and the Transcendental Logarithmic Production Function," Paper presented at the Second World Congress of the Econometric Society, Cambridge, England, September 1970.

Spann derived the following two estimable equations:

$$\mu_K = b_1 + b_2 \log K + b_3 \log F + \log L + \lambda z \tag{4-24}$$

and

$$\mu_F = b_3 + b_6 \log K + b_7 \log F + b_8 \log L, \tag{4-25}$$

where

$\mu_K = \frac{rK}{PQ}$, the required payments to capital as a fraction of total revenue,

$\mu_F = \frac{fF}{PQ}$, the payments to fuel as a fraction of total revenue,

$z = \frac{sK}{PQ}$, the allowed payments to capital as a fraction of total revenue.

Equations (4-24) and (4-25) form a system of two equations that can be used to estimate . All of the parameters of equations (4-24) and (4-25) are not functionally independent. Dividing b_6 by b_3, and rearranging terms, yields

$$(1 - \lambda)\, b_6 = b_3 \tag{4-26}$$

Spann contends that (4-26) is a constraint on the two-equation system for factor shares that is implied by profit maximization subject to the regulatory constraint.

Spann points out that the A-J thesis can fail to apply to regulated firms if either or both of two conditions hold:

(1) the regulatory constraint does not enter the firm's objective function, or (2) regulated firms do not maximize profits.

Condition (1) implies the following testable hypothesis:

$$H_1: \quad \lambda = 0$$

Condition (2) implies

$$H_2: \quad (1 - \lambda)b_6 \neq b_3$$

Two approaches can be used to test for overcapitalization. First, equations (4-24) and (4-25) can be estimated jointly, subject to the constraint, $(1 - \lambda)b_6 = b_3$, and by testing the hypothesis $\lambda = 0$. Second, equations (4-24) and (4-25) can be estimated jointly, subject to 0, and testing the hypothesis $(1 - \lambda)b_6 = b_3$.

Two sets of data were used by Spann in his empirical analysis. The first set consisted of observations on the first year of operation for all new steam-electric plants built by regulated firms between 1959 and 1963. This sample consisted of 35 plants. The second set of data consisted of a select sample of large electric companies in 1963. This sample was restricted to firms for which nonhydroelectric plants represented at least 90 percent of generating capacity, at least 90 percent of kilowatt-hour sales were generated by the firm instead of purchased from others, and no more than 10 percent of kilowatt-hour generation was resold to other power companies. The sample set, in this case, consisted of 24 regulated companies.

Table 4-6 shows Spann's ordinary least squares estimates of equations (4-24) and (4-25), subject to the constraint given by equation (4-26). The subscripts 1 and 2 on capital refer to whether megawatt capacity or total assets were used to estimate the size of the capital stock. The plant regressions were weighted by the inverse of the standard errors of estimate of the unconstrained fuel and capital share equations, since these variances were significantly different.

TABLE 4-6

RESTRICTED REGRESSIONS USED TO TEST THE A-J THESIS*

FIRM DATA

(1)
$$s_K = 0.3796 - 0.03093 \log K_1 + 0.0348 \log L - 0.00865 \log F + 0.6848Z$$
(0.02657) (0.01965) (0.01145)

$$s_F = - 0.00682 - 0.02744 \log K_1 + 0.00389 \log L + 0.02870 \log F$$
(0.03632) (0.01975) (0.03489)

$R_2 = 0.6506$ SSE = 0.04789

(2)
$$s_K = 0.11233 - 0.01493 \log F + 0.01028 \log L + 0.007623 \log K_2 + 0.66174Z$$
(0.01101) (0.02228) (0.029503)

$$s_F = 0.35298 - 0.044164 \log K_2 + 0.014917 \log L + 0.03067 \log F$$
(0.032556) (0.022195) (0.028535)

$R^2 = 0.5664$ SSE = 0.049371

PLANT DATA

(3)
$$s_K = 0.95150 - 0.010259 \log L + 0.05670 \log K_1 - 0.06614 \log F + 0.56504Z$$
(0.02914) (0.028520)

$$s_F = - 1.2388 - 0.15208 \log K_1 + 0.15241 \log F - 0.01781 \log L$$
(0.05853) (0.044143)

$R^2 = 0.7849$ SSE = 0.2180

(4)
$$s_K = 0.69630 - 0.04258 \log L + 0.09481 \log K_2 - 0.08245 \log F + 0.50094Z$$
(0.02884) (0.02637) (0.19976)

$$s_F = 0.23859 - 0.1649 \log K_2 + 0.13371 \log F + 0.03412 \log L$$
(0.03995) (0.03858) (0.04467)

$R^2 = 0.75582$ SSE = 0.18381

*SSE = Sums of error squared. SSE and R^2 refer to the sums of error squared about both regression equations. Standard errors are in parentheses.

Source: Spann [74]

Test of Hypothesis #1:

First a chi-square statistic is calculated as follows:

$$\chi_1^2 = - T \log \frac{L(HA)}{L(Ho)}$$

where

T = the number of observations,

$L(HA)$ = the sum of errors squared under the alternative hypothesis,

$L(Ho)$ = the sum of errors squared under the null hypothesis

$\lambda = 0$.

Acceptance of this null hypothesis would indicate that the Lagrangian multiplier for the rate-of-return constraint was zero. This would imply that regulation is not effective, and the A-J effect is inoperative.

Table 4-7 shows the chi-square statistic computed for both plant and firm data. Table 4-7 shows that λ is significantly different from zero at the 0.01 level in all cases. Furthermore, $\bar{\chi}$ is within the range predicted by the model ($0 \leq \chi \leq 1$). Thus, the A-J thesis cannot be rejected.

Table 4-8 shows that the χ^2 statistics are below the critical level when K_1 is the measure of the capital stock but when K_2 is not used. Spann concluded that this confirms the A-J thesis in the plant data but raises some doubts concerning the validity in the firm data.

Spann's findings are consistent with Courville's findings and Peterson's findings. In Spann's paper, the A-J thesis is expressed as a set of nonlinear restrictions on the factor share equations. The A-J thesis was accepted by Spann in almost all cases.

TABLE 4-7

χ^2 STATISTICS FOR TESTING $\Theta = 0$

Data Set	Measure of Capital Stock	Estimated	To Test $\Theta = 0$
Plant	K_1	0.5650	26.48*
Plant	K_2	0.5009	25.77*
Firm	K_1	0.6848	43.98*
Firm	K_2	0.6617	34.43*

*Significant at 0.01 level

Source: Spann [76]

TABLE 4-8

χ^2 STATISTICS FOR TESTING PROFIT-MAXIMIZATION RESTRICTION GIVEN EFFECTIVE REGULATION

Data Set	Measure of Capital Stock	χ^2
Plant	K_1	0.987
Plant	K_2	1.414
Firm	K_1	3.586
Firm	K_2	9.877*

*Significant at 0.01 level

Source: Spann [76]

* * *

It is obvious from the above three chapters that neither the theory nor empirical findings are sufficiently detailed or conclusive to permit policy prescriptions concerning new regulatory practices. This conclusion is perhaps somewhat too conservative in light of the fact that almost all past institutional reorganizations were preceded by an absence of perfect knowledge.

Furthermore, while our knowledge of utility behavior is imperfect, it is suggestive of future research needs. Chapters 2, 3, and 4 lead to the conclusion that if the current practice of regulation is viewed as a system of incentives, then the productivity of this system is not fully known. The following chapter outlines a variety of other specific incentive mechanisms that have been tried out and that could be tried out. Chapter 5 suggests another line of research that can suggest a direction for experimentation.

CHAPTER 5

EXAMPLES OF INCENTIVE MECHANISMS APPLICABLE TO REGULATED FIRMS

Classification of Incentive Mechanisms

While the current practice of rate-of-return regulation can be thought of as a particular system of incentives, there are many other regulatory practices that could be introduced to motivate specific performance by utilities. There are undoubtedly several ways to group such incentive mechanisms. For purposes of this study, mechanisms are classified according to the source of information flows they generate and the reward structure by which agents are paid. The types of possible mechanisms discussed in subsequent sections of this chapter are indicated in table 5-1. It is noteworthy, however, that the assignment of practices to classes of incentives is somewhat arbitrary and suggestive at best.

TABLE 5-1

CLASSIFICATION OF INCENTIVE MECHANISMS

	Source of Information	
Reward Structure	Agent Only	Agent and Others
Agent's productivity only	A	B
Total system productivity	C	D

Incentive mechanisms falling into categories A and C use information that is acquired from the agent directly, that is, behavior is to be motivated; while mechanisms falling into categories B and D use information acquired from the agent as well as from sources other than the agent. Incentive mechanisms are designed to reward the agent, based upon either the productivity of the agent alone, that is, mechanisms falling into categories A and B, or the productivity of the agent plus additional measures of performance, that is, mechanisms falling into categories C and D.

Type A Mechanisms

The Incentive Rate-of-Return Mechanism (IROR)

This mechanism allows the company to realize a rate of return on common equity based upon a cost performance ratio (CPR) calculated in the following manner:

$$CPR = \frac{\text{Actual Construction Costs}}{\text{Estimated Construction Costs}} \qquad (5\text{-}1)$$

The CPR is adjusted for scope changes on the project and for inflation, when deemed appropriate. A hypothetical IROR schedule is shown in figure 5-1.

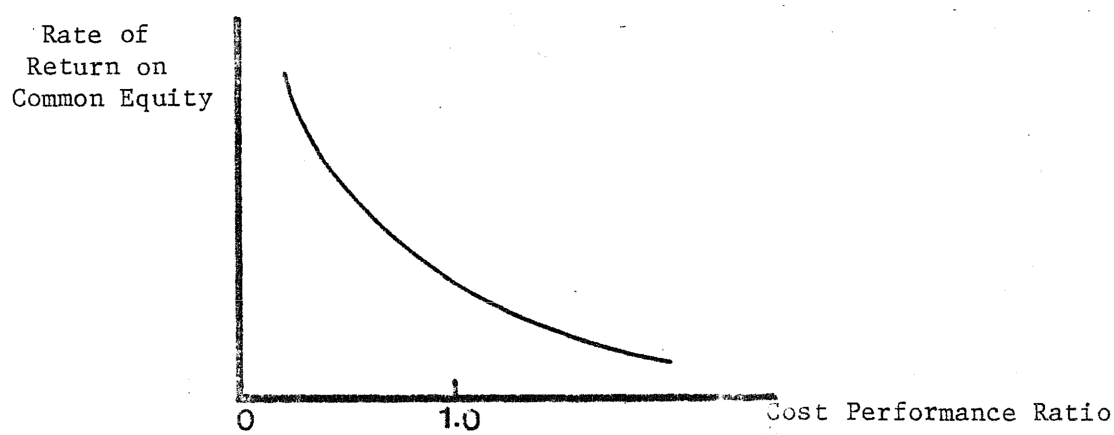


Figure 5-1 Hypothetical IROR incentive mechanism

Figure 5-1 shows that the rate of return on equity is a decreasing function of the CPR. The CPR measures the success of management in controlling costs, which is a direct result of the manner in which the ratio is calculated, that is, the ratio compares "actual" to "estimated" costs.

This mechanism was devised by the Federal Energy Regulatory Commission, FERC, for use on the Alaska Natural Gas Transportation System construction project. Two principal criteria were identified in designing the IROR.[1] First, increases in costs due to economic factors outside the control of management should not be included in the CPR. Thus, the actual cost of construction is carefully calculated to purge the CPR of these exogenous cost increases. Second, cost increases due to project delays can be attributed to management errors.

In this mechanism, the burden of revealing accurate information rests ultimately on the agent. The allowed rate of return on equity is clearly a function of the agent's assessment of expected costs. Information the agent possesses ex ante, which is the basis for his prior beliefs regarding project costs, will ultimately determine the agent's reward. However, without additional constraints imposed on the agent, there is an unambiguous incentive for the agent to inflate estimates of costs. This would have the effect of biasing the CPR downward, resulting in a higher allowed return on equity than would be possible if the agent revealed the "truth." This illustrates the effect that informational asymmetry can have on ex post rewards. Only if additional devices are integrated into the incentive mechanism to penalize the agent somehow for revealing "false" information, can this mechanism achieve its objective. For a discussion of the effect of information on decisions, see appendix A.

The IROR does, however, simplify the normal review process. Two additional reasons have been identified for its adoption. First, if the project undertaken is perceived to have high risk, suppliers in the capital markets will require higher interest payments on debt capital; by establishing an incentive rate of return, the lenders may perceive that the pipeline company will exert efforts to keep cost overruns in control. Second, the incentive rate of return is alleged to induce the necessary flow of needed equity capital and to provide incentives for the owners to perform efficiently under the terms of the contract. Exactly how people active in the capital markets view this mechanism is not known. Intuitively, it seems that the capital markets would place some value on

---

[1]These criteria are found in [41,pp. 18-19]. The discussion presented in this section is taken largely from this source.

the potential improvements in efficiency that this mechanism could generate.

The Managerial Labor Market As an Incentive Mechanism

In a recent article by Fama [25], the need to implement incentive mechanisms is seriously questioned. Fama contends that the managerial labor market is capable of inducing managers of both regulated and non-regulated firms to choose acts that are in the best interests of security holders.

The main thesis presented in Fama's work arises from his belief that management and risk bearing should be considered separate factors of production. The set of contracts, or the "firm", is disciplined by competing firms, and each factor of production is disciplined by the opportunities provided by the markets for its services. Previous contributions of a manager to firm productivity are signals to the managerial labor market used to determine his opportunity wage. The previous associations a manager has had with firms, and his resulting successes or failures, give the market the information that is needed to assess accurately the manager's productivity, consequently allowing the market to determine the rental rate for the manager's human capital. Fama contends that self-interest gives the manager a stake in the success of the management team to which he belongs currently.

Fama recognizes that although the managerial labor market may be able to base an opportunity wage on a manager's past performance, there is still the uncertainty about the manager's choice concerning appropriation of nonpecuniary benefits over a period of time. This may result in a deviation from contract between the manager and other factors of production. Specifically, a manager consumes resources through shirking, incompetence, or consumption of perquisites to the point where marginal-expected utility is equal to the additional dollar of wealth that may be used outside the firm. This causes the firm's value to be less than maximum.

When the manager is also the firm's sole security holder, he cannot avoid full ex post settling up with himself; that is, he must pay for his deviation directly. Fama also states that some form of ex post settling up must also exist when the manager is not the sole security holder if incentive problems are to be avoided. Briefly, this is accomplished in Fama's world by assuming that the following two conditions are operative: (1) the managerial labor market is efficient in that it processes current and past information to revise future wages, and furthermore, the market understands any enforcement power inherent in the wage revision process, and (2) full control of managerial behavior through wage changes is accomplished by assuming that the weight of the wage revision process is sufficient to resolve managerial incentive problems. In short, rational managerial labor markets are cognizant of shortcomings in available mechanisms for enforcing ex post settling up, and assessments of deviations from contract will be incorporated into contracts on an ex ante basis, presumably through wage adjustments. Wage adjustments, or revaluations of managerial human capital, are a form of full ex post settling up if it is assumed that the manager perceives the present value of likely changes in his human capital to be at least as great as the cost of his deviations from contract.

The stochastic process responsible for generating the manager's measured product is presented by Fama as,

$$\tilde{Z}_t = Z_t + \tilde{\varepsilon}_t, \tag{5-2}$$

where $\tilde{Z}_t$ is the uncertain end-of-period measured marginal product (MP), Z_t is the expected value of the MP, and $\tilde{\varepsilon}_t$ is random noise. Equation (5-2) is a random walk plus white noise. Writing it in its inverted form yields,

$$\tilde{Z}_t = (1 - \Phi)Z_{t-1} + \Phi(1 - \Phi)Z_{t-2} + \Phi^2(1 - \Phi)Z_{t-3} + \ldots + \tilde{\varepsilon}_t \, , \tag{5-3}$$

and

$$\overline{Z}_t = (1 - \Phi)Z_{t-1} + \Phi(1 - \Phi)Z_{t-2} + \Phi^2(1 - \Phi)Z_{t-3} + \ldots, \qquad (5\text{-}4)$$

where $\overline{Z}_t$, the manager's expected MP, equals his current wage.

Equation (5-3) expresses conveniently the current measured MP as an autoregressive process of infinite order (AR(∞)). From the preceding discussion, it should be clear that equation (5-3) has the desired properties promulgated by Fama. The current MP is generated by a weighted average of past observations going back an infinite number of periods. Equations (5-3) and (5-4) have an infinite order, while a manager has a finite working life. This discrepancy causes no problem as long as the manager's current MP is almost fully absorbed by the stream of wages over his future working life.

Equivalently, the expected MPs shown in equation (5-4) are adjusted on the basis of all past deviations of MPs from their expected values, that leads to a precise form of full ex post settling up, by which

> ...any potential managerial incentive problems in the separation of riskbearing, or security ownership, from control are resolved. The manager can contract for and take an optimal amount of consumption on the job. The wage set *ex ante* need not include any allowance for *ex post* incentives to deviate from the contract since the wage revision process neutralizes any such incentives.[2]

Equations (5-3) and (5-4) also show that full ex post settling up for deviations from contract during period t, need not necessarily occur at the end of period t. The value of Φ $(0 < \Phi < 1)$ will determine how the wage revision process allocates past MPs across periods. The point is, however, that all MPs (hence, deviations from contract) are accounted for in the manager's current wage.

---

[2] Taken from [25, p. 301].

Fama's analysis provides a valuable framework for conceptualizing the manner in which information regarding a manager's MP might be absorbed by the managerial labor market and used in assessing the manager's future performance. In Fama's world, a full ex post settling up will result, assuming the manager perceives he will pay for his deviation from contract sometime in the future, via a process of wage adjustments.

Fama's analysis is presented in this section because it provides an insightful contrast to the main stream of thought regarding the principal-agent relationship. To date, no explicit tests of Fama's hypothesis have appeared in the literature. Of course, in the context of regulation, the interest is not in the firm's value maximization, but social welfare; and furthermore, the relationship between the firm's security holders and managers is altered by the presence of regulation. This is further explored in chapter 4 below.

Collection Mechanisms in Fuel Adjustment Clauses

Some states impose a lag in the collection of revenues through fuel adjustment clauses. Florida, for example, requires a company to wait two months before it can collect revenues generated by fuel adjustments; North Carolina has imposed a three-month lag.[3]

The primary incentive created by lagged recovery of revenues is improved asset management. Moreover, the company has the incentive to maintain efficient cash flow management by controlling fuel costs. The fuel adjustment clause, without lagged recovery, has a tendency to weaken the utility company's efforts to adopt cost control innovations. These weakened incentives are a result of the reduced penalties incurred by utilities for poorly managing fuel costs and maintaining baseload plants.

---

[3]The lags are identified in [41, p. 35]. Much of the discussion presented here is taken from the same source.

The implementation of lagged recovery in some states suggests that regulators are aware of the weakened incentives arising from fuel adjustment clauses. Most states and the FERC include purchased power costs in adjustment clauses; this is to provide utilities with incentives to purchase cheaper power when it is available. Here, as in the above discussed mechanisms, there is a limited amount of empirical research that tests various assertions.

Type B Mechanisms

Cost-Sharing Contracts

A cost-sharing contract can be represented as follows:

$$\pi_0 = \pi_N + \gamma(C_a - C_0) \tag{5-5}$$

where

π_0 = the final contract profits,
π_N = the estimated (negotiated) contract profits,
γ = the cost-sharing rate,
C_a = actual costs,
C_0 = estimated costs.

Equation (5-5) indicates that net profits realized by the agent are determined by two components: (1) the profits established ex ante as being a fair and reasonable return on assets, and (2) an adjustment based upon deviations of actual costs from estimated costs. The cost-sharing rate, , takes a value from 0 to 1, inclusive and is determined by the parties to a contract prior to its enforcement. Thus, γ determines the risk that each party will bear during the enforcement of the contract. Asymmetric information on the part of the negotiating parties could lead clearly to a non-optimal-cost-sharing rate; for example, one party might be forced to accept a level of risk, generating a situation where the marginal cost (adjusted for risk) exceeds the marginal benefit (adjusted for risk). This will lead to a non-Pareto optimal solution.

This type of contract is sometimes written between utility companies and construction companies. Plant expansion efforts often lead to cost overruns that are shared by the utility and the construction company in predetermined proportions. The utility has a disincentive to control costs if it can easily include costs, resulting from poor management, in construction work in progress (CWIP). Thus, the regulatory authority must take steps to monitor the firm's activities and to implement other types of incentive mechanisms designed to overcome these potential cost overruns.

Insurance Contracts with Deductible Clauses

An insurance policy with a deductible clause is appropriately characterized as a risk-sharing contract. Generally, this type of contract stipulates that the insured party is responsible for paying damages below some stipulated amount; the company will only pay for damages when they exceed the cutoff point.

The deductible clause is analogous to a truncated probability distribution. Figure 5-2 presents a hypothetical probability distribution that indicates that the insured party will pay for damages below p^*, the truncation point. Asymmetrical information can easily lead to a nonoptimal p^* for one or both parties.

The deductible clause gives the insured party an incentive to avoid damages. Furthermore, it is common for an insurance company to give lower insurance rates to an insured party whose past record, or current circumstances, indicates a higher probability that damages will be avoided.

This type of incentive mechanism currently has wide appeal in the insurance industry, evidenced by the large number of companies that include deductible clauses in insurance contracts.

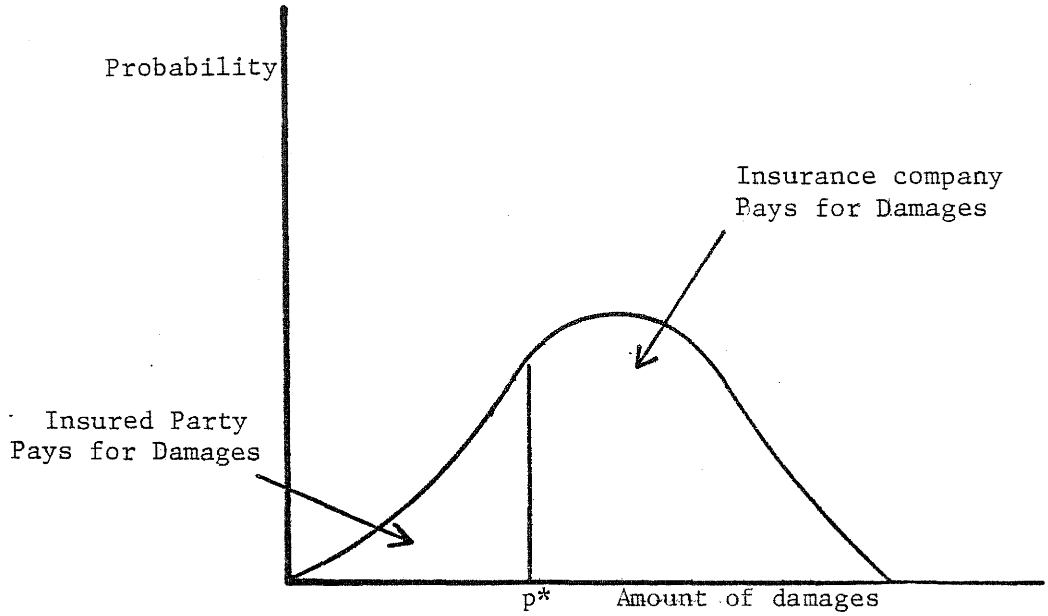


Figure 5-2 Hypothetical probability distribution characterizing an insurance policy with a deductible clause

Type C Mechanisms

Managerial Incentives and Corporate Debt Policy

Both mechanisms reviewed in this section focus on the existence of asymmetric information between "insiders" and "outsiders." The term insiders refers to corporate managers, while outsiders refers primarily to investors, and for our purposes, the regulatory authority. Managers, possessing inside information, convey information to outsiders about the firm's business risk and profitability by a process known as financial signaling. To induce valid financial signals it is necessary to establish incentive mechanisms for managers. Ross [72] has suggested one such incentive mechanism, or "schedule", with the following structure:

1. At time 0, the manager gets paid a wage that is proportional to the current market value of the firm, V_0; and
2. At time 1, he receives compensation that depends on the terminal value of the firm, $\tilde{X}$.

The total compensation to the manager is

$$\tilde{M} = \gamma_0 V_0 + \tilde{Y}^m, \tag{5-6}$$

where

$$\tilde{Y}^m = \begin{cases} \gamma_1 \tilde{X}, \text{ if } \tilde{X} \geq \hat{y}, \\ \gamma_1 \tilde{X} - L, \text{ if } \tilde{X} < \hat{y}, \end{cases} \tag{5-7}$$

and

$\tilde{M}$ = the manager's uncertain end-of-period compensation,

V_0 = the current market value of the firm,

γ_0 and γ_1 = constants,

L = the penalty assessed to the manager if the firm goes bankrupt at the end of the period,

$\hat{y}$ = the payments promised to the firm's debt holders at the end of the period.

In this model, the key to the notion of financial leverage signaling is the term L, the penalty imposed on the manager if bankruptcy occurs. In a nonregulated industry, an increase in $\hat{y}$, higher leverage, implies a higher probability that the manager will have to pay the penalty and receive a smaller percent value of compensation in time 2.[4] In a regulated industry, higher leverage may impose less threat of bankruptcy through adjustments in the firm's allowed rate of return. Nevertheless, the regulatory authority and investors will observe the level of debt in the firm's capital structure and monitor the signals received by the firm's managers regarding their ability to service the debt payments, $\hat{y}$. Since investors and the authority know that it is in the self-interest of the manager to avoid the penalty, leverage adjustments have informational content; a leverage increase has positive informational content, while a decrease in leverage may carry some negative content.

[4]This interpretation is found in [13].

With asymmetrical information, the present market value of the firm, V_0, is determined by the participants in the capital markets. It is the capital market's perception about the firm's future prospects of solvency that determines V_0. Hence, an increase in y will increase V_0, which causes the present value of the penalty, $\Delta V(L\tilde{b})$, to increase, while on the other hand, it causes an increase in the manager's current wage, $\gamma_0 V_0$. The term $\tilde{b}$ is one if $\tilde{X} < \hat{y}$ and zero if $\tilde{X} \geq \hat{y}$. The optimal leverage to the manager is the Y*, at which $\Delta V_0/\Delta \hat{y}^* = \Delta V(Lb)/\Delta \hat{y}^*$ that indicates the marginal increase in the manager's current wage equals the marginal increase in the present value of the penalty.

Ross argues that V_0 will equal the pure value of investment, $V(\tilde{X})$. For this to be true, the penalty must go to bondholders in the event of bankruptcy. The gross returns to stockholders and bondholders may be expressed as:[5]

$$\tilde{y}^S = \begin{cases} (1 - \gamma_1)\tilde{X} - \hat{y}, & \text{if } \tilde{X} \geq \hat{y}/(1 - \gamma_y), \\ 0 & \text{, if } \tilde{X} < \hat{y}/(1 - \gamma_1), \end{cases} \tag{5-8}$$

and

$$\tilde{y}^B = \begin{cases} \hat{y} & \text{, if } \tilde{X} > \hat{y}/(1 - \gamma_1), \\ (1 - \gamma_1)\tilde{X} + L, & \text{if } \tilde{X} < \hat{y}/(1 - \gamma_1), \end{cases} \tag{5-9}$$

where $\tilde{y}^s$ is gross returns to stockholders and $\tilde{y}^B$ is gross returns to bondholders. Adding (5-7), (5-8), and (5-9), we obtain,

$$\tilde{y}^M + \tilde{y}^s + \tilde{y}^B = \tilde{X}, \tag{5-10}$$

[5]See [13] for this explanation.

that in equilibrium, yields the pure value of investment:

$$V(\tilde{X}) = X(\tilde{y}^{S}) + V(\tilde{y}^{B}) + V(\tilde{y}^{M}). \tag{5-11}$$

Equation (5-11) indicates that the present value of the manager's compensation, $V(\tilde{y}^{M})$, is a component of the market value of the firm. Furthermore, the market value of the firm is equal to the pure value of investment.

The incentive mechanism illustrated by Ross implies that financial managers will search for optimal capital structures to maximize their own wealth. This is a result of the assumption that a penalty is imposed on the manager if bankruptcy occurs and makes the manager's end-of-period compensation a function of his choice of capital structure. Only when managers find the firm's optimal capital structure, will capital markets be able to discriminate among different firms.

Although Ross deals with corporate debt policy, his incentive mechanism is an example of an effort to overcome the hazards of informational asymmetries. In Ross's framework, it is in the agents' interest to provide valid signals to "outsiders" in order to achieve firm value maximization. The optimal decision is to equate the marginal cost, $\Delta V(L\tilde{b})/\Delta\tilde{y}^{*}$, to the marginal benefit, $\Delta V_0/\Delta\tilde{y}^{*}$. As in numerous other partial equilibrium models, the optimal decision in choice of action is defined in a marginal benefit-cost context.

Chen and Kim [13] point out that Ross's model may break down because there is an economic incentive for shareholders to make side payments to managers to induce false signaling. These side payments can be easily disguised as part of the normal managerial compensation. There is also an incentive for both shareholders and bondholders to make side payments to managers to give false signals to the regulatory authority. These false signals may give rise to unwanted allowed rates of return or inflated rate bases.

A Signaling Model Developed by Leland and Pyle

Leland and Pyle [46] developed a simple model of capital structure and financial equilibrium in which entrepreneurs seek financing of projects whose true qualities are known only to them. This model assumes that entrepreneurs will signal capital markets by their retention of a fraction of the equity in the project. The implication of the model is that agents must give valid signals to outsiders if the market value of a project is to be maximized. Even if an agent is prohibited from having an ownership interest in the project, there are other conceivable ways to tie the manager's welfare to the success or failure of a project.

Leland and Pyle assume that capital markets are competitive and that there is no uncertainty about the projects. The total market value of the project, V, can be expressed as:

$$V = \frac{1}{1+\tau}\left[\mu(\alpha) - \lambda\right] , \qquad (5\text{-}12)$$

where

τ = the riskless rate of interest,

$\mu(\alpha)$ = the market valuation schedule, expressing the market's perception of the true expected return as a function of α,

α = the signal, representing the fraction of equity retained by the entrepreneur,

λ = the market's adjustment for the risk of the project with returns $\tilde{X}$ about the mean.

In the case of the capital asset pricing model, $\lambda = \lambda^{*}COV(\tilde{X},\tilde{M})$, where λ^{*} is the "market price of risk" and $\tilde{M}$ is the return on the market portfolio. Equation (5-12) indicates that the market value of the project is a function of the manager's retention of ownership in the project, α, and is equal to the certainty equivalent, $\mu(\alpha) - \lambda$, discounted at 1 plus the risk-free rate.

The agent is assumed to maximize his expected utility of wealth with respect to (1) the financial structure of the project or firm; (2) his holdings of equity in the project or firm; and (3) his holding of the market portfolio and the riskless asset. The agent's objective is to

$$\text{Maximize} \quad E\left[U(\tilde{W}_1)\right] \tag{5-13}$$

where $\tilde{W}_1$, = the agent's uncertain end-of-period wealth, $U(\cdot)$ is his preference function over W_1, and E is the expectations operator. Maximizing equation (5-13), given $\mu(\alpha)$, yields an optimal portfolio that depends on μ:

$$\alpha^* = \alpha^*(\mu), \tag{5-14}$$

$$\beta^* = \beta^*(\mu), \tag{5-15}$$

where α^* and β^* are the optimal holdings of the project and the market portfolio, respectively.

A market valuation schedule $\mu(\alpha)$ is said to be an equilibrium valuation schedule if the entrepreneur's true evaluation is correctly identified by the market for all values of μ for which the entrepreneur undertakes the project. That is,

$$\mu\left[\alpha^*(\alpha)\right] = \mu \tag{5-16}$$

for all levels of μ that induce the entrepreneur to undertake the project, given the schedule $\mu(\alpha)$.

Figure 5-3 provides an example of an equilibrium-signaling schedule. It shows that the market's perception of the expected return on the project is an increasing function of α. The agent will choose α in such a way as to maximize equation (5-13). Given competitive capital markets, if

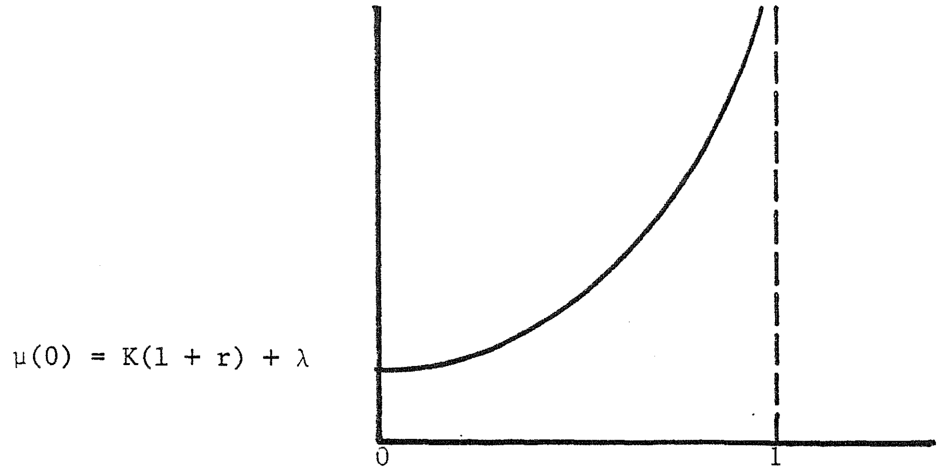


Figure 5-3 Hypothetical example of an equilibrium signaling schedule

Source: Adapted from [46, p. 378]

the market's perception of $\mu(\alpha)$ is greater than the actual μ (known only to the manager), outside investors would, on average, receive less than the return required for the project's risk, and equity financing would not continue on such terms. On the other hand, if $\mu(\alpha)$ consistently underestimated the true μ, given α, excess returns would exist for outside investors. Competitive forces will eliminate any excess returns; thus, for levels of μ for which entrepreneurs undertake their projects, it must be true that $\mu\alpha^*(\mu) = \mu$, in equilibrium.

The entrepreneur has a clear incentive to choose an optimal in order to assure that equity financing will continue and to maximize his own utility. An equilibrium occurs, when taking account of the market's evaluation of alternative equity positions, the entrepreneur signals correctly.

Ross has pointed out that Leland and Pyle have developed a model in which the market must know the risk preferences of the entrepreneur. It is in the general interest of managers to misstate these preferences, resulting in the need to institute additional signaling mechanisms.

Type D Mechanisms

Cost-of-Service Indexing

The first incentive mechanism reviewed in this section is the New Mexico cost-of-service-indexing method (COSI). See [43] for an extensive analysis of this indexing method.

The COSI method allows a rate increase to occur if the utility company earns less than a minimum allowed rate of return during a quarter or triggers a rate decrease if the rate of return on common equity rises above a maximum allowed rate of return. The same adjustment on a per kWh basis is also applied to the energy charge for each class of service.

Bazant [7] explains the objectives in designing and implementing rate indexing of this type. First, the COSI mechanism was designed to establish earnings stability and reliability by reducing investor risk and cost. In addition, it was anticipated that this method would attract necessary new capital. Second, it was hoped that the COSI method would encourage and enable the company to choose acts most beneficial to the public interest; that is, "...provide for demand growth, improved system reliability and comparatively lower energy costs to the consumer in the future."[6] Third, it was the intention of those developing this method to put in place a mechanism that would establish strong management incentives to control cost increases and enhance the efforts of management to create economies while maintaining the ability to plan and prepare for the future. Fourth, COSI attempts to provide for automatic and immediate benefits to consumers where net service cost savings are realized or for benefits to the company when cost increases are incurred. Fifth, a more realistic allocation of capital costs attributable to construction work in progress was sought between the current and future ratepayers. Sixth, and of considerable importance, this method was designed to reduce unnecessary demands upon the time, energies, and other resources of the company and the commission in traditional rate proceedings.

[6]Taken from [7, p. 44].

This method was conceived, designed, and recommended by Charles D. Olmsted, commission special counsel, and Stanley Bazant, Jr., as case director in a rate proceeding initiated by Public Service Company of New Mexico (PNM) in the fall of 1974. Since 1974, the COSI method has undergone extensive review, and certain adjustments to the original concept have been made during the actual implementation of COSI.

The mechanics of COSI are best explained by example. Service rate adjustment dates are set by the commission's order on the first day of February, May, August, and November of each year. On December 29, 1978, the commission ordered elimination of the quarterly adjustment and instituted an annual adjustment. At least 10 days before each adjustment date, the company is required to file a "cost-of-service index report form" with the commission.

The key figures in such statements are the average jurisdictional electric common equity investment and the jurisdictional net income available for common equity. Dividing the net income figure by the equity investment yields the annual return on jurisdictional common equity (ROE). The resulting ROE is then subtracted from the nearest edge of the allowed band, or range of return, for example, 13.5 percent to 14.5 percent. With this percentage difference between the current ROE and the appropriate range, the calculation of the current cost-of-service index is straightforward, as shown in lines 62 through 66 of table 5-2.

This model of service rate regulation is designed to maintain a level of revenue sufficient to cover the company's debt service, preferred stock dividends, operating costs, amortization, tax costs, and to provide a stable return on common equity.

Kaufman and Profozich [43] provide an analysis of the effect of COSI on service rates as compared to established ratemaking practices. They point out that in comparing the typical bill, PNM's bill is generally below those of other utility companies used in the analysis. The percentage change in PNM's typical electric bills over the period 1974-77 compares favorably with those of the other companies used in the analysis.

TABLE 5-2

ABSTRACT OF PUBLIC SERVICE COMPANY OF NEW MEXICO
COST OF SERVICE INDEX REPORT FORM

	Column B
Part IV - Return on Jurisdictional Common Equity	
60. Annual return on jurisdictional common equity (ROE) ((line 59 38) x 100%) x 2.0	15.743%
61. Percentage difference between current ROE and 13.5%-14.5% range	(1.243%)
62. Revenue differential ((line 38 x 61) ÷ 48.59% 2.0)	(953,982)
63. Jurisdictional kWh sales during period	1,381,356,115
64. Incremental index factor (line 62 ÷ 63)	(.000691/kWh)
65. Previous index factor	.002688/kWh
66. Current cost of service index (line 64 + 65)	$ .001997 kWh

Source: Taken from [7, p. 53]

It is difficult to determine the impact of COSI on the improvement of regulatory efficiency. The company indicated that COSI would free management from much of the burden imposed by the rate case cycle. On the other hand, the commission estimated that unification of the COSI data requires two to four times the effort required for a similar function in a traditional rate case. Furthermore, since the institution of COSI, the company has realized a substantial increase in its expenditures for legal services and outside consulting services; some of these cost increases are no doubt a result of COSI. Also, the commission does not appear to have realized substantial savings since the advent of COSI.

The question of increased efficiency of the regulatory process resulting from COSI must go unanswered at this time. There appears to be the potential for savings inherent in COSI but to date these savings have not been identified. This is partly due to the time and energy required of the company and commission to investigate and reflect on the several components that might signal increased efficiency; it appears at this time that necessary efforts have not been directed toward this end.

A Reward-Penalty Incentive Mechanism

Another incentive mechanism falling into category D was pioneered by Weitzman [86]. The information source for the mechanism is the agent, in as much as the agent chooses a target he feels can be achieved. The model is purposely presented here in regulatory framework that deviates from the approaches found in the literature. Initially, the regulatory authority will announce a set of targets, $T = (T_1, T_2, \ldots, T_N)$, and a bonus, or reward, based upon a number of criteria. The agent then formulates a set of targets, $\hat{T} = \hat{T}(\hat{T}_1, \ldots \hat{T}_N)$, and a bonus. The actual bonus, B', is calculated as

$$B' = \begin{cases} \hat{B} + \alpha(T - \hat{T}) & \text{if } T > \hat{T}, \\ \hat{B} + \beta(\hat{T} - T) & \text{if } T < \hat{T}. \end{cases} \qquad (5\text{-}17)$$

where

B' = the actual bonus calculated ex post,

$\hat{B}$ = the manager's announced bonus,

$\hat{T}$ = the manager's announced target,

T = the output achieved ex post,

α = a reward coefficient set by the authority,

β = a penalty coefficient set by the authority.[7]

Equation (5-17) indicates rewards to agents for performing better than their announced performance target and penalizes agents for falling short of their announced target. The authority has the ability to set α and β according to some criteria it feels will induce optimal agent behavior.

The targets may be output objectives, cost ceilings, productivity measures, or other measurable objectives. Allowing the agent to choose $\hat{T}$ and $\hat{B}$, ex ante, ideally induces the agent to define achievable and realistic objectives. The information source is the agent and other objective sources. Any incentive the agent has to bias T and B in order to reap benefits by easily surpassing his stated objective is mitigated by the asymmetry of α and β. A situation of $\beta < \alpha$ penalizes the agent less for falling short of his target than he is rewarded for surpassing his target. While this mechanism may reduce the agent's propensity to hide his true prior beliefs regarding achievable objectives, it is doubtful that this mechanism will totally solve the moral hazard problem.

To our knowledge, this mechanism has not been implemented explicitly by regulatory authorities in any sort of rate of return incentive mechanism. However, some reflection may leave the reader to recognize this mechanism as being similar to bonus compensation plans instituted by many corporations. In particular, if $\beta = 0$, a simple performance contract results, commonly used by many sales organizations. No doubt, a scheme analogous to the simple performance contract is used by utility companies to reward managers for superior decision making. Variations of equation (3-17) can provide the foundation for several types of incentive designs.

---

[7]The order of the magnitude of the coefficients is described by Weitzman as: $0 < \alpha < \beta$.

Summary

Table (5-3) restates each of the incentive mechanisms that have been reviewed in some detail and adds several others in the context of one of the categories defined in table 5-1. Mechanisms not mentioned in the text are also listed.

The majority of incentive mechanisms fall into categories A and D, especially those mechanisms designed specifically for regulated firms. Table 5-3 suggests that incentive mechanisms designed to obtain information from the agent have reward structures based on the agent's productivity. Mechanisms designed to obtain information from the agent as well as other sources have reward structures based on the agent's productivity, in addition to other performance measures.

Most of the incentive mechanisms identified in the table do not belong to their assigned category in all circumstances. Clearly, arguments can be made and justifications developed to move mechanisms from one category to another, depending upon how one perceives information is transmitted in the relevent socioeconomic environment. The framework used in this study is intended to serve as a guide in bringing together the notion of informational asymmetries and reward structures.

TABLE 5-3

INCENTIVE MECHANISMS IDENTIFIED IN THIS STUDY

Reward Structure	Source of Information: Agent Only	Source of Information: Agent and Others
Agent's Productivity only	1. The Incentive rate of return (IROR)	1. Cost-sharing contracts
	2. The managerial labor market	
	3. Insurance contracts	
	4. Fuel adjustment clauses with lags	
Total system productivity	1. Signaling model--Ross	1. The New Mexico cost-of service index
	2. Signaling model--Leland and Pyle	2. Reward-penalty mechanism
		3. Plant certification
		4. Contract renegotiation
		5. Regulatory lag provisions
		6. Michigan three-part incentive mechanism
		7. Total factor productivity

CHAPTER 6

THE PROBLEM OF PREDICTING THE FEASIBILITY OF NEW INCENTIVE MECHANISMS

For cost control efforts to be successful, it is not enough to design new incentives or to rework old ones. It is imperative to find out whether the incentives are likely to be used effectively, or at all, by the commissions.

In earlier chapters of this report, an attempt was made to analyze the problem of providing incentives and to lay down a foundation of knowledge upon which the development of new incentives may proceed. The successful introduction of new cost control mechanisms requires (1) research and development, (2) testing in the field, (3) communicating the existence and advantages of the mechanisms to the commissions, (4) adoption and experimentation by the commissions in the use of the mechanisms, and (5) incorporation of the mechanisms into the routines of commission behavior. Progression through these stages is far from automatic. Major problems that have shown up repeatedly in attempts to move from research and development to actual use of a new policy or program are as follows:

1. competing interests of potential users and the promoters of the innovation
2. needs for redesign of the innovation to suit the particular situation of a state or local organization
3. resistance to the innovation within the organization and
4. inadequate guidance on monitoring of implementation [91]

Many researchers have attempted recently to learn more about how and why a public organization takes on a new policy or program. Thus, even though our research is not yet at the stage where the practicality of alternate mechanisms can be determined directly, it is possible to provide a frame of reference for assessing the conditions under which successful implementation of incentives might be expected to take place. In this chapter, state public utility commissions are viewed as adaptive systems constantly facing both routine and novel problems.

Since the questions for development of incentive mechanisms are how and why commissions deal with the novel and nonroutine, this chapter provides a description of several traditional approaches for studying organizational innovation. Chapter 7 presents a review and an analysis of findings of existing studies of adoption and implementation of innovation as they may apply to the commissions.

The Routines of Commission Regulation

Like most organizations, a great deal of what commissions do is highly programmed and repetitive. Figure 6-1, A general model of a public organization, provides a framework for an introductory description of commissions as they exist today.

Commission Environment

As with all public organizations, the commissions are intended to be accountable to the public through elected officials or officials appointed by elected ones. "Suprasystem control organizations" are the vehicles for asserting public control.

Unlike other public organizations, most of the commissions are "independent," meaning that only the courts can review their decisions, not the governor or the legislature. "Independence" is relative, however. The governors, often with legislative approval, generally appoint public utility commissioners; and the state legislatures review their budgets.

Contextual Environment

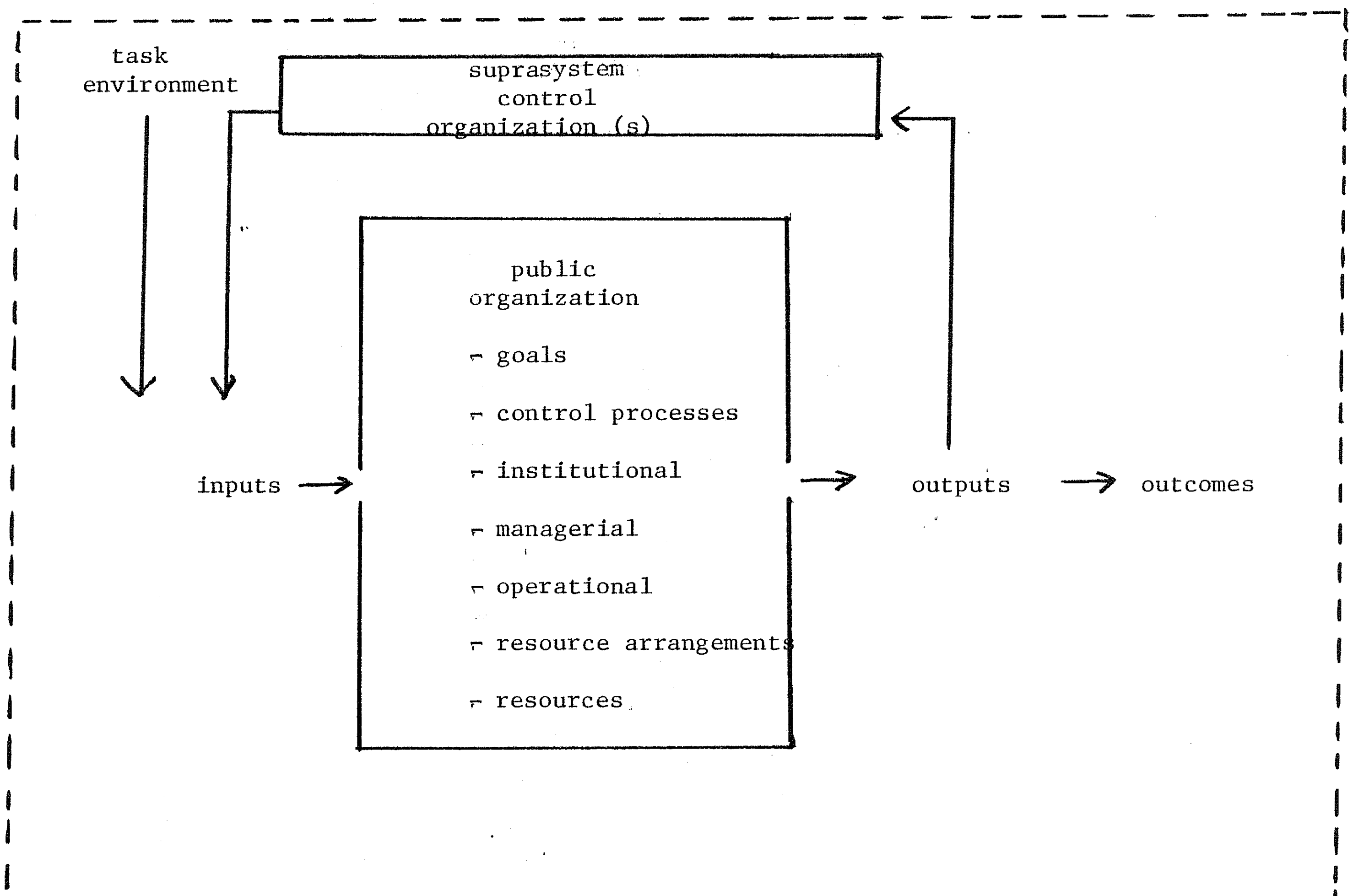


Figure 6-1 General model of a public organization

Thus, the control organizations depicted in figure 6-1 include the governor's office and the legislature as well as the judiciary.

The environment of an organization may be divided into "contextual" and "task" components. The "contextual environment" is composed of elements in the environment that are not directly related to the organization being observed. The "task environment" is composed of elements of the environment directly related to the organization [3].

The control organizations of the governors' offices, the state legislatures, and the state and federal courts are elements of the commission's task environments. So are other organizations, including the U.S. Department of Energy and other federal agencies, the state energy offices, private suppliers of goods and services to the commissions, professional associations like the National Association of Regulatory Utility Commissions (NARUC), research centers like The National Regulatory Research Institute (NRRI), and public interest groups. Social, economic, and political trends and events are also part of the commissions' task environments. The world price of oil is an example.

Components of the Commissions

Any organization is a purposeful system that uses the resources at its command to produce goods or services. Goals or policies set the general direction of an organization's activities. The overriding formal goal of state public utility commissions is to limit the market power of businesses that through economic fact and legal choice have come to be designated as public utilities. The power, telecommunications, water, and transport industries have formed the core of the jurisdictions of the 50 state commissions, the District of Columbia's commission, and the 8 federal commissions. These industries have, with recent exceptions, been determined through the public policy-making process to have undue financial power in dealing with their consumers and to be essential to the public well-being. It followed that they should not be allowed to function without detailed supervision of their most fundamental financial decisions. Yet, reflecting American mistrust of both big business and big government, outright ownership of the utilities has for the most part been eschewed.

Instead, the commissions were established and charged with three fundamental tasks: (1) to set fair ceilings on utility prices, (2) to set fair ceilings on utility profits, and (3) to ensure adequate service to consumers.

The resources that make up an organization include people, machines, materials, and information. To illustrate, the total budgets of the commissions (in 1976) ranged from a low of $241,000 for one particular commission to a high of $17 million for another. Total staff size ranged from 7 in to 855. (See table 6-1, Characteristics of State Regulatory Agencies.) Staff specialists include lawyers, accountants, statisticians, engineers, and rate/tariff analysts. Their training may be considered a commission resource that enables interpretation and analysis of information from utility reports, audits, citizen complaints, and hearings.

Resource arrangements are "the way in which resources or elements of an organization are combined or mutually arranged" [3]. The concept is essentially synonymous with "structure" but reminds us that structure refers to the arrangement of resources at a particular time. The resource arrangements of commissions vary considerably, so that it is not possible to provide an organization chart of a typical commission. The attached diagram of the Michigan Public Service Commission shows one arrangement. (See figure 6-2.) The Michigan commission is organized both by organizational function, such as data collection or tariff analysis, and areas of regulation, such as railroads, gas production, and transmission.

The hierarchical levels of the commissions are commonly distinguished as "commissioners" and "staff." Thompson [77] and Parsons [61] suggest that organizations exhibit institutional, managerial, and technical, or operational levels of responsibility and control. At the institutional level, the relationship of the organization to its environment is articulated. At the managerial level, the technical suborganization is serviced through (1) mediation between the technical suborganization and users of its products, (2) procurement of resources; and (3) administration, or control of the technical suborganization (Thompson [77, pp. 10-11]).

TABLE 6-1

CHARACTERISTICS OF STATE REGULATORY AGENCIES

STATE	MEMBERS	TERM	SELECTION	SELECTION CHRM.	GAS & ELECTRIC STAFF	TOTAL STAFF	BUDGET
Alabama	3	4	Popular Vote	Popular Vote	55	57	610,000
Alaska	3	6	GS	G	29	29	641,000
Arizona	3	6	Popular Vote	Comm. Elect	25	148	672,000
Arkansas	3	6	GS	G	52	47	438,000
California	5	6	GS	Comm. Elect	154	855	16,878,000
Colorado	3	6	GS	G	17	85	1,329,000
Connecticut	5	6	G, A&C	Comm. Elect	64	82	1,559,500
Delaware[1]	5	6	GS	G	9	7	682,800
Florida	3	4	Popular Vote	Comm. Elect	114	346	7,100,800
Georgia	5	6	Popular Vote	Comm. Elect	40	104	2,312,866
Hawaii	3	6	G	G	14	17	481,000
Idaho	3	6	GS	Comm. Elect	28	67	337,000
Illinois	5	5	GS	G	132	226	6,040,900
Indiana	3	4	G	G	59	80	1,621,000
Iowa	3	6	GS	Comm. Elect	55	108	1,033,000
Kansas	3	4	GS	Comm. Elect	29	144	2,765,407
Kentucky	5	4	G, A&C	G	60	52	1,344,000
Louisiana	5	6	Popular Vote	Popular Vote	43	93	1,574,000
Maine	3	7	GL	GS	31	55	296,000
Maryland[2]	5		G, A&C	G	65	67	2,413,000
Massachusetts	3	4	G	G	51	112	1,700,000
Michigan	3	6	G Senate-Reject	GS	112	265	8,806,000
Minnesota	5	5	GS	Comm. Elect	33	134	2,959,000
Mississippi	3	4	Popular Vote	Comm. Elect	23	73	
Missouri	3	6	GS	G	84	189	2,298,000
Montana	5	4	Popular Vote	Comm. Elect	13	27	
Nebraska	5	6	Popular Vote	Comm. Elect	19	61	1,200,000
Nevada	3	4	G	G	30	66	1,000,000
New Hampshire	3	6	G, A&C	G, A&C	18	34	262,000
New Jersey	3	6	GS	GS	89	221	3,764,000
New Mexico	3	6	GS	G	20	17	571,000
New York	7	6	GS	G	334	627	17,262,000
North Carolina	7	8	GL	G	102	111	1,008,000
North Dakota	3	6	Popular Vote	Comm. Elect	48	33	775,000
Ohio	3	6	GS	G	103	295	3,446,000
Oklahoma	3	6	Popular Vote	Comm. Elect	14	209	2,825,000
Oregon	1	4	G	G	55	296	11,450,000
Pennsylvania	5	10	GS	G	263	476	10,240,000
Rhode Island	3	6	G, A&CS	G	18	20	241,000
South Carolina	7	4	Elect Gen. Assy	Comm. Elect	36	96	2,241,000
South Dakota	3	6	Popular Vote	Comm. Elect	15	23	482,000
Tennessee	3	6	Popular Vote	Comm. Elect	20	123	1,742,000
Texas	3	6	G, A&CS	Comm. Elect	59	47	1,200,000
Utah	3	6	GS	Comm. Elect	6	22	708,585
Vermont[3]	3	6	G, A&C	G	25	27	398,000
Virginia	3	6	Legis. Elect	G	15	480	10,600,000
Washington	3	6	GS	G	36	169	4,300,000
West Virginia	3	6	GS	G	59	121	3,090,000
Wisconsin	3	6	GS	G	90	143	2,286,000
Wyoming	3	6	GS	Comm. Elect	17	29	450,526
Dist. of Columbia	3	3	Mayor/City Council	Comm. Elect	19	28	989,400

[1]All five Delaware Commissioners are part-time

[2]Three Maryland Commissioners are full-time and two are part-time.

[3]Vermont has one full-time commissioner, the chairman, and two commissioners are part-time.

G - Appointed by governor
GS - Appointed by governor, approved by senate
GL - Appointed by governor, approved by legislature
G, A&C - Appointed by governor, advice and consent of legislature
G, A&CS - Appointed by governor, advice and consent of senate

SOURCES

The States and Electric Utility Regulation, (Alan E. Finder, ed.) The Council of State Governments. Lexington, Kentucky, 1977.

Lehman's Regulatory Analysis Service, 1978, Public Utilities Group Lehman's Securities, 120 Broadway, New York, N.Y.

Annual Report on Utility and Carrier Regulation, (Geneva Beierlein, ed.), National Association of Regulatory Utility Commissioners, Washington, D.C., 1976.

Source: [55, pii]

In Parsons's and Thompson's terminology, commissioners are located at the institutional level of their organizations. However, to the extent that they personally preside over the ratemaking proceedings of the commissions, they are part of the technical core, or operational level, as well. Moreover, the institutional level of the commissions is not exclusively the domain of the commissioners. Because their terms of office and their initial expertise with public utility regulation are often limited, key top level staff may also be considered members of the institutional level of commission control.

An unusual feature of commission organization at the institutional level of control is its collegial form. Instead of being headed by one chief executive, the commissions are typically composed of three, five, or seven commissioners having equal votes. (See table 6-1.) Most commissioners are appointed. The commissioners in 12 states are elected by popular vote. In 2 states they are elected by the state legislature. Commissioners' terms last from 4 to 10 years, with 6 the mode.

Organizational control processes are processes through which "the goals of the organization are generated, the horizontal and vertical differentiation devised and directed, the work done in the organization is accomplished at the operational, managerial and institutional levels of problem solving, performance is measured and evaluated, adaptation to internal and external disturbances is directed and controlled, and operations or activities are executed which produce the products of the organization" (Backoff, [3, p. 29]).

A fundamental process of a public utility commission, one that goes on at the commission's technical core, is ratemaking. The ratemaking process includes three stages. First is a request for a rate change. Usually the request is from a regulated utility, and usually the change involves a rate increase. However, rate changes may be initiated by commission staff as well. Out of the request, a rate case develops. A rate case is a formal hearing to consider the arguments for and against the proposed change. The outcome of a rate case is a formal order stating what new rates will be charged to whom.

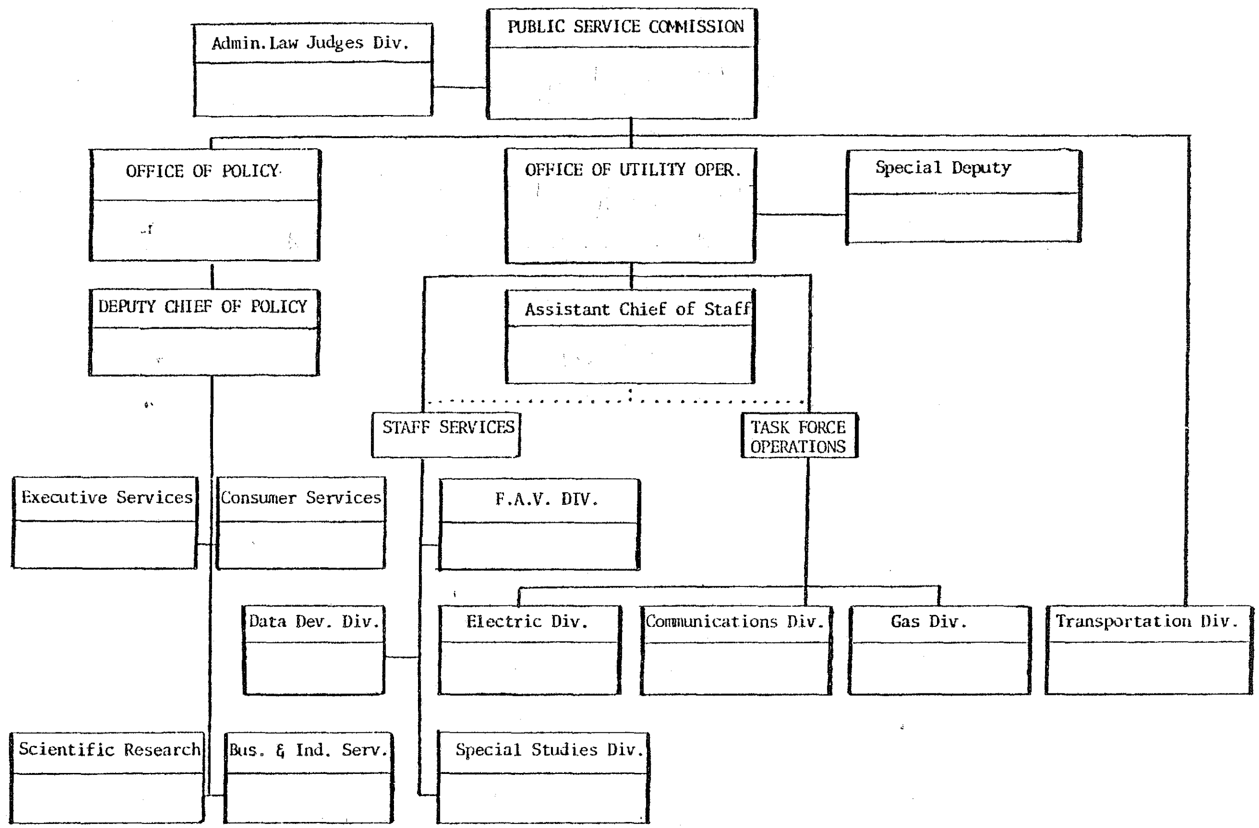


Figure 6-2: Organization Chart of the Michigan Public Service Commission

Source: [54]

Besides hearings to consider specific rate requests, commissions may conduct "generic hearings" to review arguments on an issue that does not require an immediate rate decision. They also process complaints, monitor the performance of the utilities through regularly required statistical reports, and conduct other formal business as well as the business of the informal organization.

Outputs are states of affairs produced by organizational processes interacting with resources and resource arrangements. Outputs have an impact on the environment. The organization also generates products that have only an internal impact.

The major outputs of a commission are intended to be rates that will produce revenues to assure the utilities a fair return on their investment and allocate costs fairly among different classes of users.

The final components of the model of an organization presented in figure 6-1 are the flows that link the organization's outputs to the environment and then feed back to the organization as new inputs. These feedback relationships "provide us with a means of conceptualizing the dynamic, working exchange between the organization and its environment." At any particular time, the structure and routines of the commissions may be fixed. Over time, the commissions are faced with nonroutine developments from within and without.

The Nonroutine in Commission Regulation

So far, state public utility commissions were introduced in the context of an open systems model of an organization. Because organizations are at least partially open systems, they are subject to internal and external disturbances that may result in change in any of their components or processes. A brief description of a simple model of organizational change, followed by a discussion of the concept of innovation, may be used as the framework for a review of changes in commission regulation and criticism of commission change or the lack of it.

A Model of System Adaptation

Adaptation is "the ability of a system to react to environmental (or internal) disturbances in a way that is favorable to the continued operation of the system" (Bobbitt et al. [9, p. 241]). Figure 6-3 depicts the process of adaptation to environmental disturbances. The model presents adaptation as a three-step sequence beginning with a stimulus from the environment, followed by the organization's response, and ending with the environment's response. The sequence then begins again as the environment provides a new stimulus to the organization. The model is shown as modified by Bobbitt et al. [9] after Feibleman and Friend [26].

The original stimulus from the environment may be of three types, distinguished according to the magnitude of their effect on the organization. A "negligible" stimulus is not strong enough to evoke a response. A "destructive" stimulus does not evoke a response either; it destroys the organization. An "effective" stimulus may be "minimal," or just sufficient to evoke a response from the organization; "optimum"; or "drastic," a large response that stops short of being destructive. The effective stimulus may produce a "maladaptive" response, an "elastic adaptive" response, or a "plastic adaptive" response. An elastic adaptive response allows the organization to return to its original state. If the organization makes a plastic adaptive response, it is no longer able to return to its original state.

The organization's response may, in its turn, have a "limited," "adjustive," or "extensive" effect on the environment. An adjustive effect results in at least a temporary steady state between the organization and those aspects of its environment affected by the response. An extensive effect reverberates through the environment, producing a new stimulus to the organization, and starting the cycle over again.

The model of organizational adaptation that has been presented may be helpful in looking more deeply at the responses of commissions to their environment. The model takes note of the possibility that potential environmental stimuli may provoke no response at all from a commission.

The model suggests that organizational choices over time will have an effect on the ability of a commission to continue to do its job, depending on whether responses are maladaptive or adaptive. And the model distinguishes two possible forms of adaptation, elastic and plastic, that are a vital dimension of the behavior of organizations. An organizational system may be characterized as particularly elastic or particularly plastic in its habitual responses to environmental stimuli. A completely elastic system could handle any change; a completely plastic system could adapt to nothing.

The Study of Innovation

Organizational innovations are new stimuli from the environment that have the potential of changing one or more aspects of the organization's components. The study of organizational innovation is concerned with describing and explaining adaptation to these new environmental stimuli. Before looking at innovation in the commissions, let us review some general conceptual issues in the study of innovation and some approaches that have been used to study the innovation process.

Innovations may be distinguished from inventions. Innovations originate outside the organization. Inventions are brought into being within the organization. The distinction is desirable, since the responses of organizations to innovations and their propensity to produce inventions appear to have different determinants.

Innovations are "new": "Only those attributes constituting 'newness' are necessary and sufficient conditions for an idea, practice or thing to qualify as an innovation" (Zaltman et al. [93, p. 46]). However, whether the innovation must be new to the environment or merely to the potential user is a subject of scholarly debate. Definitions of innovation have relied on both aspects of "newness." Barnett [5, p. 7], quoted in Zaltman, defines an innovation as "any thought, behavior or thing that is new because it is qualitatively different from existing forms." Similarly, the Federal Trade Commission has stated "that consumer and industrial products can be called 'new' only when they are either entirely new or have been

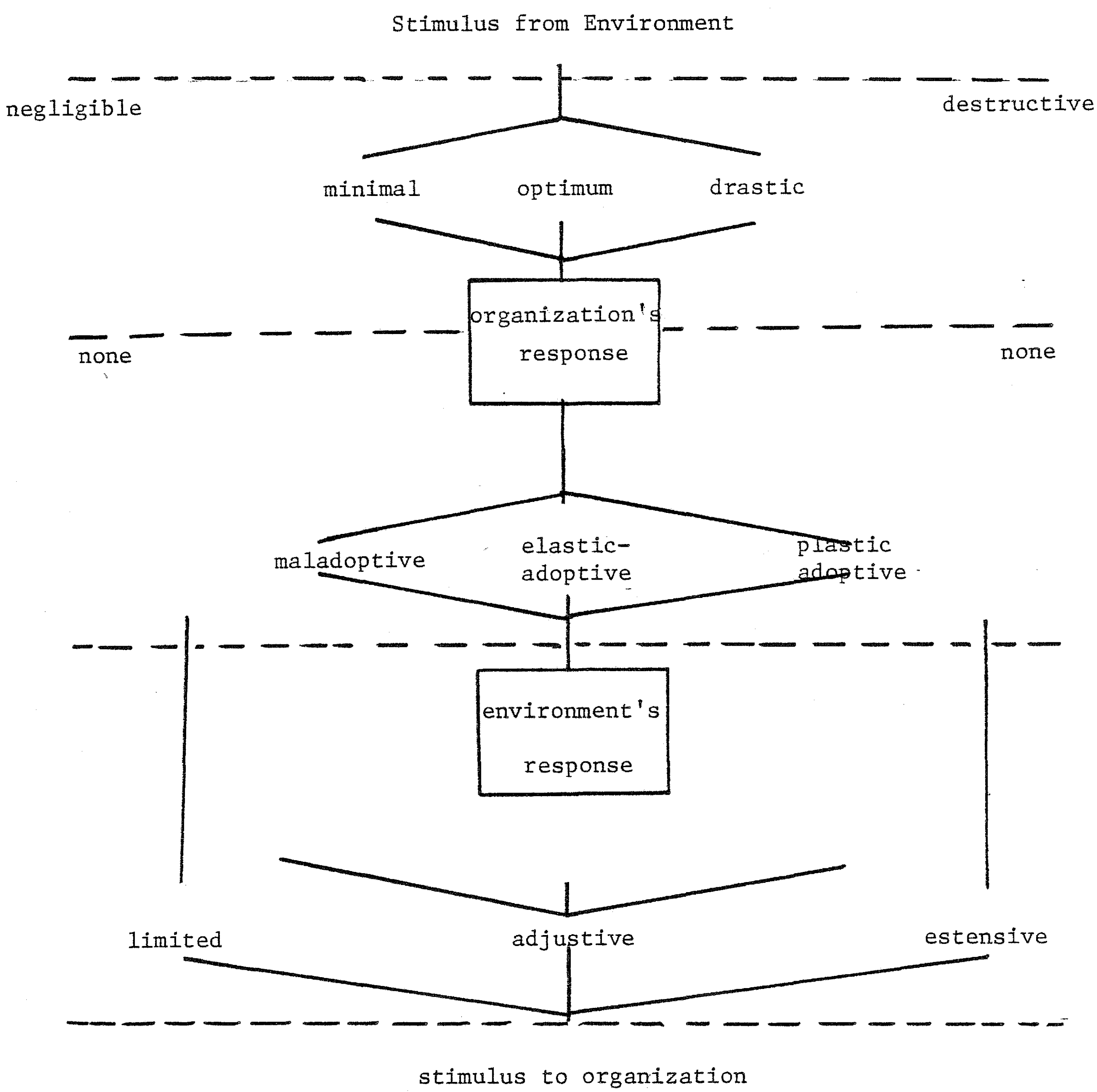


Figure 6-3: Organizational Adaptation

Source: [9, p.242]

changed in a functionally significant and substantial respect" (Zaltman [93, p. 9]). Others have said something may be called an innovation only while it has existed in the environment for a short period of time or been tried by a small percentage of prospective users.

The alternative use of newness as depending on perceptions rather than on objective qualities has been expressed by Rogers and Shoemaker:

> An innovation is an idea, practice, or object perceived as new by the individual. It matters little, as far as human behavior is concerned, whether or not an idea is "objectively" new as measured by the lapse of time since its first use or discovery......If the idea seems new and different to an individual, it is an innovation. [70, p. 19]

The Rogers and Shoemaker definition refers to individual perceptions. Their approach to "newness" has also been used where organizations are the unit of analysis (Walker [84], and Hage and Aiken [34], for example).

Innovations, if used, will change some aspect of the organization. How much change an environmental stimulus must be likely to induce to be called an "innovation" is another conceptual issue in the study of innovation. Robertson [68, p. 7], writing from a marketing perspective, distinguished among innovations by "how continuous or discontinuous their effects are on established consumer patterns." "Continuous" innovations generally involve alteration of existing products, rather than creation of new ones, says Robertson, and have the least effect on established consumption patterns. "Dynamically continuous" innovations have more disruptive effects than continuous innovations. "Discontinuous" innovations involve "the establishment of new consumption patterns and the creation of previously unknown products" (ibid.). Wilson [90, p. 196] says "an innovation (or, more precisely, a major innovation, since we are not concerned with trivial changes) is a 'fundamental' change in a 'significant' number of tasks." March and Simon [49] take a similar tack, saying "initiation and innovation are present when change requires the devising and evaluation of new performance programs that have not previously been a part of the organization's repertory and cannot be introduced by a simple application of programmed switching rules" [49, p. 175]. Downs [22] looks at the degree of change in a public organization by the structural depth to which the change reaches. From shallowest to deepest, the layers suggested by Downs are (1) specific actions taken by the bureau, (2) the decision-making rules it uses, (3) the institutional structure it uses to make those rules, and (4) its general purposes [22, pp. 167-68].

Although "innovation" in ordinary language connotes change that is more than trivial, amount of change, like the complementary concept of newness, depends on the perceptions of the individual adopter. For the researcher in organizational innovation, neither newness nor amount of change can be assumed to inhere to the particular stimulus being studied but should be allowed to vary. An organization that does not perceive a particular innovation as very new or very major may respond differently from one with a different opinion.

Clearly, it is not possible to develop a monothetic typology of innovations from the central characteristic of newness or from the degree of change that might be provoked by the innovation. Indeed, we agree with Daft and Becker [19, pp. 120-22] that most typologies that have been suggested for the classification of innovations use attributes or correlates that may vary, rather than a relatively invariant partitioning.

A usable monothetic classification of innovations may, however, be derived from the initial focus of the innovation. Zaltman et al. [93, p. 31] lists several such typologies. Dalton [20] proposed that innovations be classified as technological, value centered, or structural (administrative). Knight [45] classified innovations as product or service, production process, organizational structure, or people. His classification appears to include technological resources, like computers, as part of production processes, although other researchers have found it useful to distinguish between things and their uses (Eveland et al. [24], and Bingham [8]). Grossman [33] distinguishes between "ultimate" and "instrumental" innovations.

Daft and Becker's [19] typology is (1) technical, divided into product and process, (2) "organizational structure" (administrative); and (3) goal-(value-) centered innovations [19, p. 123]. One problem with this typology is that technical processes do not occur without "people processes." A second is that administrative processes are considered a subtype of administrative structure. We suggested earlier that it may be desirable to distinguish between processes occurring over time and time-specific resource arrangements. Finally, the Daft and Becker typology subsumes "people" under organizational structure when, again, it may be desirable to consider them as resource components that may be grouped and linked in a variety of ways.

The Backoff model of public organizations (figure 6-1) provides a way of typologizing innovations by initial focus that we believe is preferable to the Daft and Becker typologies as well as to the others reviewed above. Backoff's classification scheme appears to exhaust the components and activities of an organization and to allow assignment of innovations to one and only one classification: (1) goals, (2) resources, (3) resource arrangements, (4) control processes, and (5) products or outputs.

It is possible to classify innovations by their source as well as by their initial focus. Researchers have looked at innovations originating in particular classes of sources but to our knowledge have not proposed a formal classification system based on the origin of the innovation. For organizations, such a classification might include (1) suprasystem control organizations, (2) peers, (3) suppliers, (4) clients, or (5) individuals and organizations outside the organization being studied but not included in the first four categories.

An innovation may be a thing, a practice, or an idea. Innovation is also a process and has been studied as such through three "traditional approaches." The material below is largely adapted from Yin, [91] and from Havelock [36].

The Research, Development, and Diffusion Approach

The first approach to the innovation process sees it as (1) discovery and laboratory testing, (2) field testing and demonstration, (3) diffusion to potential users, (4) user testing, and (5) adoption or rejection of the innovation by the users.

"Diffusion" is the spread of something. The research, development, and diffusion approach considers the diffusion process as a social phenomenon that can be summarized by temporal or spatial patterns of adoption or nonadoption in a population of potential adopters of an innovation (Robertson [68, p. 32] and Brown [12, p. 6]). The cumulative pattern of diffusion over time that has received the most theoretical and empirical support is an S-shaped or logistic curve. Robertson attributes the idea of the S-shaped diffusion function to the sociologist Gabriel Tarde. Evidence for such a generalized pattern has since been found in a number of contexts, such as the diffusion of Friendly ice cream stores, Planned Parenthood affiliates [12], a new drug among physicians, and new agricultural technologies among farmers [68],[69].

In marketing, the concept of diffusion is closely related to that of the product life cycle. The difference is that "diffusion refers to the percentage of potential adopters within a social system or market segment who adopt over time, whereas the product life cycle is based on absolute sales levels over time" (Robertson [68, p. 30]).

While the research, development, and diffusion approach has been widely used to attempt to explain innovative processes, it has not been successful in saying why innovations are or are not used by public organizations:

> [At] least in the case of local organizations, the approach has not satisfactorily explained why some innovative efforts succeed and others fail (House [40]). The main problem is that experience has shown little relationship between the successful progression of an innovation through the first three stages and the ultimate extent of adoption. One

> reason for this failure to explain innovative efforts in local organizations is that the approach assumes, first, the transferability of innovative experiences from one site to another and, second, a passive role by the ultimate adopter or user. Because of these two assumptions, the approach grossly overlooks implementation factors, and these very factors may be the reason that otherwise demonstrably worthy innovations are not adopted. (Yin, [91, p. 8])

Yin's assessment is limited to studies of local organizations. However, the arguments appear to apply to state organizations as well. Further, as noted by Yin, it would not be enough to insert an implementation stage into the original approach. Initiatives by potential adopters may influence or even precede research and development. In fact, "the basic features of an organization and its local setting, as well as the manner in which an innovation is introduced, may even outweigh the importance of research, development, or diffusion" (Yin [91, p. 9]).

The Social Interaction Approach

"Social interaction" is the term used by Havelock [36] and Yin [91] to describe the classic "adoption perspective" on diffusion.

The adoption perspective is "the dominant and most completely developed area of research on the diffusion of innovations, representing an effort spanning about 45 years" [11, p. 7]. This is the area of innovation research typified by the work of Everett Rogers. In the Rogers model, the individual passes through a series of stages, beginning with knowledge of the innovation. At the "persuasion stage" he forms a favorable or unfavorable attitude toward the innovation. At the "decision stage" he "engages in activities which lead to a choice to adopt or reject the innovation." At the "confirmation stage" he "seeks reinforcement for the innovation decision he has made. But he may reverse his previous decision if exposed to conflicting messages about the innovation" (Rogers and Shoemaker [70, p. 103]).

The social interaction approach or adoption perspective is compatible with the research, development, and diffusion approach. The difference is in point of view. The diffusion approach looks at diffusion as a social phenomenon; the adoption approach looks at diffusion from the standpoint of the individual.

Because of its focus on individual perceptions, the adoption perspective emphasizes the communications system in which an individual adoption takes place. Indeed, Rogers and Shoemaker view innovation as the subset of communications research that deals with new ideas [70, p. 12].

The dependent variable in social interaction studies is often individual "innovativeness." The independent variables are then characteristics hypothesized to be correlated with repeated early adoption of innovations.

Yin criticizes attempts to use the social interaction approach to describe and explain the innovation process in organizations:

> The approach tends to obscure the problem of the shifting nature of the innovation from site to site, the problem that innovative efforts are not the result of a unitary act of adoption by a single adopter, and the possibility that implementation factors may once again play an important role much earlier in the staged sequence and not merely after the decision to adopt has been made. An alternative conclusion would be to challenge the overall utility of the social interaction approach rather than to modify it, and to look for other ways of studying innovation in local organizations. (Yin [91, p. 10])

The Problem-Solver Approach

The problem-solver approach avoids the conceptual inadequacies of the first two approaches for the study of organizational innovation. However, it presents several difficulties in operationalization. The approach looks at the innovative process from the standpoint of a decisionmaker or decisionmakers in the organization. The innovation process is seen as beginning with a "performance gap," a perceived discrepancy between what the organization is doing and what it ought to be doing (Downs [22, p. 191]). The perception of a gap stimulates search for a course of action that will close the gap. The performance gap may arise from (1) slow adjustment of criteria of satisfaction to actual achievement (March and Simon [49, p. 183]); (2) a natural tendency of criteria of satisfaction to adjust upwards over time (ibid.); (3) changes in the organization's internal environment, such as new personnel; and (4) changes in the organization's external environment, such as changes in demand for the organization's output (Zaltman et al. [93, p. 169]; HUD [82, p. 14]; and Menzel [50]).

Recognition of a need for change is considered the first stage in a sequence followed by (1) identification of alternative solutions, (2) selection of an alternative, (3) implementation of the alternative, and (4) evaluation of the results (Yin [91, p. 11]).

The problem-solver approach fits nicely into a general model or organizational adaptation. In emphasizing the choice process within organizations, it highlights the potential both for maladaptive adoption of innovations and for changes in the innovation after the initial adoption.

The operational shortcomings of the problem-solver approach, as identified by Yin are (1) difficulties in specifying the organizational events that are to be the units of analysis, and, (2) difficulties in specifying end states to be used as the dependent variable [91, p. 12]. He suggests that it may be for these reasons that the problem-solver approach has not been greatly used for empirical research.

The Implementation Approach

Yin distinguishes from the problem-solver approach a newer "implementation" approach that includes his own work. However, a focus on implementation is closely related to the third tradition of innovation research. The major difference appears to be in the emphasis on solving the problems of generating internal support for change, and rearranging resources and acquiring new ones as necessary to use the innovation, as well as the problem of deciding whether to use it. Recent work by Eveland and Rogers [24], exploring the "specification" of an innovation after it has been adopted, may also be considered research using the implementation approach. In this, we disagree with Yin who says Rogers and Eveland have merely added an "implementation stage to the social interaction model." He finds their model no less susceptible to criticism than other "elaborations" of the social interaction approach. Like the problem-solver approach, the implementation approach presents difficulties in operationalization.

The Innovation-Decision Design Approach

Noting "the extreme variance" in findings on organizational innovation, Downs and Mohr [23, p. 700] have proposed using what they call an "innovation-decision design." The 1976 article did not present a new theory of innovation but suggested ways to go about developing a better one. In their emphasis on adoption rather than on implementation, Downs and Mohr carry on the tradition of the social interaction approach to the study of innovation. In their emphasis on different organizational responses to different innovations, their work is in the tradition of the problem-solver approach.

Downs and Mohr begin their argument with the suggestion that we "reject the notion that a unitary theory of innovation exists and postulate the existence of distinct types of innovations whose adoption can best be explained by a number of correspondingly distinct theories" [23, p. 701].

They say that to begin to untangle types of innovations and delineate the appropriate theories, we need to pay attention to the "secondary attributes" of innovations. Primary attributes, say Downs and Mohr, are those which in a particular piece of research show no variation from one organization to another. All the organizations in a study might perceive an innovation as low cost, for example, making cost a primary attribute. Secondary attributes are perceived differently from one organization to another. "Secondary attributes can be viewed as variables that characterize the circumstances surrounding a particular decision to innovate" [23, p. 706].

Downs and Mohr are particularly critical of studies that combine adoption of a number of innovations as the dependent variable. They say that since the impacts of variables are different across organizations, the multiple-innovation studies are creating an "aggregation bias." In addition, they say that the studies do not aid in studying "innovativeness" that may vary by type of innovation [23, p. 707].

The innovation-decision design "eliminates any confusion that might stem from volatile secondary attributes" [23, p. 706]. The relationships of many innovations and many organizations would be the units of analysis. For example, 10 innovations and 100 organizations might be studied, giving a sample size of 1,000.

Studies that follow Downs's and Mohr's suggestion that many innovations be studied for their varying impacts, rather than single innovations or aggregations of many innovations, may well clarify differences in adoption behavior. The innovation-decision design, however, demands a great deal of data from the responding organizations if it is to be usable.

From all the above, it can be seen that a theory (or theories) adequate to understanding and aiding commission innovation is still lacking. One noted public utilities economist said the following:

> The theory of the firm under regulatory constraints provides some insight into corporate behavior and possible inducements to change, but a parallel theory of regulatory behavior is far from complete and can only be pieced together from an examination of regulatory techniques and practices--clearly, both theories are needed to permit a rigorous analysis of regulatory accommodation to change. (Trebing [79, p. 42])

Trebing's assessment of the state of development of a theory of innovation is, like Downs's and Mohr's, a fairly gloomy one. However, considerable empirical research has been done on public sector innovation that can provide insights into regulatory accommodation to change. In the next chapter we will review some of that research, and we will try to assess the implications of the findings for development of new incentives to be used by the commissions to promote control of costs by the electric utilities.

CHAPTER 7

RECENT STUDIES OF INNOVATION IN PUBLIC ORGANIZATIONS

Introduction

A review of existing studies may be useful to cost control efforts by focusing attention on those variables that contribute to the processes of organizational innovation and are susceptible to manipulation by policy makers. The overall findings do not provide much encouragement. It seems safe to say that the largest commissions are more likely to try an innovation. Otherwise, the only generalization that is universal to all the studies is Gray's conclusion [31] and the confirming evidence of Bingham [8] and Feller [27] that innovativeness is specific to time and issue. It may be that regional leader-follower relationships exist among the commissions. It would be of interest to the cost control effort to know if such patterns exist. If they did, this might aid selection of states for field testing of proposed innovations. However, as Flaherty [29] points out, we do not have detailed information on the communications network of the commissions.

General socioeconomic characteristics, organizational characteristics, and characteristics of individuals within the organizations can only be manipulated indirectly, if at all. One might propose increasing the commissions' budgets or expanding training programs for commission staff as a way of fostering commission adoption and incorporation of innovations, but such actions would be expensive and the outcomes highly uncertain. It is in the areas of characteristics of the relationship of the commissions to proposed innovations and factors related to the communications process that we find variables that are both influential and potentially manipulable.

To be interested in adopting an innovation, a commission must perceive it as necessary, inexpensive, and free of risk. These are common sense conclusions, but they do suggest some variables to consider in developing and implementing cost control incentives, particularly as they are related to risk:

- Modifiability: Innovations are more likely to be adopted if they are designed to be modifiable, divisible, reversible, and thus, capable of being implemented incrementally
- Degree of change: Innovations are more likely to be adopted and incorporated if they are perceived as clearly related to existing tasks and historical mission
- Visibility: Innovations are more likely to be adopted and incorporated if they are highly visible and their outputs can be measured directly
- Task specificity: Innovations that have one clear application or minimal competition between service and administrative applications are more likely to become part of the routines of commission behavior

Assuming that an incentive for cost control had been designed and that its developers felt it could improve significantly commission regulation of the electric utilities, and the developers fully believed that any rational commission would see the advantages of the innovation and adopt it, the news must still be spread to the commissions. The means must be found to aid the commissions in becoming aware of the innovation and knowledgeable about it, in experimenting with it themselves and in developing the internal expertise to use it repeatedly as an increasingly routine regulatory tool. Articles in professional journals, seminars, individual and group conferences, field trials, and expert guidance during the postadoption, preroutinization stage of the innovation process will all be needed as inventors and adopters learn more about the innovation and modify it to suit the needs of particular commissions.

The study of innovation in public organizations has grown in sophistication over the years. Ever finer distinctions have been made in types of innovations and their impacts. The number and types of influences to be considered in studying innovation have increased. Increasingly, researchers have been concerned not only with the act of adoption but the successful incorporation of the innovation into the organization's routines as one aspect of organizational decision-making processes. At the same time, researchers have grown if anything more hesitant to infer causal processes in their subject matter. An exploratory approach, often relying on case studies, seems to have to some extent replaced statistical studies, at least for the time being.

After an extensive review of the literature on innovation, 12 studies that seemed particularly rich in insights relevant to the implementation of new incentives in cost control were selected for detailed analysis. Each study is a contribution to empirical knowledge of innovation in organizations or the American states, or both. Analysis of the 12 studies aids in identifying factors that may help or inhibit adoption of incentive mechanisms in the commissions as well as weaknesses and pitfalls that may limit the relevance of the existing research. The earliest of the studies was published in 1967 [34].

There are 2 of the 12 studies that focus almost entirely on diffusion patterns--those of Walker [84] and Gray [31]. Two more use the diffusion approach as a jumping-off point to a study of adoption (Feller et al. [27] and [28]. Perry and Kraemer [63] use both a diffusion focus and an adoption focus that relies on the Downs and Mohr [23] innovation-decision design, or "adoptability" rather than adoption. Four of the studies are concerned with incorporation or implementation of innovations (Eveland et al. [24]; Yin et al. [92]; and Yin, [91]).

For a majority of the studies, the units of analysis were state or local public organizations (agencies), or both. Two of the studies use regional agencies (Daft and Becker, and Eveland). Walker, Gray, and Downs analyzed innovation in the American states. Feller discusses innovation at the level of the state [27] and the American city [28] as well as at the level of state and city agencies.

The governmental functions of the state and local agencies that were studied included housing, public education, library services, police, fire protection, sanitation, census taking, highway transportation, traffic engineering, and air quality control. Of the agencies studied, those dealing with regulation of air pollution probably have the goals, tasks, and environment most similar to public utility commissions.

The policy areas included in the state laws studied by Walker were health, welfare, education, conservation, planning, administrative organization, highways, civil rights, corrections and policy, labor, taxes, and professional regulation. Gray used legislation in education, welfare, and civil rights. Downs focused his attention on one area of state policy--juvenile correction.

Half of the studies focused on innovations in technological resources and their use. Computers and computer use were especially popular (Eveland; Feller et al.; Perry and Kraemer; Yin). The three studies that looked at innovation in the American states were concerned with policy outputs of the state legislatures. Daft and Becker used innovations in outputs and control processes. Bingham used technological resource innovations, including computers, but studied control process innovations, too. Hage and Aiken aggregated many innovations without specifying their type.

While appendix B provides a synopsis of these studies, a broad review is provided in the following pages.

Major Findings

The classical S-curve has been shown to exist for many innovative techniques diffused to American cities (Feller [28]). A spatial pattern of diffusion of policies to the American states has also found empirical support (Walker [84]). At the same time, diffusion paths have been found to differ for different categories of innovation. Gray [31] found differences in diffusion paths of state laws in the areas of welfare, civil rights, and education. Perry and Kraemer differentiated a number of diffusion patterns for computer applications in city governments. Both Feller et al. [27] and Walker found that the average speed of diffusion among states of new programs has been increasing.

Except in the studies of multiple innovations where an innovation score was assigned (Walker; Hage and Aiken; and Daft and Becker), and thus where "innovativeness" was an assumption of the study, innovativeness has not been found to be a consistent characteristic of states or cities (Feller et al. [27],[28] and Bingham [8]). Feller suggests that innovativeness may be consistent among states for particular functions of state government, but his findings are not fully supportive of this conclusion.

Yin et al. [92] found evidence for the existence of two separate processes of successful incorporation of innovations. The first he called the production efficiency process that led to service improvement. The second was the bureaucratic self-interest process, leading to incorporation. Similarly, Daft and Becker found quite different processes at work in school systems. They were able to distinguish a process of adoption of innovations that benefited college-bound students from a process benefiting terminal students. In addition, they found that administrative innovations and educational innovations followed entirely different paths to adoption. Since these were the only two studies where an attempt to make such distinctions in processes were made, the results are extremely interesting for the further study of innovation.

Few of the studies systematically differentiated among types of innovations. Where such differences were allowed for, they tended to play a role. Gray [31] and Feller et al. [27] [28] found that patterns of adoption differed by the functional field of government to which they applied. Yin found that type of innovation was associated with its successful incorporation. Hardware innovations were more likely to be successfully incorporated than software or data-processing ones [92]. He also found that transitive innovations (those aimed at creating changes in service practices with clients) were more likely to be successfully incorporated than other types [91].

Flaherty [29] found that each of the innovations studied appeared to have largely different antecedents. She had not classified her innovations according to a preset typology and did not offer a systematic explanation of the differing results except to say they tended to prove Downs and Mohr were right. Thus, all that can be said with some certainty is that different types of innovations are likely to diffuse in different patterns, to have different antecedents of adoption, and to have different influences, in and of themselves, on their use.

Higher cost innovations and those with higher costs relative to other factors were found to be negatively associated with innovation in several of the studies. The significant cost-related variables were costs (Bingham); relative advantage (Yin et al. [92]); costs relative to benefits (NRRI [56]); costs relative to other agency applications (Perry and Kraemer); and differential costs of client-oriented outputs (Downs).

Need for the innovation was a significant factor in explaining adoption of innovation (Perry and Kraemer; Yin [91]; Flaherty). Innovations that implied less change from the past were readily adopted (NRRI [56]; Perry and Kraemer). Innovations that were more divisible (Yin et al. [92]) and could thus be implemented incrementally were more likely to be incorporated. Reversibility was also positively associated with innovation success (Yin et al. [92]).

Risk and uncertainty (Perry and Kraemer) and doubts about effectiveness (NRRI [56]) were negatively associated with adoption of innovations. Similarly, "specificity of evaluation," or the degree to which an innovation's outputs can be measured directly, was negatively associated with adoption (Perry and Kraemer).

Other factors in the relationship between organizations and innovations that were found to be important were minimal competition between administrative and service applications (Yin [91]); visibility (Yin et al. [92]; and Perry and Kraemer); and task maintenance orientation of the innovation (Perry and Kraemer).

The amount of resources available to the adopting unit was repeatedly found to be positively associated with adoption. Overall resources (NRRI [56]; Feller et al. [27]; Daft and Becker), size of the adopting unit (Flaherty; Feller et al. [28]; Eveland and Rogers; and Bingham); and financial resources (Bingham, Eveland and Rogers; and Flaherty's case and study) were found to play a significant role.

Several variables related to resource arrangements were found significant: centralization (Hage and Aiken; Yin et al. [92]; Daft and Becker); formalization (Hage and Aiken; Yin et al. [92]); and organizational complexity (Hage and Aiken). Age of the commission and whether it had been restructured were found significant by Flaherty (study and case, respectively). Bingham found two output variables significant for predicting innovativeness in public schools. The first was lowness of dropout rates. The second was the percentage of college-bound graduates. Yin et al. [92] found that transitive innovations, aimed at creating changes in service practices with clients, were more likely to be successfully incorporated. The agency history of innovation was found significant by Yin et al. [92]. Presence of a service output goal was positively associated with successful incorporation (Yin [92]).

The presence of a key individual supporting the innovation was mentioned repeatedly as a factor in the innovation process (Bingham; Eveland and Rogers; Yin [91]). Professional qualifications or activities were positively associated with innovation in six of the studies. Bingham and Daft and Becker found staff "professionalism" an important explanatory variable. Flaherty found both professional activities of the commission chairman and the range of commission staff expertise important. Staff competence (Perry and Kraemer), the hiring of new, skilled individuals (NRRI), and practitioner training in use of the innovation (Yin et al. [92]) were found to be associated with adoption and use of innovations.

Daft and Becker found that the composition of the organization's staff was an important influence on the innovation process in school districts. Looking at factors that included support staff ratio and vocational teacher ratio, Daft and Becker found different factors correlated with different types of innovations. The finding is an interesting one for the study of state public utility commissions, the staff of which include several types of highly trained professionals. Hage and Aiken found a slight correlation between job satisfaction and the adoption of innovations.

Community size (Walker; Eveland and Rogers; Bingham; and Yin et al. [92]); wealth (Walker; Yin et al. [92]; Gray; Flaherty; and Bingham); industrialization (Walker, Flaherty), and urbanness (Eveland and Rogers); have all been found to be positively correlated with the adoption or incorporation of innovation. However, it is noteworthy that a positive relationship between the wealth of the state and adoption of innovation was not confirmed in the NRRI study of innovation in Arkansas [56]. Also, Downs found that none of these socioeconomic variables was correlated with adoption of his innovation, deinstitutionalization of juvenile offenders in the United States.

Downs did find that socioeconomic heterogeneity was correlated with deinstitutionalization; and he found that environmental stability, measured by the number of reorganizations of the state bureaucracy, was positively correlated with his innovation.

Walker found a positive correlation between state innovativeness and an index of judicial prestige in the states.

Demand for the particular government functions being studied was positively correlated with adoption of innovation in several of the studies. Bingham [8] found police computer use to be positively associated with community crime rates. Daft and Becker found the adoption of innovations in the educational system to be positively associated with the growth rate of the school districts and the educational level of the community.

Increases in ratepayers' bills, increased costs of energy, a desire to stabilize utility revenues, and lack of instate energy resources have been mentioned as influences on innovation in an NRRI study of innovation in the commissions [56]. Flaherty noted increases in the number of rate cases as a factor in innovation in her case study but did not use that as a variable in her correlational analysis.

The relationships of the adopting units to groups in their environments was found important to the innovation process in several of the empirical studies. General interest, activity, and influence of groups in the environment was stressed by Downs and the NRRI. Eveland and Rogers and Yin [91] found community support important. Flaherty identified high publicity about the commission, growth in consumerism, and legislative oversight of the commissions as important in her case study but did not use these variables in her general analysis of commission innovation.

Cooperation with the regulated utilities was identified by the NRRI study [56] as a factor aiding the adoption of innovation. Activity and proximity of supplier of innovations was positively associated with adoption of innovation (Feller et al. [27]; Bingham; Perry and Kraemer).

Yin et al. [92] found client participation a significant explanatory variable. Feller et al. [27] found the influence of intermediary associations important. Availability of external funding (Perry and Kraemer) and in particular, federal influence and financial support were found to be significant factors affecting the adoption of innovation (Bingham; Perry and Kraemer; and Feller et al. [27]). In all three of these cases, the studies themselves were federally funded. Yin's finding that no federal factors were positively associated with the successful incorporation of innovations may be equally important.

Factors related to the supply of innovations were mentioned in three studies. Yin et al. [92] suggested that a rich innovative environment played a role in the use of innovations; availability of technology was mentioned in an NRRI study; and the computer use already occurring in a city ("ADP score") was found by Bingham to be a factor in adoptions of computers by city agencies that were not already using them.

Factors related to the communication of innovations were mentioned in many studies, particularly in those dealing with innovation in the commissions.

Gray found that her model of interaction between adopting and non-adopting states successfully fit the actual pattern of behavior in the three areas of legislation she used. Downs found that an average of the extent of adoption of an innovation by the agency director's reference group of states was a good predictor of extent of adoption in the director's own state.

Contact with the federal agency acting as change agent, in this case the Census Bureau, was associated with adoption of Eveland's innovation. Eveland also found knowledge of the innovation by the potential adopters was correlated with adoption.

Flaherty and the NRRI studies found that a number of communication factors were important for the innovation process in the commissions. Factors Flaherty mentioned in her case study were the commissioners' search activities, their active interest in bringing innovations to the commissions, outside training courses for commission staff, and presentations of speakers invited to the commissions. Factors mentioned by the NRRI study were availability of information on energy shortages and awareness of studies supporting the innovation.

Summary

A review of empirical studies in innovation in public organizations and the American states has revealed a plethora of ideas and a scarcity of proven and accepted theory. In the case of the public utility commissions, Flaherty's work was inadequately anchored to the propositions and models that do exist. Partly because of this lack of grounding, her study does not do as much as it might to advance an understanding of adoption of innovation by the commissions.

The major conclusions that can be drawn from the studies are the following:

1. Classical diffusion patterns are often discernible in the spread of innovations among public organizations

2. The speed of diffusion of innovations among the American states is increasing
3. "Innovativeness" is not a consistent characteristic of American states or cities
4. Many different processes of innovation may exist, but relatively little conceptual or empirical work has been done to discover and map the processes
5. Adoption or implementation of innovation may be correlated with many factors, including those listed below:
    (a) characteristics of the organization's environment, such as the size, wealth, and urbanization of the community or state in which the organization is located
    (b) changes in demand for a governmental function
    (c) changes in the supply of innovations and activities of promoters of the innovation
    (d) relationship with interested organizations in the organization's environment
    (e) factors related to the communication of innovation, such as interaction of adopters and nonadopters
    (f) the type of innovation
    (g) the organization's perceptions of the innovation, including its modifiability, risk, need, relative advantage, and visibility
    (h) characteristics of the organization, especially the amount of its resources
    (i) the presence of a key individual or group of individuals within the organization who support the innovation
    (j) the professional expertise of individuals within the organization

CHAPTER 8

TOWARD FUTURE RESEARCH

The cost control problem was defined in this report as the socially undesirable propensity of a regulated firm to (1) underproduce its output, that is, display technical inefficiency; (2) combine factors of production in a manner that does not lead to the minimum attainable costs in light of the currently known technology and prices, that is, display allocative inefficiency; (3) pay for factors of production in excess of amounts that are needed to secure the factors' services; and (4) underinvest in new technology and organizational change so that future cost minimizing combinations of factors can be assured. The presence of this problem is related directly to the environment within which all the relevant decisions are made and the internal organization of utilities. The extent to which the cost control problem can be alleviated, or even eradicated, depends on the feasibility of new regulatory practices.

Ideally, any cost control policy recommendations must be preceded by a comparison of anticipated cost savings with the cost of implementing new practices. Such a comparison presupposes that the existing cost control problem lends itself to quantitative evaluation; indeed, that there exists a benchmark against which cost performance can be evaluated. In the absence of an agreed upon appropriate benchmark, it is necessary to test a number of hypotheses that relate various environments to cost performance. In other words, either there exists at the outset a clearly defined operational policy objective, or the choice of the preferred action emerges from the comparison of many options along the way.

The relevant hypotheses, such as those examined in the previous chapters, represent competing assertions concerning the true reactions of utility decision makers to alternate regulatory actions. Since test tube experiments are not possible, it is imperative that assumptions be made concerning test conditions and that past performance be rich enough to permit statistical inference. As chapter 4 illustrates, there is a general paucity of existing tests, and those that exist are not conclusive. Thus, despite a relatively long history of regulation, little is proven about the potential reactions of utilities to new regulatory practices.

Finally, the acceptability of new practices to regulatory agencies is not known. Indeed, little is known about the changing of practices in public agencies in general. The current state of knowledge, as reviewed in chapters 6 and 7, permits limited assessment of the extent to which the current practice can be altered. Without such knowledge, the design of new incentive mechanisms remains an uncertain endeavor.

The above suggests that the appropriate research strategy should include three separate, but related, activities. First, diminishing returns have not been reached in the conduct of theoretical research concerned with the behavior of utility decision makers and the structure of incentive mechanisms. In both cases, the purpose of research is to discover the logical structures of relationship. Without knowledge of such structures, it is not possible to predict either the potential impact and mode of implementation or incentive mechanisms or the behavior of utilities that may follow implementation. Second, there is a need for empirical research that would distinguish between plausible hypotheses and behavior that can be accepted as factual. It is only such behavior that can serve the needs of policy making. Finally, there is need for experimentation in the field. Once a particular incentive mechanism is identified as a candidate for implementation, the precise form in which it should be implemented and the changes that may be needed in each specific case should be anticipated through field tests.

APPENDIX A

INFORMATIONAL ASYMMETRY AND INCENTIVES

The objective of this appendix is to describe basic structures of information flows in incentive mechanisms. Such mechanisms can be implemented so that regulated firms behave in a manner deemed appropriate by both consumers and the regulatory authorities. In order to understand the structure of any specific incentive mechanism, it is necessary to examine first the role of information in the process of decision making. Information theory explains how informational asymmetries may generate nonoptimal behavior on the part of agents, the regulated firms, and why such asymmetries may give rise to the need for incentive mechanisms.

The Economics of Information

The economics of information is defined often as the production, dissemination, and manipulation of information in a market context. (See for example, [32], [37], [38].) Information is best regarded as events that generate changes in the characteristics of individuals' subjective probability distributions. These distributions are mathematical expressions that summarize the individuals' beliefs about outcomes of possible states of the world. Uncertainty is summarized by the dispersion of the individuals' subjective probability distributions.

This general view of information and uncertainty provides the essential framework for the vast majority of studies dealing with the microeconomics of information. The economics of information is actually an outgrowth of the economic theory of uncertainty. It views many of the economic variables of concern in neoclassical microeconomic theory, for example, price, costs, and output as random variables that can be characterized appropriately by probability distributions. In this context, economic units choose optimal acts, or behavior, by adapting to a given state of ignorance. The economics of information includes the further possibility of collecting new information prior to choosing a new "terminal" action. Thus, information gathering will necessarily cause individuals to revise prior subjective beliefs or prior probability distributions; these revisions of prior beliefs generate new probability distributions or posterior probability distributions. Given newly formed beliefs, economic units will again choose a course of action, taking into consideration additional increments of information. This general description of revised subjective beliefs is formally presented below.

When two or more economic units interact, the existence of heterogeneous information, or "asymmetrical" information, gives rise to each unit choosing an action based upon different information. This may lead to private as well as to social costs. In the context of regulation, asymmetrical information may lead to a social welfare loss that the regulatory commission could eliminate if it had access to the same information as the regulated firm.

The management of a regulated firm has access to information that cannot be acquired costlessly by the commission or by consumers. Furthermore, the objectives of the regulated firm may be multiple in scope; publicly held corporations are at least in part motivated by the need to maximize the market value of the firm, thereby maximizing the welfare of its security holders. In addition, and to the extent that managers make decisions, corporations act in the best interest of managers and of consumers. At the least, it is a function of regulators to obtain sufficient information to ensure that society's interests are not ignored and damaged excessively. Since collection of such information is costly and the cost of obtaining and processing complete information prohibitive, an important question that is not fully answered in this report concerns the optimal amount of resources that should be devoted to the acquisition of such information.

An alternative to acquiring the information necessary to achieve an ideal solution is to design mechanisms that will give firms the incentive to behave in a manner deemed appropriate by the commission. Incentive mechanisms are devices that can be used to change the behavior of economic agents. For example, assume a commission puts in place a simple reward structure that entitles the regulated firm to earn a return on investment above the currently allowed rate of return if the firm keeps costs from rising above some stated level. What effect will this have on the firm's choice of behavior? Putting the incentive mechanism in place reduces the firm's uncertainty about the outcomes of chosen actions. Based upon this new information, it will revise its subjective probability distribution and choose a new course of action. Thus, incentive mechanisms change the firm's perception of likely future outcomes. This will cause the firm to adapt its behavior to its new expectations.

A Basic Model

More formally, following Hirshleifer and Riley [38], it can be assumed that the set of acts, a, that could be chosen by the agent are a = (1,...,A). Let the set of states of the world be designated s = (1,...,S). It is assumed that the set of consequences of an act and all states of the world, C(a,s), are known to the individual. Also, the individual is assumed to have a prior probability distribution of initial beliefs regarding possible states of the world, $_s$. Furthermore, the individual can acquire information by receiving one of a known set of possible messages m = (1,...,M). The acquisition of a particular message may lead the individual to revise his probability beliefs, and consequently, to revise his choice of action.

These definitions provide the critical elements for analyzing the optimizing choices of the decision-making unit. A revision of an individual unit's probability belief after receiving a message, m, is conceptually represented by Bayes theorem. (See [39], p. 61.)

$$\pi_{s,m} = \Pr\{s/m\} = \frac{\Pr\{m/s\}\Pr\{s\}}{\Pr\{m\}}$$

$$= \frac{g_{m,s}\pi_s}{g_m} \quad , \qquad \text{(A-1)}$$

where

$_{s,m}$ = the individual units' probability beliefs after receiving message m,

Pr m/s = the probability of receiving message m in state s,

Pr s = the probability of state s occurring,

Pr m = the probability of receiving message m.

g_m = the probability of receiving message m,

$_s$ = the prior probability of state s

The term $g_{m,s}$ is a conditional probability, or "likelihood" function, and is related to g_m by:

$$g_m = \sum_{g=1}^{s} g_{m,s}\pi_s .$$

The essence of probability revisions is graphically shown in figure 2-1. The posterior probability reflects the weighted average of the likelihood function and the prior probability function; specifically, for each state s, the prior probability $_s$ is multiplied by the likelihood function $g_{m,s}$ and then rescaled by g_m to ensure that the probability integral is equal to unity. Figure A-1 shows how a revision of initial beliefs shifts the prior probability distribution function to a posterior probability function according to equation (A-1). This is a rigorous way of showing how an individual's initial beliefs are revised on the basis of new information.

The more important question that must be addressed is how do individuals choose acts? A general approach to answering this question can

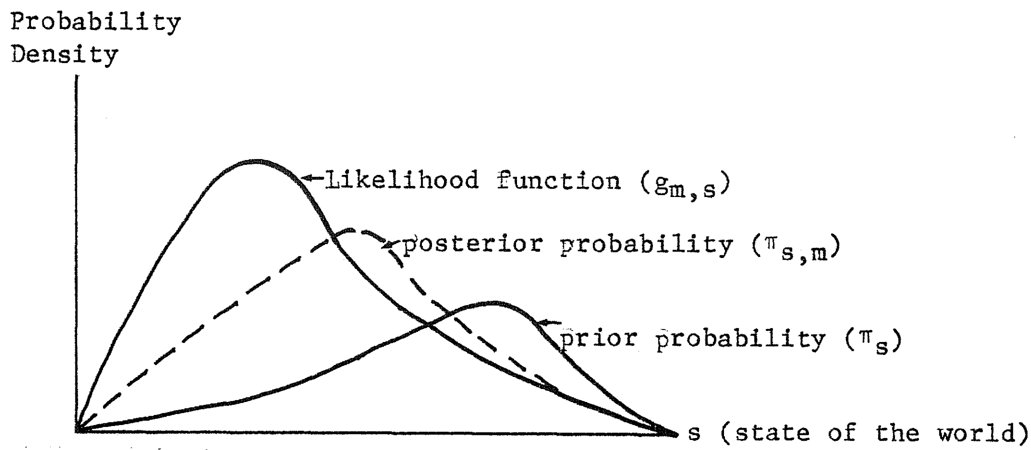


Figure A-1 The Bayesian relationship between prior and posterior probability functions (Bayesian probability recalculation)
Source: [38, p. 1394]

be developed using the Newmann-Morgenstern "expected-utility rule," according to which the utility of each act, u(a), is the mathematical expectation of the utilities of the associated consequences, $v(C_{aS})$. Letting $i = (i = 1,\ldots,s)$ represent the probabilities of the associated consequences, $C_a = (C_{a1},\ldots,C_{aS})$, the "expected-utility rule" may be written as:

$$U(a) = \pi_1 v(C_{a1}) + \ldots + \pi_S v(C_{aS})$$

$$= \sum_{s=1}^{S} \pi_s v(C_{as}). \qquad \text{(A-2)}$$

In this framework, the individual unit's objective is to maximize its expected utility, given a set of prior probability beliefs, $_s$. Such an optimization process can be formalized as:

$$\operatorname*{Max}_{a} v(a,\pi) = \sum_{s=1}^{S} \pi_s v(C_{as}). \qquad \text{(A-3)}$$

Suppose that maximizing equation (A-3) will determine an optimal terminal action, denoted as σ_o. This optimal action is based on the units' prior probability beliefs, π_s. Upon receiving a message, m, the decision maker will revise his probability beliefs. The revision of initial probability beliefs will yield a new set of beliefs, the posterior probability beliefs, $_{s,m}$. (See figure A-1.) Based upon the new posterior probabilities, the decision-maker will choose again via a process such as that described by equation (A-3). These activities determine his optimal action. Thus, upon receipt of new information, the decision maker will perform the operation of maximizing his expected utility according to equation (A-3), using his new set of probability beliefs.

The process of repetitive decisions based upon new probability beliefs is an essential part of informational economics. Assuming that new information is received at discrete points in time, the value of message m is:

$$\Delta_m = v(a_m, \pi_{s,m}) - v(a_o, \pi_{s,m}) \qquad \text{(A-4)}$$

where

Δ_m = the ex post value of message m,

a_m = the optimal action chosen after receiving message m,

a_o = the action chosen prior to receiving message m,

$\pi_{s,m}$ = the individuals' probability beliefs after receiving message m.

Thus, the ex post value of message m, $_m$, represents the expected gain from changing actions, estimated in terms of revised probabilities.

Informational Asymmetry

This basic framework provides a structure for conceptualizing the impact that information has on a decision-maker's choice of optimal terminal actions. Informational asymmetry causes parties to a contract to have different prior probabilities and/or different posterior probabilities. As new information is acquired, the parties to a contract will revise their prior beliefs and take actions to optimize their own welfare within the constraints imposed by the terms of the contract.

In the context of a regulatory environment, it is useful to imagine that there are two primary parties to a contract: (1) the regulatory authority, and (2) the regulated firm. Informational asymmetry arises because each party has access to information that the other does not. In addition, consumers may be viewed as a third party.

The asymmetries examined by Harris and Raviv [35] and Ross [72] arise because one party to a contract may not be able to observe acts chosen by other parties. For example, the regulatory authority may have the same information as the regulated firm at the time a contract is signed. However, the firm may take actions during the contract period that the authority may not be able to monitor. Furthermore, the authority may not be able to recognize the state of nature that evolved, and consequently, be unable to determine if the firm took an appropriate action, given the terms of the contract.

An additional source of informational asymmetry may arise after contracting, but prior to the firm taking its action. That is, both parties may recognize that the firm will acquire additional information before taking the actions necessary to fulfill the contractual obligations. In this case, incentives must be provided that will induce the firm to behave in an appropriate manner. This type of asymmetry arises because the authority does not receive the same information as the firm during the interval between the time the contract is signed and the time at which the firm takes its action. Informational asymmetries of this type have been studied by Weitzman [86] and Baron [6].

Myerson [53] examined asymmetries that arise prior to the signing of a contract. In a regulatory context, the firm may have access to technical information, for example, production technologies, costs etc., that the authority does not have. Since this type of asymmetry exists prior to contracting, the firm will negotiate terms that it considers satisfactory in light of its private information. In this case, the firm has the ability to refuse contractual terms if it determines that the reward structure for good performance is inadequate.

In a regulatory setting, there is a third party to a contract--the consumer. It is the authority's job to design socially optimal contracts. This implies that the authority has knowledge of how consumers "truly" value a service provided by utility companies. Green and Laffont [32] studied mechanisms that may be termed "truth revelation mechanisms." These revelation mechanisms are designed to extract information from the public rather than from information from an agent. The specialized character of such mechanisms makes their application to extracting information from an agent difficult, although perhaps not impossible. For this reason, truth revelation mechanisms are not included in the examples of incentive mechanisms presented in this report. These mechanisms are very complex and have been studied in an abstract form only. They were designed to induce consumers to reveal their true preferences for a commodity or service. Such information enables public authorities to determine the true social value of a public good. The ultimate objective of obtaining information from the public is to achieve a socially desirable outcome; in the terms of welfare economics, the ultimate objective is to achieve a "Pareto optimal" solution, described further below. If incentive mechanisms are put in place by the regulatory authority to induce optimal agent behavior, it is important to give some concrete meaning to the term "socially optimal state." The next section contains an elaboration of one definition of Pareto optimality developed by Green and Laffont.

Pareto Optimal Social States

Green and Laffont [32] studied the means by which privately held information can be elicited for public use and the effects of such elicitation process on economic efficiency. This approach allows determination of the extent to which problems of imperfect information can be mitigated. The general problem they address is that of obtaining private information useful for social decisions. Several of the models presented by Green and Laffont can be adapted for use by regulatory authorities in extracting privately held information from both utility companies and consumers.

To illustrate the basic idea, define

X_i = the consumption of a transferable resource by agent i,
d = the decision taken (or the project selected),
$\{0,1\}$ = the set of possible projects, where
d = 0 means "reject alternative project,"
d = 1 means "accept alternative project."

A preference function, or utility index, for individual i can be written as

$$U_i = U_i(d,X_i). \qquad \text{(A-5)}$$

Since d is a binary choice variable, equation (A-5) can be divided into two equations to represent the utility indices for the private good under the two possible decisions: $U_i^0(X_i)$ and $U_i^1(X_i)$, where the superscripts denote the value of d. Both functions are assumed to be strictly monotone increasing in X_i. $U_i^0(X_i)$ and $U_i^1(X_i)$ are ordinal representations of the underlying preferences. A transformation to cardinal measures can be represented as

$$\zeta = R \to R, \qquad \text{(A-6)}$$

given by

$$\zeta = \begin{bmatrix} 0 \\ U_i \end{bmatrix}^{-1} \qquad \text{(A-7)}$$

This transformation of the original utility function does not change the preferences of the individual. Writing

$$U_i^0 = \zeta\left(U_i(0,X_i)\right) = \zeta\left(U_i^0(X_i)\right), \qquad \text{(A-8)}$$

and

$$U_i^1 = \zeta\left(U_i(1,X_i)\right) = \zeta\left(U_i^1(X_i)\right) \qquad \text{(A-9)}$$

we see, by definition, that $U_i^1(X_i) = X_i$. Subtracting $U_i^1(X_i)$ from $U_i^1(X_i)$ yields $U_i^1(X_i) - X_i$, the amount of the private good that would make the individual indifferent between this amount of additional consumption, keeping the level at X_i, the individual's initial assumption level, or having the project adopted instead.
Define

$$(X_i) = U(X_i) - X_i, \qquad \text{(A-10)}$$

and refer to this quantity as the willingness to pay for the project at income level X_i.

In order to derive a solution to the preference revelation problem, it is necessary further to restrict the utility functions. We must assume that $V_i(X_i)$ is additively separable:

$$V_i(X_i) = V_i + X_i, \qquad \text{(A-11)}$$

where V_i is a constant. This indicates that the project is independent of any income effects.

Proceeding, the goal of the decision-making unit, in our case the regulatory authority, is to choose a social state in an optimal manner. If there are N individuals, $a = (d; X_i,\ldots,X_N)$ is a social state in which X_i is the consumption of individual i, $i = 1,\ldots,N$. In order to achieve Pareto optimality, it may be necessary for the authority to transfer some X_i to other individuals. Thus, it is assumed that the authority is able to draw on its own stock of X_i if necessary. This implies that the net transfer from the agents to the decision-maker is $\Sigma(X_i - X_i)$, where X_i is the initial endowment of individual i. This results from a feasibility requirement that $\sum_i V_i \leq \sum_i X_i$. Consequently, $U_i(a)$ in equation (A-8) respects the transfer to the authority. It is now possible to write

$$U_i(a) = U_i(d, X_i). \quad \text{(A-12)}$$

We are now in a position to define a Pareto optimal social state.

Definition: A social state a A, where A is the set of alternative social states, is said to be Pareto optimal if for any $a' \varepsilon A$,

$$U_i(a') \geq U_i(a),\ i = 1,\ldots,N, \quad \text{(A-13)}$$

which implies

$$U_i(a') = U_i(a),\ i = 1,\ldots,N. \quad \text{(A-14)}$$

The nature of the utility functions and social states described allow us to derive the following useful results:

1. If $\sum_i V_i > 0$, then the set of Pareto optimal states is

$$\{a/a = (d, X_1, \ldots, X_N),\ d = 1\ ,\}$$

2. If $\sum_i V_i < 0$, then the Pareto optima are

$$\{a/a = d, X_1, \ldots, X_N,\ d = 0\ .\}$$

These results indicate that the set of Pareto optima can be completely characterized by the public decision taken. The decision, d, is, therefore, Pareto optimal, if

$$d = 1 \text{ whenever } \sum_i V_i > 0,$$

and

$$d = 0 \text{ whenever } \sum_i V_i < 0.$$

In the case where $\sum_i V_i = 0$, the set of Pareto optima coincides with the entire set A.

This example gives a rigorous description of Pareto optimal social states. In essence, we can say that the authority should see to it that the agent (the company) undertakes a project if the consumers' willingness to pay, V_i, is greater than zero, and that they reject the project if their willingness to pay is less than zero.

Extending the results, we can say that the agent, if acting in the best interest of consumers, must be cognizant of the consumers' willingness to pay for additional services or new innovations. An important question must be now addressed. Current financial literature has built its foundation on the net present value rule that states that a firm should undertake a project if the project's net present value (NPV) is greater than zero. Then, and only then, will the welfare of the firm's owners (security holders) be maximized. Under certain assumptions, the net present value rule and the consumers' willingness to pay will lead to the same optimal choice of behavior. Whether or not this is true in a regulatory setting is unknown. A great deal of literature concludes that managers are behaving optimally if they are seeking out projects with positive NPVs and subsequently undertaking these projects. But, intuitively, it seems that some projects could be devised that could ultimately lead to social ruin, in spite of having a positive NPV.

APPENDIX B

ANALYSIS OF EMPIRICAL STUDIES OF ORGANIZATIONAL INNOVATION

Table B-1 lists the explanandum and explanans for each of the 12 studies reviewed in this appendix. The explanandum refers to the population that the studies attempt to explain and the samples actually used in the studies. The explanans are general categories for independent variables that attempt to explain the behavior of the explanandum. The explanans in table B-1 include three categories or "levels": communal, organizational, and individual.

The explanandum for three of the studies is the behavior of 48 American states. Walker and Gray used purely state level variables as the explanans, although Walker supplemented his empirical investigation with a model of the behavior of the individual decision makers within the state. Consequently, Downs supplemented state level variables in his study of state innovativeness with organizational and individual variables.

Feller, et al. [28] and Perry and Kraemer [62] used local government behavior as the primary level to be explained. However, they used local government suborganizations (agencies) as well and applied variables from the communal, organizational and individual levels to do the explaining.

The seven remaining studies tried to explain behavior on the organizational level. In three cases, however, the samples are a small subset of the population purportedly being studied. Hage and Aiken [34] and Daft and Becker [19] claim to shed light on innovation in organizations, yet their samples are limited to local semipublic social welfare organizations and public school districts. The sample size is also extremely small for both these studies. Eveland and Rogers [24] attempt to illuminate the innovation process for public organizations in general,

TABLE B-1

POPULATIONS, SAMPLES, AND LEVELS OF ANALYSIS IN 12 EMPIRICAL STUDIES OF ORGANIZATIONAL INNOVATION

	Explanadum			Explanans		
	Population	Sample	Sample Size	Communal	Organizational	Individual
Walker	All American states	Most states	48	(X)		
Gray	All American states	Most states	48	(X)		
Downs	All American states	Most states	48 (N=16 for some variables)	(X)	X	X
Bingham	Local government organizations	Local government organizations	4 types of agency; 310 cities	X	(X)	
Eveland	Public organizations	Regional government organizations	257		(X)	X
Feller et al. [28]	State organizations	States and state agencies	50 states; 2 types of agency	(X)	(X)	X
Feller et al. [29]	Cities	Cities and city agencies	465 cities; 4 types of agency	(X)	(X)	X
Daft and Becker	Bureaucracies	Regional government organizations	13	X	(X)	X

TABLE B-1

(Continued)

	Explanandum			Explanans		
	Population	Sample	Sample Size	Communal	Organizational	Individual
Hage and Aiken	Organizations	Local governments; local government organizations	16		(X)	
Perry and Kraemer	Local government	Local government; local government agencies	713	X X	(X)	X
Yin et al. [92]	State and local government organizations	Case studies of state and local government organizations	48	X	(X)	X
Yin [91]	Local government organizations	Local government organizations	3 types of organizations; 16 cities	X	(X)	X

( ) = level explained

but their sample is limited to one type of regional, special purpose planning body.

Of the seven studies focusing on organizational behavior, only Hage and Aiken used organizational level variables alone as explanans.

Boundaries: The Innovations

Table B-2 shows innovations by source for the 12 studies where that information was made explicit. The original source is indicated. The original source may or may not be the same as the marketer of the innovation.

The studies of the American states all emphasized peers as the source of innovative policies. Peer associations were mentioned by Walker [85] as a source of information about innovations, although not necessarily as the original source. Feller et al. emphasized the role of private suppliers in his studies. Bingham [8] used at least one innovation developed and marketed in the private sector; the nature of many of the other innovations suggests that a large majority originated in this sector. Eveland's innovation in computer software was developed as well as marketed by the federal government. Both Walker and Gray [31] emphasized the role of the federal government in developing and disseminating policies. Feller et al. and Yin discussed the role of the federal government as a disseminator of policy though not necessarily as the original source.

The only study to use an innovation explicitly developed by an independent research center was individualized instruction in public schools, used by Bingham.

None of the studies discussed innovations that originated with the organizations' clients; nor were innovations that originated with control organizations discussed, at least explicitly. Yet for regulatory agencies, many of the innovations adopted originated through legislative or judicial mandate.

TABLE B-2

STUDIES OF INNOVATION IN PUBLIC ORGANIZATIONS BY SOURCE OF INNOVATION

	Peers	Research Centers	Peer Assn.	Private Suppliers	Clients	Control Organizations	Federal Government
Bingham (local agencies)		X		X			
Daft and Becker (regional agencies)							
Downs (American states)	X						
Eveland & Rogers (public organizations/ regional agencies)							X
Feller [27] (state agencies)				X			
Feller [28] (city agencies)				X			
Gray (American states)	X						X

TABLE B-2

(Continued)

	Peers	Research Centers	Trade Assn.	Private Suppliers	Clients	Control Organizations	Federal Government
Hage and Aiken (organizations/ local organiza- tions)							
Perry and Kraemer(local agencies)							
Walker (American states)	X						X
Yin [92] (state and local agencies)							
Yin [91] (city agencies)							

Table B-3 categorizes the innovations considered in the 12 studies by their initial focus. The table uses the general model presented in chapter 1 rather than Daft and Becker's model. The information in Table B-3 indicates that a majority of the studies restrict their attention to innovations in material resources or operational processes that have a strong material component.

Three studies distinguish between "technology" and how technology is used. Eveland and Rogers speak of closely linked "tool" decisions and "use" decisions. Perry and Kraemer treat computer applications as a complex technological package composed of people, apparatus, and techniques. They emphasize that they are not dealing with computers solely as physical entities. Bingham distinguishes between physical products and processes. A "process" innovation requires changes in methods and may or may not include a product. In each study, using the terminology of this paper, the authors discuss the interconnection of physical resources and operational processes.

Yin, in his study of routinization, categorized each innovation according to "the breadth and nature of its functional applications." Yin defined "task specific" innovations as those having only one possible application and "task diverse" innovations as those which could be used for at least two applications. Two kinds of task diverse innovations were distinguished: those which are distinctly limited to service applications and those which have both service and administrative applications. Yin's categorization may well be a useful addition to the conceptualization of innovation types.

Downs [21], Gray, and Walker use outputs of the state policy-making process as their innovations. Gray and Walker use the outputs of the state legislatures. However, Downs's innovation is operationalized at another level. Deinstitutionalization in Downs's study has not only been set as a policy but to various extents has been implemented.

Only Daft and Becker partition their innovations by whether they have to do with administrative processes or organizational outputs. Further,

TABLE B-3

FOCI OF STUDIES OF INNOVATION
IN PUBLIC ORGANIZATIONS

	Control Processes			Resources		Resource Arrangements	Outputs
	Opera-tional	Admin-istrative	Stra-tegic	Machines	People, Information, Other		
Bingham (local agencies)	4*,D**			4,D			
Daft and Becker (regional) (agencies)		A					78,A
Downs (American states)							1
Eveland & Rogers (public organi-- zations/region-- al agencies)				1			
Feller [27] (state agencies)				4,D			
Feller [28] (city agencies)				40,A/D			
Gray (American states)							12,D

TABLE B-3

(Continued)

	Control Processes			Resources		Resource Arrangements	Outputs
	Operational	Administrative	Strategic	Technological	People, Information, Other		
Hage and Aiken*** (organizations/ local organizations)							
Perry and Kraemer (local agencies)				10,D			
Walker (American states)							88,A
Yin et al. [92] (state and local agencies)				140,D			
Yin [91] (city agencies)				6,D			

* Numbers are number of innovations studied in each category.
** D = Disaggregated; A = Aggregated; A/D = Aggregated by type of organization.
*** Innovation types were not distinguished.

they divide outputs into two categories: those benefiting terminal students in the high schools and those benefiting students who are going to college.

None of the studies deal explicitly with changes at the strategic level of the organization or with innovations initially focused on personnel or processes involving people exclusively (rather than machine-people combinations). None of the studies deal with resource arrangements as the initial focus of the innovation, although Eveland and Yin are concerned with the ultimate impact of innovations on resource arrangements.

The Dependent Variables

Table B-4 shows a progression from the type of study to the concepts defined in each study to each study's operational definitions.

In the nine studies that dealt with diffusion and/or adoption of innovations, rather than with their incorporation, the emphasis was on the Rogers and Shoemaker [70] definition of "innovation." Walker, for example, defines an innovation as "a program or policy which is new to the states adopting it, no matter how old the program may be or how many other states may have adopted it" [84, p. 881]. Downs, Gray, Feller, and Eveland and Rogers (as a basis for a study of incorporation) also use this view of innovation. Bingham, Daft and Becker, and Perry and Kraemer employ definitions that rely more on objectively determined "newness." Hage and Aiken do not differentiate between subjective and objective "newness." They define "program change" as "new programs or services" added to the organization [34, p. 503].

The 12 studies developed a variety of operationalizations of the several concepts to be measured. Figure B-1 views complex measures of diffusion patterns, innovativeness, incorporation, and adoptability as stemming from the simple dichotomous measure of adoption versus nonadoption. Table B-5 shows each operational definition used by the authors.

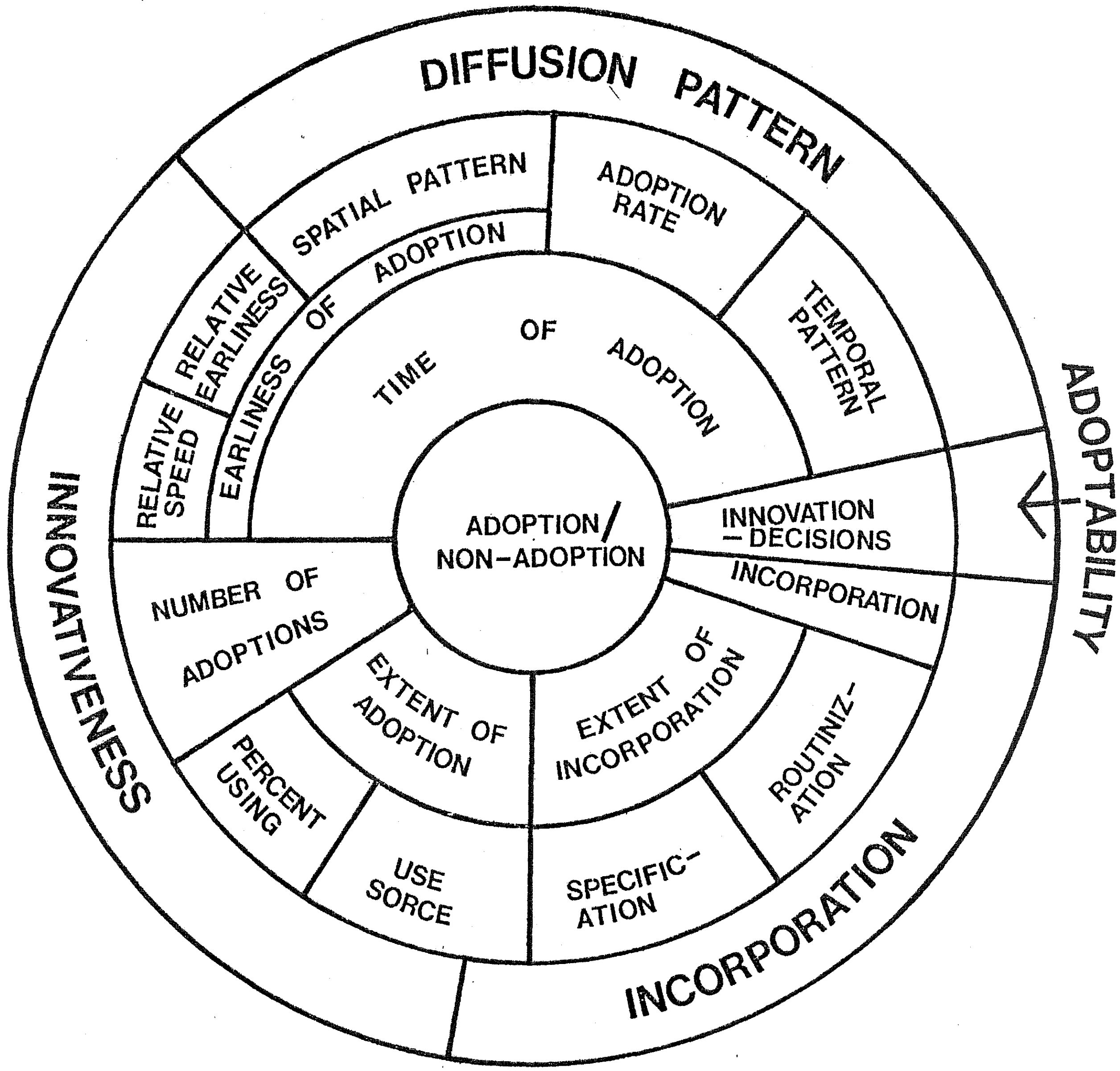


Figure B-1: Relationship of operational
definitions to focus of research

TABLE B-4

CONCEPTUALIZATION AND OPERATIONALIZATION OF DEPENDENT VARIABLES IN 12 STUDIES OF ORGANIZATIONAL INNOVATION

Author	Type of Study	Concept	Definition	Type of Operational Definition
Bingham	Adoption	"Adoption"	The first or early use of an idea by one of a set of organizations with similar goals	- Adoption/non-adoption - Extent of adoption (use)
Daft and Becker	Adoption	"Innovation"	Organizational change that is new to a group of organizations that share the same goals and technology	- Number of adoptions
		"Innovativeness"	Frequent early adoption	
Downs	Adoption	"Policy innovations"	Policies new to the political units (here states) adopting them no matter how long they have been around or how many comparable units may have adopted them--(Walker, 1969)	- Extent of adoption
Eveland & Rogers	Incorporation	"Innovation process"	Sequence of changes in key characteristics of an organizational system	- Adoption/non-adoption - Time of adoption - Extent of adoption

TABLE B-4

(Continued)

Author	Type of Study	Concept	Definition	Type of Operational Definition
Gray	Diffusion/ adoption	"Innovation"	An idea perceived as new	- Temporal pattern of adoption
		"Diffusion"	The process by which something spreads, consisting of the communication of a new idea in a social system over time	- Earliness of adoption
Hage and Aiken	Adoption	"Program change"	New programs or services added to the organization	- Number of adoptions
Feller et al. [27]	Diffusion/ Adoption	"Diffusion"	Initial and subsequent adoption of the techniques in our study by the organizational units in our sample	- Time of adoption - Earliness of adoption
		"Receptivity"	Defined operationally by earliness and extent of adoption	- Extent of adoption (use)

TABLE B-4

(Continued)

Author	Type of Study	Concept	Definition	Type of Operational Definition
Feller et al. [28]	Diffusion/ adoption	"Diffusion"	"Rate and extent of acceptance and use of innovations among a class of adoptors and the processes by which individual adoptors interact with one another and with other change agents" (p. 2)	- Temporal pattern - Number of adoptions - Earliness of adoptions
		"Adoption"	"Behavior of a single adoptor with respect to acquisition of either a single innovation or a group of innovations"	
		"Innovativenes"	Rank order among a set of adoptors (p. 2)	
Perry and Kraemer	Diffusion/ adoptability	"Diffusion"	Defined operationally	- Extent of adoption - Rate of adoption - Pattern of adoption
		"Adoptability"	Probability that a local government will decide to adopt a particular innovation given the characteristics of the innovation and given the innovation's relationship to organizational characteristics	

TABLE B-4

(Continued)

Author	Type of Study	Concept	Definition	Type of Operational Definition
Walker	Diffusion	"Innovation"	"A program or policy which is new to the states adopting it, no matter how old the program may be or how many other states may have adopted it" (p. 881)	- Earliness of adoption - Spatial pattern of adoption
Yin [92]	Incorporation	"Successful innovative experience"	Degree to which an outcome of attempts at technological innovation were found to be beneficial	- Incorporation
		"Incorporation"	"The ways in which new changes became a per--manent part of an organization's bureaucratic fabric" (p. 13)	
Yin [91]	Incorporation	"Routinization"	Stage in the life history of an innovation at which the innovation no longer is an innovation but has become part of the common services routinely provided (p. 5)	- Routinization

TABLE B-5

OPERATIONAL DEFINITIONS OF DEPENDENT VARIABLES

A. DIFFUSION

Type of Indicator	Indicator Subtype	Measurement	Author
Time of adoption	General temporal pattern	- Diffusion curves: plots of frequencies and cumulative frequencies of adoption over time	Feller [27]
		- Dumulative proportion of states that had adopted a new law	Gray
	Adoption rate	- Number of years for a computer application to diffuse to three percent of the local government population	Perry and Kraemer
		- Number of adoptions of computer applications per year over the 10 consecutive years of most rapid adoption for that innovation	Perry and Kraemer
	Elaborated temporal pattern	- Cluster analysis of six variables: (1) mean year of adoption (2) standard deviation (in years) of the adoption distribution (3) peakedness (kurtosis) of the distribution (4) skewness (5) range (in years) of the adoption period (6) cumulative percentage of adoptions	Perry and Kraemer
	Spatial pattern	- Varimax factor analysis, using a matrix of pair-wise comparisons of all state innovation scores (see below) on all 88 issues	Walker

TABLE B-5

(Continued)

B. INNOVATIVENESS

Type of Indicator	Indicator Subtype	Measurement	Author
Adoption versus non-adoption		- Whether or not agency had filed a letter of intent to use the innovation	Eveland and Rogers
		- Innovation was or was not used	Bingham
Number of adoptions		- Number of new programs added in a five-year period ("rate of program change")	Hage and Aiken
		- Number of innovations adopted	Feller et al. [28]
		- Number of innovations adopted within a time period appropriate to each innovation	Daft and Becker
Time of adoption		- Date of "letter of intent"	Eveland & Rogers
Earliness of adoptions	Relative earliness of adoptions	- First 10 percent of adopters	Feller et al. [28]
		- First 10 states adopting a law	Gray
		- For each innovation, classification of states as early adopters, followers or late adopters based on inspection of diffusion curves	Feller et al. [27]
		- Comparison of rank order of a city's innovativeness in any given field with its rank order in any other field, using (1) total number of innovations adopted and (2) weighted measures of "novelty" of the innovation	Feller et al. [28]

TABLE B-5

(Continued)

B. INNOVATIVENESS

Type of Indicator	Indicator Subtype	Measurement	Author
Earliness of adoptions (cont.)	Relative speed of adoption	- Correlation coefficients computed between the date a given technology was first adopted by any city and the date of its adoption by each city that used it, and the computed time lag for each of the technologies adopted by the city	Feller et al. [28]
		- Innovation score prepared by (1) counting the number of years that elapsed between the first and last recorded legislative enactment of a program, (2) giving each state a number for each list of dates that corresponds to the percentage of time that elapsed between the first adoption and its own adoption, (3) subtracting the average of the sum of the states' scores on all issues from 1.000	Walker
Extent of adoption	"Use scores" from survey items	- Sum of "yes" responses to a number of questions concerning the innovation's use (e.g., housing computer use score= sum of "yes" responses to 15 questions concerning computer use in housing)	Bingham
		- One point was awarded for each different user agency in the area served by the regional agency that the agency reported was using the system	Eveland and Rogers
		- Ten-item index of implementation	Eveland and Rogers

TABLE B-5

(Continued)

B. INNOVATIVENESS

Type of Indicator	Indicator Subtype	Measurement	Author
Extent of adoption (cont.)	Percentage of potential users	- Percentage of total using the innovation (e.g., percentage of elementary teachers using individualized instruction)	Bingham
		- Deinstitutionalization rate: number of juvenile offenders placed in state-operated or state-funded community-based residential facilities divided by total number of offenders in these faciliies, plus those in institutions, multiplied by 100	Downs
C. ADOPTABILILTY			
Innovation-decisions		- Adoption versus nonadoption of each innovation by each member of the sample	Perry and Kraemer
D. INCORPORATION			
Level of specification		- Five-point scale on which each of 53 applications of the innovation was rated	Eveland and Rogers

TABLE B-5

(Continued)

D. INCORPORATION

<table>
<tr><th>Type of Indicator</th><th>Indicator Subtype</th><th>Measurement</th><th>Author</th></tr>
<tr><td>Successful incorporation</td><td></td><td>- Reader-analysts of case studies categorized the outcome of each innovative experience as follows:
<table><tr><td></td><td></td><td colspan="2">Incorporated</td></tr><tr><td></td><td></td><td>Yes</td><td>No</td></tr><tr><td rowspan="2">Service Improvement</td><td>Yes</td><td>Success</td><td>Failure</td></tr><tr><td>No</td><td>Failure</td><td>Success</td></tr></table></td><td>Yin et al. [92]</td></tr>
<tr><td>Passages and cycles</td><td></td><td>- Number of passages or cycles achieved by the innovation. (A "passage" occurs when a formal transition from one organizational state to another has taken place; a "cycle" is an organizational event that occurs repeatedly during the lifetime of the organization.) Passages and cycles for five types of resources were counted: budgetary resources, personnel resources, training programs for service personnel, organizational maintenance, and supply and governance</td><td>Yin [91]</td></tr>
</table>

Diffusion Patterns

Where the studies were attempting to discern and explain patterns of diffusion, time of adoption was the fundamental data. Simple diffusion curves showing frequencies and cumulative frequencies were used. Perry and Kraemer developed a highly elaborate measure of the temporal pattern of diffusion, using a cluster analysis of six statistics. Perry and Kraemer also used measures of adoption rates to represent the speed at which different innovations diffused. Only Walker looked at the spatial pattern of diffusion.

Innovativeness

Adoption versus nonadoption was used as a measure of innovativeness in two studies. Eveland and Rogers used it in conjunction with several other measures. Bingham used it for two of his eight innovations. This is an extremely crude measure, particularly in a cross-sectional analysis.

The number of adoptions was used by Hage and Aiken and Feller et al. as a measure of innovativeness. Daft and Becker also used a number of adoptions but adjusted for newness by counting only those adoptions that took place within one of two time periods appropriate for the innovations.

Time of adoption was used by Eveland and Rogers, along with other operational definitions, and served as the basis for measures of earliness of adoption. Relative speed of adoption was measured by Feller et al. and Walker, using the date of the first adoption as the starting date.

Extent of adoption was used by Bingham, Eveland and Rogers, and Downs. Bingham and Downs used the percentage of actual to potential use to represent extent of adoption. Eveland and Rogers awarded points for actual users within the area of their regional agencies. Bingham, Eveland and Rogers also constructed and used indices of extent of adoption that may be viewed as measures of extent of incorporation. Bingham summed the "yes" responses to a number of questions on the innovation while Eveland and Rogers used a 10-item index of "implementation."

Incorporation

Three states were defined in Yin [91] as measurement of successful innovative efforts. In state one, "Success 1" was assigned by a reader-analyst who judged that the innovative effort resulted in both incorporation and service improvement. "Success 2" was assigned if it was found that the innovation had not been incorporated but had not led to service improvement. The other two states were both classed as "failure." One problem with Yin's categorization is that it is not made clear whether and to what extent "Success 2" was preceded by actual trial.

Eveland and Rogers [24] and Yin et al. [92] attempted to measure what one might call "extent of incorporation." Eveland and Rogers used a five-point scale of specification levels, and Yin used a measure of "passages and cycles" successfully negotiated by the innovation.

Adoptability

Perry and Kraemer were the only researchers who attempted to use adoptability as the dependent variable. They used adoption versus nonadoption of each innovation for each member of their sample to form matrices of innovation-decisions. Each cell in the matrix was an "innovation-decision."

The Independent Variables

Table B-6 groups the independent variables used in the 14 separate investigations, comprising the 12 studies, and can be grouped into eight categories: (1) characteristics of the innovations, (2) characteristics of the organizations, states, or cities, (3) organizational perceptions about innovation, (4) characteristics of individuals with the organization; (5) communication factors; (6) characteristics of the environment; (7) "performance gap" factors; and (8) characteristics of decision units.

TABLE B-6

SCOPE OF THE INDEPENDENT VARIABLES

	Innovation	Organizational Perceptions	Organization	Individuals	Communication	Environment	"Performance Gap"	Decision Units
Bingham	X	0	X	X	0	X	X	0
Daft and Becker	0	0	X	X	0	0	0	0
Downs	0	0	X	X	X	0	X	0
Eveland & Rogers (Stage 1)	0	0	X	0	X	X	0	0
Eveland & Rogers (Stage 2)	0	0	X	X	X	X	0	0
Feller et. al [27]	X	0	X	X	X	X	0	X*
Feller et. al [29]	X	0	X	X	X	X	X	0
Gray	X	0	X	0	X	0	0	0
Hage and Aiken	0	0	X	X	0	0	0	0
Perry and Kraemer (diffusion)	X	0	X	X	X	X	0	0
Perry and Kraemer (adoptability)	0	X	0	X	X	X	X	0
Walker	0	0	X	0	X	X	0	0
Yin et al. [92]	X	0	X	X	0	X	0	0
Yin [91]	X	X	X	X	0	X	0	0

X = variables in that category were included in the study
0 = variables in that category were not included in the study

Seven of the studies differentiated among types of innovations. However, only Perry and Kraemer used innovation attributes as independent variables to any great extent. Perry and Kraemer measured the relationship between organizations and innovations in their study of "adoptability." Hage and Aiken's study, on the other hand, used only organizational characteristics and aggregate staff characteristics, such as job satisfaction.

Characteristics of individuals or aggregate characteristics of individuals were looked at in 11 of the 14 investigations. The exceptions were the studies by Gray, Walker, and the first stage of the Eveland and Rogers study. Aspects of communication links between the organization or state and its environment were included in 9 of the 14 investigations. Interaction among the states is the driving force in Gray's model of diffusion.

Environmental variables were studied in 10 of the 14 investigations. Downs distinguishes between characteristics of the unit of analysis and the environment of the unit of analysis. Downs looked at general socioeconomic characteristics of the states that in other studies were treated as environmental factors predicting the innovative behavior of organizations. Daft and Becker, Gray, and Hage and Aiken did not use environmental variables.

Four of the studies incorporated some variant of the "performance gap" concept into their studies as independent variables. For example, Downs used four measures of "performance gap." Bingham used elements of "specific demand" for innovation to explain adoption of his product and process innovations. Feller et al. [28] defined performance gaps in physical terms, for example, tons of refuse per household per year were taken as an indication of demand for sanitation services. Only Perry and Kraemer approached the performance gap from the point of view of the organization. Their definition of "need" for an innovation included both objective factors and subjective assessments of the organizational requirements to be met by the innovation [62, p. 124].

Feller et al. [27] focused on characteristics of "decision units" in his study of adoption of innovations by state agencies. The size of the decision unit, its prestige, and the characteristics of the decision makers in the unit were used as independent variables.

The Models

As befits the exploratory bent of our studies, more than 12 models were proposed. Feller et al. [27] and [28] used diffusion models as a stepping stone to his adoption models. Eveland and Rogers used an adoption model to pave the way for a model of incorporation. Walker investigated a diffusion model and then proposed an adoption model to complement and enrich his findings. Perry and Kraemer, on the other hand, used entirely separate models in their studies.

Diffusion Models

Walker, Gray, and Perry and Kraemer present detailed diffusion models. Feller et al. [27] used frequencies and cumulative frequencies of adoption of innovations by state agencies to categorize states by earliness of adoption. In his 1976 study of adoption of innovations by the cities, he investigated the characteristics of the diffusion curves both within and across functional areas of city government.

Walker looked at spatial patterns of diffusion among the American states. In his model, state "innovativeness," aggregated over many innovations and many years, is reflected in identifiable geographical groupings. He suggests that some states are national leaders, however, regional leaders and their followers may also be identified.

Gray's model of diffusion relies on temporal patterns in three legislative areas and uses interaction between adopters and nonadopters over time to predict the cumulative proportion of adopters.

Perry and Kraemer's study is notable for its disaggregation of diffusion patterns. Ten separate patterns of adoption of computer application were identified and were used as one of the dependent variables. This model views diffusion outcomes as the result of innovation attributes and policy interventions affecting the behavior of local governments. Furthermore, Perry and Kraemer used innovation attributes and policy interventions as their categories of independent variables because both kinds of variables may be manipulated by policy makers. Moreover, such factors as environmental characteristics and organizational structure are not easily manipulated.

Both Perry and Kraemer's model and their findings are of particular relevance to the cost control project:

> If innovation attributes do indeed influence diffusion, then they should be considered when choosing incentive systems to encourage diffusion or when designing or redesigning an innovation for a particular system of potential users. In essence, innovation attributes represent a potentially manipulable, additional aspect of diffusion processes for consideration by policymakers. [62, p. 20]

They make a similar argument for policy intervention variables:

> Policy interventions may be viewed as [attempts] "to manage infrastructure (manipulate fields) and in so doing make it desirable for other organizations to behave in ways they would not have otherwise [77, p. 20]....Since policy interventions are, like innovation attributes, multi-dimensional, their study can provide insight into diffusion tactics....[77, pp. 25-26]

Adoption Models

Downs, Feller et al., Hage and Aiken, and the first phase of the Eveland and Rogers study used the sparest, least complex models among the studies chosen for analysis. Their models are in essence conceptual frameworks providing categories within which to group the independent variables. Hypothesized correlations are carefully explored, but for the most part interrelationships among the variables are not.

In contrast, notions of process are built into the Bingham, Daft and Becker and Walker adoption models. Bingham finds that organizational characteristics and characteristics of the organizational environment act directly on the local government adopter, while characteristics of the general community environment act indirectly through specific demands [8, p. 516]. A peculiarity of the Bingham model is that although "product" and "process" innovations are differentiated, the model does not easily allow comparison of the variables or of the processes of innovation adoption.

Daft and Becker [19] see incentives and innovation ideas in the environment resulting in proposals for innovation within the organization. With the aid of "enabler variables," the proposals are transformed into decisions to innovate.

Walker [84] sees adoption of a new program as resulting from a perceived performance gap. Awareness of adoption of a new program "by a state taken as a legitimate reference group" reduces a state's "sense of relative well being." Agitation for adoption of a similar program then develops. Expert support and evidence favoring the new program feeds the agitation, and as a result, the state adopts. The specialization and professional development of the state government bureaucracy adds to awareness of expert support and awareness of evidence favoring new programs and thus indirectly fosters adoption [84, p. 898].

Adoptability

Perry and Kraemer's study of adoptability is aimed at "the assessment of the type and nature of environmental and organizational influences on the probability that a particular (computer) application will be adopted by an organization" [62, p. 120]. Four abstract dimensions were selected that were believed to encompass key determinants of adoptability. The first was the relation of the innovation to the organization's domain, or sphere of activity. Domain includes the technologies used, populations served and services rendered by an organization ([62, p. 123] citing [77]). The second conceptual dimension was integration, or "factors which facilitate the discovery and implementation of an innovation by an organization" (Ibid.). A risk dimension was intended to incorporate "factors which influence the perception that the innovation will produce expected results," including relative financial costs, the availability of slack resources and the specificity with which the innovation can be evaluated [62, p. 124]. Finally, the need for the innovation was conceptualized as including both objective factors and subjective assessments of the organizational requirements to be met by the innovation. Only the relationship of organizational domain and adoptability were hypothesized to be linear, increasing steadily as the innovation related less to a task orientation and more to a maintenance orientation. An S-shaped pattern was hypothesized from low to high integration. Adoptability was predicted to drop precipitously at some point as the level of risk increased, and need was hypothesized to have little effect on adoptability at low levels but to increase considerably as need became more severe.

Incorporation

"Incorporation," routinization," and "specification" are the terms used by Yin et al. [92] and Eveland and Rogers [24] for what happens to an innovation after adoption.

Eveland and Rogers propose a process model of innovation that sees the innovation moving through five stages of increasing specification. These five stages are listed below:

1. Agenda setting: "The stage at which the general problems of the organization are defined and commonly recognized" [24, p. 72]. Eveland and Rogers view this stage as a continuous process and thus not, strictly speaking, part of the innovation sequence. However, they include agenda setting in their model, as a necessary precursor of the sequence.

2. Matching: "The stage at which a general problem from the agenda and a possible solution are brought together within the organization" (ibid.). "Matching" seems to approximate the concept of "adoption" used in other studies.

3. Redefining: "The stage at which the primary attributes of the innovation are defined in terms relevant to members of the organizational unit." [24, p. 73]

4. Structuring: "The stage at which participants engage in conjoint activity to establish the innovation within the structure of the organizational unit....Structuring is characterized by the emergence of recognizable organization arangements [sic] for the innovation." (ibid.)

5. Interconnecting: "The stage at which the organizational structure formed for the innovation defines its relationships with the rest of the organization." [24, p. 74]

These five stages are viewed as sequential. However, backtracking may occur, and the sequence may move at different speeds. The terminal point of the innovation process is "the point at which the innovation is no longer recognizable to the participants in the organization as an 'innovation,' but rather is seen as a normal part of the organization" [24, p. 74].

Eveland and Rogers hypothesized that different variables would have different impacts at different stages of specification. These impacts include the following: (1) professionalism at the agenda-setting and matching stages, (2) system support and innovativeness at the matching stage, (3) accountability and resources at the redefining stage, and (4) communication and effectiveness at the structuring and interconnection stages.

Yin [91] views innovations as having a "life history" that begins with their adoption and ends at the terminal point defined by Eveland and Rogers. There are three stages in Yin's model: (1) improvisation, "the initial period during which an innovation began to operate following adoption"; (2) expansion, "marked by both the continued growth of the innovation and the achievement of several passages or cycles"; and (3) disappearance, the period in which the innovation continued to be used but was no longer thought of as an innovation [91, pp. 73-74]. "Routinization" was defined in terms of "passages" negotiated and "cycles" survived for five types of resources needed to sustain the innovation. A passage was said to occur when a formal transition from one organizational state to another took place. For example, an entry related to the innovation in the agency's operating manual might be a passage. A cycle was defined as an organizational event that occurred repeatedly, such as preparation of the agency budget. The more cycles an innovation passed through, the more routinized it was considered to be [91, p. 57].

The Yin et al. [92] study relied largely on a static model in which individual variables were sequentially correlated with successful incorporation, nonincorporation, and failure. The variable categories included innovation attributes, external conditions, agency characteristics, and implementation factors [92, p. 25]. The approach was thus quite similar to those of many adoption studies, despite Yin's criticism of the social interaction approach as a way of understanding what organizations do with innovations.

However, sandwiched in between the bundles of determinants, Yin proposed two separate models of innovation. The first model was a production efficiency process leading mainly to service improvement. The second model was a "bureaucratic self interest process" leading mainly to incorporation. Yin suggested that different variables would be associated with each process [92, pp. 87-108].

Research Design

Table B-7 shows the objectives, types of research, and research strategies of the 12 studies.

Churchill [14] distinguishes three research types. In exploratory research, the emphasis is on discovery of ideas and insights. Descriptive research is "concerned with determining the frequency with which something occurs or the relationship between two variables." Typically, he says, descriptive research is guided by hypotheses. Finally, causal research is concerned with determining cause and effect relationships.

All of the studies examined here can be considered both exploratory and descriptive. Gray's interactive model of diffusion is the only model that can be considered causal. The strategies and methodologies employed in the remaining studies are suitable for exploration and description.

Research Strategies

Except for Gray's study and to some extent Eveland and Rogers's study, two research strategies were used: the field survey and the field study. A field survey attempts to be representative of some known universe and emphasizes the generation of summary statistics [14, p. 63]. A field study is an in-depth review of a few typical situations and is less concerned with the generation of summary statistics than in evaluating the interrelationships among a number of factors [14, p. 62].

Walker's study can be considered a field survey. Five of the studies begin with a relatively extensive field survey and move from there to a

TABLE B-7

DESIGNS OF RESEARCH IN 12 EMPIRICAL STUDIES OF ORGANIZATIONAL INNOVATION

Author	Objective	Type of Research	Use of Time DV	Use of Time IV	Strategy	Methodology
Bingham	Explanation	Exploratory, descriptive, and causal	C	C	Field survey Field study	Correlation regression qualitative
Daft and Becker	Description, explanation, control	Exploratory, descriptive, and causal	L	C	Field study	Correlation regression
Downs	Explanation	Exploratory, descriptive, and causal	C	C	Field survey Field study	Correlation regression
Eveland and Rogers	Description	Exploratory	L	C	Field survey Field study Model building and testing	Multiple classification qualitative
Feller et al. [27]	Description, explanation, prescription	Exploratory, descriptive	L	C	Field survey Field study	Correlation qualitative

TABLE B-7

(Continued)

Author	Objective	Type of Research	Use of Time		Strategy	Methodology
			DV	IV		
Feller et al. [28]	Description, explanation	Exploratory, descriptive	L	C	Field survey Field study	Correlation qualitative
Gray	Description, explanation	Descriptive, causal	L	L	Model building and testing	Logistic functions correlation regression
Hage and Aiken	Explanation	Descriptive	C	C	Field study	Correlation
Perry and Kraemer	Description, explanation	Exploratory, descriptive	L	C	Field survey Field study	Regression discriminant analysis
Walker	Description explanation	Exploratory, descriptive	L	L	Field survey	Correlation factor analysis
Yin [27]	Description, explanation, prescription	Exploratory, descriptive	L	L	Case survey	Correlations regression
Yin [91]	Describe, explain	Exploratory, descriptive	L	L	Case survey	Simple associations

narrower, more in-depth field study. Eveland and Rogers begin with a field survey and subsequently moves to a field study on the basis of the survey's results. On the basis of the field study, they then designed and tested relevant models. Gray's study cannot be classified as either a field survey or a field study. Gray engages primarily in model building and testing, using data to test her deductively generated model. Also, Yin's study cannot be considered a field survey or field study, since he relies on aggregations of case studies for his data.

Type of Data Analysis

The methodology most used in the 12 research examples is correlation analysis, despite criticism of the usefulness of this method for dealing with the complexities of the innovation process.

There have been efforts to try alternatives to corrrelational analysis. Perry and Kraemer attempted a more sophisticated quantitative analysis, using discriminant analysis and linear representations of "adoptability," generated from matrices of "innovation decisions."

Others have become convinced that it is too early in the development of research on innovation to use more sophisticated quantitative techniques. Eveland and Rogers approached their study with the belief that the study of innovation in organizations needs to go back to the stage of variable identification, rather than determining the relationships between independent variables and organizational innovativeness.

Eveland and Rogers attempt to use a combination of etic and emic approaches to develop their model of the innovation process in organizations.

> The etic approach consists of using the researcher's socio-cultural system to define and categorize behavior in the respondent's socio-cultural system. This research approach imposes external order on the behavior being investigated. In contrast, the emic approach consists of using the respondents' socio-cultural system to define and categorize the respondents' behavior [24, p. 13].

Similarly, Yin [91] says he was disenchanted with his earlier work [92] because thoughts had been led by the data, instead of the reverse. In his later study, he gives more emphasis to in-depth review of particular cases with a minimum of quantification.

Treatment of Time

It is a strength of diffusion research that the time variable is built into the concept of innovation. Unfortunately, many researchers have neglected to build time into the measurement of innovation in a manner allowing time-varying comparisons of innovation adoption.

Both Gray and Walker's studies look at both the dependent variable and independent variables longitudinally. Walker divides his data into three historical periods and measures the appropriate socioeconomic and political data for each of them. In Gray's model, temporal change is an essential variable.

That part of the Eveland and Rogers study dealing with development of the model of specification uses a longitudinal approach to both the dependent and independent variables. Also, Yin uses a longitudinal approach in exploring the concepts of incorporation and routinization.

Three of the studies use only cross sectional data for both dependent and independent variables, making it impossible to discern possible modifications of the innovations or changes in the attitudes of the innovators or to spot changes in rates of adoption.

Daft and Becker's study uses two separate time periods for aggregation of an innovativeness score. This, at least, permits some comparison. To the extent possible, they also used independent variables appropriate to the two periods.

A drawback of the two Feller studies and the Perry and Kraemer study is that although they show the dependent variables longitudinally, as diffusion curves, many of the independent variables are measured cross-sectionally and at the end of the diffusion periods. This approach allows the proposition that the adoption of innovation caused attitudes and attributes of organizational personnel and their organizations, but not the reverse.

Recent Research on Innovation in State Public Utility Commissions

None of the 12 studies reviewed above dealt with agencies regulating public utilities. Nor, to our knowledge, had any other empirical study until 1980. Only 2 of the 12 studies looked at innovation in any regulatory agency: Feller et al. used innovations in air pollution control.

Kathleen Flaherty [29] in Innovation in State Public Utility Commissions sought to fill the gap in the existing literature by investigating the commissions' responses to innovations. She uses two models of innovation in the commissions. The first is a typical correlates-of-adoption or determinants approach. The second is a Downs and Mohr "adoptability" approach. As the only empirical investigation directly applicable to the problem of the feasibility of promoting new incentive mechanisms for use by the commissions, Flaherty's study deserves the closest scrutiny. The study is an exploratory one and undertakes to cover a great deal of previously uncharted territory. Thus, it is not surprising that careful analysis reveals flaws in Flaherty's work.

The set of organizations studied was public utility commissions in the American states. From the entire population of states, the response varied between 27 and 44 commissions. For each commission, the commission chairman and the staff member "most responsible for the procedures and principles used by the staff in electric utility matters" [129, p. 134] were interviewed. Variables at the level of the commissions environment and the individual were used to explain variance in adoption as well as variables at the level of the unit of analysis.

Four innovations were used in the study. We will list the innovations exactly as defined by Flaherty:

COMPUTERIZED ANALYSIS is defined as use of a computer (but not a programmable calculator) to process calculations such as summations, regressions, projections, simulations, and financial models for preparation of staff testimony or cross examination in an electric utility rate case. Excluded from this definition are use of computers for personnel functions such as storage of employee records, and calculation of the commission payroll. Also excluded are other uses for operational rather than technical/analytical aspects of utility commission functions, and use of computers as storage/retrieval devices in the absence of calculations on the data as described above.

PRODUCTIVITY ANALYSIS is defined as the use of a ratio of the output of an electric utility (for example, kilowatt-hour sales) to one or more of the inputs (for example, labor, capital, fuel) for purposes of measuring performance of that utility. Productivity analysis may fall on a continuum from "single factor" comparisons where output is examined in relation to a single input, to "total factor" comparisons where ratios involve all the factors of electric utility production. A given utility's ratios may be compared with its previous performance on these factors or with the ratios of "similar" utilities.

TIME-OF-DAY-PRICING involves establishing rates which reflect the cost of supplying electricity at different times, as these costs vary by the hour of the day, the day of the week, and perhaps the seasons of the year. The rates are made higher in peak periods, when loads are greater, and lower in off-peak periods when loads are reduced and the costs of providing capacity to meet higher demand periods is charged against those who are responsible for the heightened demand.

FUTURE OR PROJECTED TEST PERIOD implies projection of part of a year (for example, six months in a partially projected test year) or a full year (fully projected test year) to estimate an electric utility's future return on operations and the adjustments necessary to allow the utility to earn a reasonable rate of return on the allowable rate base. (Flaherety [29])

Productivity analysis and rates for time-of-day pricing may be classified as output innovations. Computerized analysis is a resource innovation. Allowing the use of a future test period to calculate a utility's future revenues is a change in the information resources of the commission that is incorporated into the operational level process of rate making.

The source of the idea of using a future test period may often be the utility. In 1978, it seems likely that the primary source for diffusion of rates reflecting marginal-cost considerations such as seasonal, interruptible, and time-of-day rates, has been the U.S. Department of Energy. The Public Utilities Regulatory Policies Act, passed in October of 1978, required the commissions to consider adopting time-of-day rates as a means of conserving energy.

The Dependent Variables

Flaherty attempted to use three measures of adoption to operationalize the dependent variable. The first was "adoption" versus "nonadoption." Flaherty apparently had sufficient data to make finer distinctions on the status of commission consideration of the innovations but decided not to use it. Table B-8 shows how responses on the status of the innovations were recorded. Several points may be made about Flaherty's categorization. First, she has included "studying" in the category of nonadoption. Since three of the four innovations had apparently diffused to only about half the commissions at the time of the study, she may have lost valuable information on commission awareness of the innovations by lumping the commissions that were studying the innovations with the ones that had never considered them. Also, it may be that the antecedents of awareness of an innovation are different from those predicting adoption or outright rejection.

In the case of time-of-day pricing, it may be argued that by including adoption and experimentation in the definition of "adopted," Flaherty has exaggerated the actual use of the innovation and is confounding two processes. Time-of-day pricing may be characterized as an authority innovation [24]. It emanates from the federal government with the authority of public law. One might hypothesize that such an innovation will be adopted both swiftly and superficially but not be successfully incorporated (to use Yin's term) nearly as fast. It may well be that lumping experimentation with adoption is not a strong commitment, at present, by the commissions to use time-of-day pricing.

Time for adoption was the second operationalization Flaherty used. It was defined as the "number of months elapsing from the order date of a

TABLE B-8

FLAHERTY CATEGORIZATION OF
ADOPTION/NONADOPTION RESPONSES [29]

	Adoption	NonAdoption
Computerized analysis	- Adopted - Proposal submitted	- Never considered - Studying - Explicit rejection
Time-of-day pricing	- Adopted - Experimental use	- Never considered - Studying - Explicit rejection
Productivity analysis	- Adopted	- Prohibited - Never considered - Studying - Explicit rejection
Future test periods	- Partially projeted - Fully projected with previous partial - Fully projected	- Prohibited - Never considered - Studying - Explicit rejection

commission's first use of an innovation and the order date of the first known use by a state commission" [29, p. 201]. For three of the innovations the date of the first adoption was unknown, and the date of the beginning of the Flaherty study (January 1973) was used.

Extent of adoption was defined as the "number of rate cases in which one of the innovations was applied divided by the number of rate cases decided in that time" [29, p. 205]. This definition poses difficulties in data gathering. It can be argued that better measurements of extent of adoption could be derived by looking at the impacts of the innovation on utilities and their customers.

The Independent Variables

Although she did not distinguish among types of innovations, Flaherty did attempt to follow the prescription of Downs and Mohr and gave consideration to several secondary dimensions of the innovations. She asked the commissioners and staff whether they found each innovation a major expense, an improvement in regulation, a major change in electric utility regulation, and/or a "political risk." She also asked whether they thought information was readily available on how to apply the innovation, whether it was merely a variation on analysis already being done by the commission, whether there was ready acceptance of the innovation within the commission, or whether the innovation represented a "radical change" in the way the commission was already doing things.

She asked the commissioners and staff whether they thought the innovation would "solve pressing problems" for the commission; and she asked them to respond to the statement that "one reason for our initial decision to apply [innovation] in rate cases was to increase our prestige with other commissions" [29, p. 225].

The organizational variables Flaherty used were the following:

Number of commissioners--adjusted for proportion of time devoted to regulating elecric utilities. [29, p. 158]

Size of staff--adjusted for amount of effort devoted to electric utility regulation.[29, p. 164]

Budget--adjusted for amount of resources devoted to regulating electric utilities. [29, p. 169]

Person resources--developed for three years (1973, 1975, 1977) by multiplying sum of adjusted commissioner plus staff figure by 2000 to get person-hours and dividing by electricity work volume. See below for definition of work voume. [29, p. 174]

Dollar resources--Adjusted budget was divided by work volume. [29, p. 174];

Hierachy of decision making within the commission--Interview question asked who in the commission made the final decision on the innovation. [29, p. 175]

Rate of turnover of upper management--One point was given for each chairman between 1973 and 1977. [29, p. 162]

Structure of commission (staff structure)--organized by utility, function, or a combination of utility and function. [29, p. 165]

Age of commission--Reconsitution of commissions was used rather than age. "Reconstitution" was defined as changes in method of selection or number of commissioners since 1970. [29, p. 157];

Work volume--Sum of volume of complaints and number of formal proceedings closed regarding electric utilities during a given year. [29, p. 172]

Flaherty did not employ the concept of slack resources. She used structure of commission but did not use centralization, a venerable variable in the study of organizational innovation [34].

Variables related to characteristics and behavior of individuals in the commissions were the following:

Training of staff to use innovation--At the time we first considered using [innovation] our staff had the training to apply it." [29, p. 221]

Professional activities (commissioners and staff)--An index of professional activities was developed based on Hage and Aiken [34]. Separate scores were calculated for chairmen and staff heads. A composite score was also prepared. [29, p. 181]

Search activities (commissioners and staff)--An index was constructed from seven questions judged to indicate the commission's level of search activity. [29, p. 179]

Education level of staff--An index was constructed using percentages and staff members with education beyond the baccalaureate degree. [29, p. 166];

Range of staff expertise--Commissions were scored according to whether particular professions were represented. [29, p. 167]

Notably lacking from this group of variables is any reference to key advocates or change agents.

Flaherty used some standard socioeconomic measures as environmental variables: state per capita income, median education, urbanization, industrialization, and expenditure levels.

She also used several variables related to the role of interested parties in consideration of the innovation:

Number of active interest groups--Interview item: "Please list the groups or individuals in your state who take an active interest in your commission's decisions on electric utility rate matters." [29, p. 187]

Frequency of contact with interest groups. [29, p. 190]

Priorities of interest groups. [29, p. 192]

Interest group activity--Respondents were asked to name the types of groups actively trying to convince the commission either to use or not to use a technique at the time it was first considered [29, p. 193]

Availability of university expertise--"Availability of university expertise was important in our decision on [innovation]."

One factor noticeably missing from the list of environmental variables is the impact of environmental turbulence as expressed in specific demand for commission action. Furthermore, there is no discussion of the role of the state legislature, the utilities, the federal government, or other commissions.

The final two variables in Flaherty's list may be categorized as aspects of the commission's general relationship to its environment:

Method of selection of commissioners--appointed or elected

Bureaucratic autonomy--staff members were asked whether or not "approval [was] required from any agency in your state government in order to use the innovation." [29, p. 176]

These are interesting variables that may have some potential for shedding light on the innovative behavior of commissions, but Flaherty does not adequately tie them to the particular commission behavior being explored.

NRRI Studies

The NRRI has not itself pursued a general study of innovation in state public utility commissions. However, much of its work has to do with the adoption of particular innovations, often by particular commissions.

The NRRI selected five states for in-depth case studies of innovations in 1978 and early 1979. The states and their innovations were as follows:

Missouri: restructuring of traditional block rates through a normal rate case rather than generic hearings
Oregon: a set of energy conservation bills
Arizona: a gas pipeline safety program
Wisconsin: rates based on time-differentiated marginal costs
Arkansas: time-of-day pricing

Other NRRI studies have also looked at innovations, though not necessarily labeling them as such. Studies of commission-ordered management audits, power plant productivity, and trends in commission regulation have explored aspects of innovation and identified factors affecting the innovation process. A review of the five NRRI case studies and the three general studies reveals a number of variables to consider in assessing the state of our knowledge of commission innovation. Many of these variables turn out to have been commonly used in the literature on organizational innovation.

References

1. Alchian, A.A., and Kessel, R.A. "Competition, Monopoly, and the Pursuit of Pecuniary Gain." In Aspects of Labor Economics, Princeton, N.J.: National Bureau of Economic Research, 1962.
2. Averch, H., and Johnson, L.L. "Behavior of the Firm under Regulatory Constraint." American Economic Review (December 1962), 1052-69.
3. Backoff, Robert W. "Innovation & Administrative Reform," Ph.D. dissertation, Indiana University, 1972.
4. Bailey, E.E. Economic Theory of Regulatory Constraint. Lexington, Mass.: Lexington Books, 1973.
5. Barnett, H.G. The Basis of Culture Change. New York: McGraw Hill, 1953.
6. Baron, D.P. "Price Regulations, Quality and Asymmetric Information" Working paper, Northwestern University, 1979.
7. Bazant, Stanley, Jr. "The New Mexico Regulatory Response to Inflation" (Mimeographed.) pp. 43-55.
8. Bingham, Richard D. "The Adoption of Innovations by Local Governments" Milwaukee: Marquette University, Office of Urban Research, NSF Grant No. DA-44312, 1975.
9. Bobbitt, H. Randolph, Jr., et al. Organizational Behavior, Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1974.
10. Boyes, William J. "An Empirical Examination of the Averch-Johnson Effect" Economic Inquiry (March 1976): 25-35.
11. Brown, L.A. "Innovation Diffusion: A New Perspective." Columbus, Ohio: The Ohio State University, Department of Geography, Discussion paper no. 60 in Studies in the Diffusion of Innovation, NSF Grant SOC 73-09058, 1979.
12. Brown, L.A., and Philliber, S.G. "The Diffusion of a Population-Related Innovation: The Planned Parenthood Affiliate." Social Science Quarterly 58, 1977, pp. 215-228.
13. Chen, A.H., and Kim, E. Han "Theories of Corporate Debt Policy: A Synthesis." Journal of Finance,
14. Churchill, Gilbert A., Jr. Marketing Research. Hinsdale, Ill.: The Dryden Press, 1979.
15. Coleman, James S.; Katz, Elihu; and Menzel, Herbert Medical Innovation: A Diffusion Study, Indianapolis, Ind.: The Bobbs-Merrill Co., 1966.

16. Courville, Leon. "Regulation and Efficiency in the Electric Utility Industry." Bell Journal of Economics (Spring 1974)); 53-74.

17. Crew, Michael A., and Kleindorfer, Paul R. "Marginal Discretion and Public Utility Regulation." The Southern Economic Journal (1979); 696-709.

18. ________. Public Utility Economics. New York: St. Martin's Press, 1979.

19. Daft, Richard L., and Becker, Delwyn W, Innovations in Organizations New York: Elsevier, 1978.

20. Dalton, G.W. The Distribution of Authority in Formal Organization. Boston: Harvard University, Division of Research, 1968.

21. Downs, Anthony. Inside Bureaucracy. Boston: Little, Brown and Company, 1967.

22. Downs, George W., Jr. Bureaucracy, Innovation, and Public Policy. Lexington, Mass.: Lexington Books, 1976.

23. Downs, George W. Jr., and Mohr, Lawrence B. "Conceptual Issues in the Study of Innovation." Administrative Science Quarterly 21 (December 1976) 700-714.

24. Eveland, J.D.; Rogers, Everett M.; and Kelepper, Constance; The Innovation Process in Public Organizations. Ann Arbor, Mich.: University of Michigan, Department of Journalism, NSF Grant RDA 75-17952, 1977.

25. Fama, E.F. "Agency Problems and the Theory of the Firm." Journal of Political Economy (April 1980): 288-307.

26. Feibleman, J., and Friend, J.W. "The Structure and Function of Organization." Philosophical Review 54)1945) 19-44.

27. Feller, Irwin et al. Diffusion of Technology in State Mission-Oriented Agencies. 1974 University Park, Pa: Pennsylvania State University, Institute for Research on Human Resources, Center for the Study of Science Policy.

28. Feller, Irwin; Menzel, Donald C.; and Kozok, Lee Ann. Diffusion of Innovations in Municipal Govts. University Park, Pa.: The Pennsylvania State University Institute for Research on Human Resources, Center for the Study of Science Policy, 1976.

29. Flaherty, Kathleen R. Innovation in State Public Utility Commissions: An Exploratory Study of Techniques in Energy Regulation. Washington: National Bureau of Standards, 1980.

30. Fox, John. "An Empirical Test of the Averch-Johnson Hypothesis in the Electric Utility Industry." Paper no. 538, Krannent Graduate School of Industrial Administration, Purdue University, Nov. 1975.

31. Gray, Virginia. "Innovation in the States: A Diffusion Study." American Political Science Review. (December 1973); 1174-85.

32. Green and Laffont. Incentives in Public Decision-Making. New York: North-Holland Publishing Company, 1979.

33. Grossman, Joel B. "The Supreme Court and Social Change." The American Behavioral Scientist 13 (1970); 535-51.

34. Hage, Jerald, and Aiken, Michael, "Program Change and Organizational Properties: A Comparative Analysis." American Journal of Sociology (March 1967): 503-19.

35. Harris, M., and Raviv, A. "Optimal Incentive Contracts With Imperfect Information." Journal of Economic Theory (April 1979); 231-59.

36. Havelock, Ronald G. Planning for Innovation through Dissemination and Utilization of Knowledge. Ann Arbor, Mich.: University of Mighigan, Center for Research on Utilization of Scientific Knowledge (CRUSK), 1969.

37. Hirshleifer, J. "Wh re Are We in the Theory of Information"? American Economic Review (May 1973): 31-39.

38. Hirshleifer, J., and Riley John G., "The Analytics of Uncertainty and Information--An Expository Survey." Journal of Economic Literature (December 1979): 1375-1421.

39. Hogg, R., and Craig, A. Introduction to Mathematical Statistics, 3d ed. New York: MacMillan Publishing Co., 1970.

40. House, Ernest R. The Politics of Educational Innovation. Berkeley, Calif.: McCutchan Publishing Corp., 1974.

41. Illinois Commerce Commission. Review of Existing Incentive Mechanisms. Prepared for the U.S. Dept. of Energy, Washington, D.C. by the Policy Analysis and Research Division of the Illinois Commerce Commission, Springfield, Illinois, 1980.

42. Jensen, M.C., and Meckling, W.H. "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure." Journal of Financial Economics (October 1976): 305-60.

43. Kaufman, Alvin and Profozich, Russell J. The New Mexico Cost of Service Index: An Effort in Regulatory Innovation. Columbus, Ohio: The National Regulatory Research Institute, 1979.

44. Klevorick, A.K. "The 'Optimal' Fair Rate of Return." Bell Journal of Economics and Management Science (Spring 1971): 122-53.

45. Knight, K. "A Descriptive Model of the Intra-Firm Innovation Process." Journal of Business 40 (October 1967): 478-96.

46. Leland, H. E. and Pyle. D.H. "Informational Asymmetrics, Financial Structure, and Financial Intermediation," The Journal of Finance (May 1977).

47. Leverich, Joseph. "An Outline of the History of Economic Regulation in the United States." June 1980.

48. Mailer, Norman. Of A Fire on the Moon. Boston, Mass.: Little, Brown and Company, 1969.

49. March, James G., and Simon, Herbert A. Organizations. New York: John Wiley and Sons, 1958.

50. Menzel, Donald. "Scientific and Technological Dimensions of Innovation in the American States." Paper presented at the 1975 annual meeting of the Midwest Political Science Association, May 1-3, 1975, Chicago, Ill.

51. Morton, Walter A. "Do Utilities Waste Capital?" Public Utilities Fortnightly, May 27, 1971, pp. 13-24.

52. Mosher, William E., and Crawford, Finla G. Public Utility Regulation. New York: Harper Brothers, 1933.

53. Myerson, R. B. "Incentive Compatibility and the Bargaining Problem." Econometrics, (January 1979): 61-74.

54. National Association of Regulatory Utility Commissioners. 1975 Annual Report on Utility and Carrier Regulation. Washington, D.C.: National Association of Regulatory Utility Commissioners, 1976.

55. National Regulatory Research Institute. Regulatory Characteristics Profile. Columbus, Ohio: The National Regulatory Research Institute, 1978.

56. Energy Management & Conservation in Arkansas: A Case Study Prepared by The National Regulatory Research Institute with the Cooperation of the Arkansas Public Service Commission. Columbus, Ohio: The National Regulatory Research Institute 8-1, 1978.

57. Commission-Ordered Management Audits of Gas & Electric Utilities, Col., Ohio: The National Regulatory Research Institute, 1979.

58. Controlling Gas Pipeline Loss in Arizona, Columbus, Ohio: The National Regulatory Research Institute, 1978.

59. The Missouri Rate Reform Initiative. Columbus, Ohio: The National Regulatory Research Institute, 1978.

60. Utility Regulation & the Legislative Process in Oregon, Columbus, Ohio: The National Regulatory Research Institute, 1979.

61. Talcott. Structure & Process in Modern Societies. New York: N.Y.: The Free Press of Glencoe, 1960.

62. Perry, James L., and Kraemer, Kenneth L. Diffusion and Adoption of Computer Application Software in Local Governments. Irvine, Calif: University of California. Public Policy Research Organization and Graduate School of Administration, NSF Grant No. PRA76-15548. January 1978.

63. ____________. "Innovation Attributes, Policy Intervention and the Diffusion of Computer Applications among Local Governments." Policy Decisions (1978): 179-205.

64. Peterson, H. Craig. "An Empirical Test of Regulatory Effects." Bell Journal of Economics (Spring 1974): 111-26.

65. Phillips, Charles F. Jr. The Economics of Regulation. Homewood, Ill.: Richard D. Irwin, 1969.

66. Pressman, Jeffrey L., and Wildavsky, Aaron B. Implementation. Berkeley, Calif.: University of California Press, 1973.

67. Rees, R., "A Reconsideration of the Expense Preference Theory of the Firm." Economica (August 1974): 295-307.

68. Robertson, Thomas S. Innovative Behavior and Communication. New York: Holt, Rinehart and Winston, Inc., 1971.

69. Rogers, Everett M. Diffusion of Innovations. New York: The Free Press, 1962.

70. Rogers, Everett, and Shoemaker, F. Floyd. Communication of Innovations, New York; Free Press, 1971.

71. Rogers, Everett M.; Eveland, Johd D. and Klepper, Constance A. The Innovation Process in Public Organizations: Some Elements of a Preliminary Model. Final Report, NSF Grant RDA 75-17952, Executive Summary, University of Michigan, Ann Arbor, 1977.

72. Ross, S. A. "The Economic Theory of Agency: The Principal's Problem," American Economic Review. (May 1973): 134-39.

73. ____________. "The Determinants of Financial Structure: The Incentive Signalling Approach," Bell Journal of Economics (Spring 1977).

74. __________. "On The Economic Theory of Agency and the Principle of Similarity," In pp. 215-37 Essays on Economic Behavior Under uncertainty, Amsterdam: North-Holland Publishing Co., 1974. 215-237.

75. Sheshinski, E. "Welfare Aspects of a Regulatory Constraint: Note," American Economic Review (March 1971): 175-8.

76. Spann, Robert M. "Rate of Return Regulation and Efficiency in Production: An Empirical Test of the Averch-Johnson Thesis." Bell Journal of Economics. (Spring 1974): 38-52.

77. Thompson, James D. Organizations in Action. New York; McGraw Hill, 1967.

78. Trebing, Harry M. "A Critique of Regulatory Accommodation to Change." In Regulation in Further Perspective. Cambridge, Mass.: Ballinger Publishing Company, 1974.

79. Trebing, Harry M. "Realism and Relevance in Public Utility Regulation." Journal of Economic Issues 8 (June 1964): 209-33.

80. Troxel, Emery. Economics of Public Utilities. New York, Holt, Rinehart, and Winston, Inc. 1947.

81. Twentieth Century Fund. Electric Power and Government Policy. Baltimore; Lord Baltimore Press, 1948.

82. U.S. Department of Housing and Urban Development. Factors Involved in the Transfer of Innovations. Prepared by Public Affairs Counseling for Office of Policy Development as Research, Dept. of HUD, HUD Contract H2282R, January 1976.

83. U.S. Federal Trade Commission. "Permissible Period of Time during which New Products May be Described as 'New.'" Advisory Opinion Digest no. 120, April 15, 1967.

84. Walker, Jack L. "The Diffusion of Innovations among the American States." American Political Science Review (September 1969): 880-899.

85. __________. "Comment: Problems in Research on the Diffusion of Policy Innovations." American Political Science Review, (December 1972.)

86. Weitzman, M. L. "Optimal Rewards for Economic Regulation." American Economic Review (September 1978): 683-91.

87. Wilcox, Clair and Shepherd, William G. Public Policies toward Business, 4th ed., Homewood, Ill.: Richard D. Irwin, 1975.

88. Williamson, D. E. "Managerial Discretion and Business Behavior." American Economic Review (December 1963): 1032-57.

89. ________________. Economics of Discretionary Behavior: Managerial Objectives in a Theory of the Firm. Englewood Cliffs, N.J.: Prentice-Hall, 1964.

90. Wilson, James Q. "Innovation in Organization: Notes toward a Theory." pp. 193-218. In Approaches to Organizational Design. Edited by Pittsburgh; University of Pittsburgh Press, 1966.

91. Yin, Robert K. Changing Urban Bureaucracies. Lexington, Mass.: Lexington Books, 1979.

92. Yin, Robert K.; Heald, A. Karen; and Vogel Mary E., Tinkering With the System. Lexington, Mass.: Lexington Books, 1977.

93. Zaltman, Gerald; Durean, Robert, and Halbek, Jonny. Innovations and Organizations. New York: Wiley, 1973.