Information, Incentives and Commitment in Multiperiod Regulatory Mechanisms

David P. Baron

Columbia Institute for Tele-Information
Graduate School of Business
809 Uris Hall
Columbia University
New York, New York 10027
(212) 854-4222

Draft
Please Do Not Reference
Without Author's Permission

Information, Incentives, and Commitment in Multiperiod Regulatory Mechanisms

David P. Baron†
Stanford University

I. Introduction

In response to pressures created by technological advances and service innovation, the telecommunications industry has undergone dramatic change during the past decade. In addition, the presence of regulated and unregulated services has complicated the already difficult task of identifying the costs relevant for pricing in the context of traditional cost-of-service regulation. These changes have strained existing regulatory policies and exacerbated the incentive problems inherent in regulation. This has generated proposals for regulatory reform, innovation, and deregulation intended to improve incentives for cost reduction, investment, and efficient pricing. The structuring of these incentives, however, is complicated by the incomplete information the regulator has about the costs of and opportunities available to the regulated firm and by the inherent limitations on the monitoring of performance.

During the period of technological and regulatory change in the telecommunications industry a theory of incentive regulation has developed that addresses the implications of incomplete information and imperfect observability for regulatory policy. This paper focuses on how incomplete information about costs, and about their evolution over time, can be incorporated into regulatory policy to provide incentives for efficient pricing and investment. This focus is not an assertion that information problems are the most important ones affecting the regulation of the telecommunications industry. Instead, this focus is intended to reflect the presence of an informational dimension in a number of important policy issues. Noll and Owen (1987) cite this as a fundamental problem confronting the FCC prior to the AT&T divestiture.

> The FCC could not determine AT&T's costs, nor could it settle on a sensible cost-based method for pricing. One set of AT&T prices, the Telpak tariff, went through nearly two decades of hearings without a final determination of its lawfulness. It was apparent that even with a fully informed regulatory policy and the best will possible, the FCC could not cope successfully within available administrative procedures with AT&T's control of the information necessary to regulate prices effectively. (page 10)

† This research has been supported by NSF Grant No. IST-8606157.

Informational asymmetries of this type are at least as important today and are at the center of the regulatory incentive problems addressed in this paper.

The information-based, incentive theory of regulation considered here developed as an alternative to prior research which largely took regulatory policies as given and focused on deducing and testing the implications of those policies. The theory is similar to its predecessor in that both focus on normative objectives, as in the theory of efficient second-best pricing, and seek to develop efficient regulatory policies and institutions given the structural features of the regulatory setting. The theory of incentive regulation, however, differs from its predecessor in a number of ways.[1] First, regulatory policies are viewed as endogenous and subject to choice by a regulator or a legislature. Second, the theory addresses the implications of incomplete information and unobservable actions for the strategic behavior of both the regulator and the regulated firm. Third, the focus is less on the relationship between cost structure and pricing policy, which has been the focus of much of the preceding literature, and more on the consequences of strategic behavior for pricing policies. Fourth, the theory specifically addresses how information obtained from monitors of performance should be used in regulatory policies. Although much remains to be accomplished before this incentive theory is adequate to address practical problems of regulation in their full complexity, it provides a number of potentially useful principles for the formulation of regulatory policy.

The incentive theory of regulation has generally been developed in the context of monopoly regulation or of supply by a publicly-owned firm and thus is most applicable to state regulation of local telephone companies.[2] Although this theory may not be directly applicable to those segments of the telecommunications industry that are, or soon will be, sufficiently competitive that regulation will not be needed, the principles identified in the theory may be applicable to the transition period from regulation to competition. The theory presented here thus is viewed as pertaining to the transition to competition and to the regulation of local telephone service which, in spite of the alternative of cellular and cable systems, is likely to continue to be regulated for the foreseeable future. The paper thus

[1] This theory has been surveyed by Baron (1987), Besanko and Sappington (1986), Caillaud, Guesnerie, Rey, and Tirole (1985), and Sappington and Stiglitz (1986).

[2] Recent work by Demski, Sappington, and Spiller (1986) and Caillaud (1986) has extended the theory to include a market alternative to regulation.

focuses on cost-based pricing policies and the associated incentive problems with particular emphasis on long-term policies that respond to information that is generated through performance.

The regulatory mechanisms considered here are in the spirit of recent policy proposals to delegate to the firm the authority to make certain decisions subject only to caps on profits or prices. The regulatory mechanisms prescribe a mechanism, or set of implementable policies, and delegate to the firm the choice of a particular policy based on the information it has about its costs. The policies cover extended periods of time, so they may include provisions that allow prices to adapt to information revealed by either exogenous events or performance in earlier periods. An important factor affecting the efficiency of such a mechanism is the regulator's ability to commit credibly to long-term regulatory policies. Commitment refers to the ability of the regulator to specify at the beginning of the regulatory horizon the policies for each future period. When the regulator is unable to commit credibly to long-term policies, it may act opportunistically either to exploit information that becomes available or to take advantage of sunk investments. The commitment issue is considered in the next section and the remainder of this section addresses the types of regulatory proposals to which this presentation is addressed. The principles applicable to those proposals are then summarized.

A number of the recently implemented or proposed policies are intended to provide at least some limited commitment to a regulatory framework. The New York Public Service Commission (NYPSC) concluded an agreement with New York Telephone Company to limit the increase in basic rates to $1 per month over two years with the company allowed to retain any cost reductions achieved. Presumably, when this agreement was negotiated, the firm had superior information relative to the NYPSC about what cost savings could be realized. The theory presented here is consistent with this policy in that it is based on an informational asymmetry between the regulator and the firm and uses the pricing policy as the principal instrument to regulate the firm.

The NYPSC (1987) has also concluded an agreement with the New York Telephone Company that precludes filing "a rate case for general rate relief that would be effective prior to January 1, 1991" in exchange for allowing the company to retain one half of the return on equity in excess of fourteen percent. The NYPSC, however, reserved the right

to adjust the agreement in light of "unforeseen circumstances."[3] Providing assurances that revisions will be made only in the event of unforeseen circumstances is at the heart of the commitment issue. From a practical perspective, a principal impediment to the application of the incentive theory considered here is in restricting revisions in policies to those cases in which events were truly unforseeable rather than just unspecified in the regulatory policy.

This agreement presumably will be renegotiated if it extends for its prescribed duration, and the performance during that period will undoubtedly be the basis for any subsequent agreement. If, for example, the firm earns a supercompetitive profit because, for example, it had prior knowledge of substantial realizable cost economies, the NYPSC might act to capture them in future periods. Recognizing this the firm may have less incentive to realize the cost savings in the current period. The inability, or political unwillingness, of a regulator to commit to a longer-term arrangement, however, can lead to inefficiencies. It is the tension between providing incentives to realize cost savings and the incentive of the regulator to confiscate those savings once they have been identified that underlies the issues considered in this paper.

A recent proposal by the National Telecommunications and Information Administration (1987) would place a cap on rates but would leave profits uncontrolled as a means of providing firms with incentives for cost reductions.[4] As Wendell Lind, AT&T's administrator of rates and tariffs, asserted, "With price-cap regulation instead of rate-of-return, prices will go down more because we'll have the incentive to really try to be more efficient. We would pass some of that (savings) on to customers."[5] Such a system is a means of delegating decisions to the regulated firm, which can be particularly important when information is incomplete. As in the case of the NYPSC agreement, however, expectations about the

[3] "The Commission reserves the authority to act on the level of NYT's rates in the event of unforeseen circumstances that, in its opinion, have such a substantial impact upon the range of earnings levels or equity costs envisioned by the moratorium as to render the Company's actual return unreasonable and unnecessary for the provision of safe and adequate service." NYPSC (1987, Appendix, p. 4)

[4] The issues here are similar to those in defense contracting where a proposal such as NTIA's would be labeled a firm-fixed-price contract. The Department of Defense employs such contracts for the procurement of standardized goods but finds it preferable to use cost-plus-incentive-fee contracts and cost-plus contracts when there is considerable cost uncertainty due either to cost randomness or to incomplete information. Cost-based regulatory policy is analogous to such contracts.

[5] *Wall Street Journal*, August 6, 1987.

renegotiation of the policy can impact its efficiency.

In its subsequent proposal of "price-cap" regulation the FCC (1987) requested comment on both the basis for the cap and the rules to govern its periodic revision. The objective of the FCC for the price-cap system was stated as: "Because it would permit the carrier to retain at least some of the profits arising from increased efficiency or creativity, this method of regulating would encourage greater efficiency and innovation than exists now, especially in less competitive markets. This approach would substantially decrease incentives to shift costs from more to less competitive services." The means of adjusting the price cap, according to the FCC, "should be related to the carrier's cost of providing interstate services, but generally beyond its ability to control." The principle of basing adjustments on exogenous factors addresses the obvious incentive problem, but as will be indicated in Section III.C adjustments initiated by exogenous factors can provide an opportunity for revisions based on endogenous information. This can cause inefficiencies when the firm has private information about its costs and the regulator is unable to commit credibly to how it will adjust the cap.

One case in which commitment has so far been maintained is in the price cap system for British Telecommunications (BT) which limits until 1989 price changes on a basket of services to the RPI index of inflation less X percent.[6] As Vickers and Yarrow (1988) correctly conclude, "The RPI - X by itself does not avoid the incentive problems of regulation, which arise essentially because of the monopoly of information of the regulated firm." (p. 71) That is, the index is based on realized consumer prices and not on the actual or anticipated cost changes for BT. Vickers and Yarrow (1985) address the issue of renegotiation of the agreement at its conclusion, "if X is determined by reference to BT's actual costs, we are back to rate-of-return regulation, with all its attendant problems. In particular, if the permissible tariffs depend on the level of costs, then the incentive to cost efficiency is blunted." (p. 43) This system also appears not to have relied on self-selection or delegation as a means of tailoring the policy to the possible costs that BT has.[7]

Furthermore, as Vickers and Yarrow (1988) state, this form of regulation may have

[6] The RPI is the equivalent to the CPI in the U.S., and X has been set at three. The price of domestic equipment rentals was also capped at RPI + 2 percent. The policy proposed by the FCC is closely related to this system.

[7] In part this is complicated by the multiple services supplied by BT.

many of the same types of problems as does cost-of-service regulation.

> RPI - X price control is closer to rate-of-return regulation than it might appear at first sight. If the DGT [Director General of Telecommunications of the Office of Telecommunications] intervened because BT's rate of return was becoming excessive, he would be acting in a manner very similar to how regulatory bodies in the United States often behave. If he does not intervene before the price control formula expires in 1989, it would also seem likely that the review at that time will be based largely upon BT's prospective rate of return, because BT has a near monopoly of information about the provision of most of the relevant telecommunications services in the UK." (p. 50)

The consequences of intervention and of renegotiation on the efficiency of regulatory policies is the subject of Sections III.C and IV.B.

The implications of the incentive theory considered here for this class of regulatory policies are summarized by the following points.

1. In a setting with incomplete information and identifiable events, the regulator prefers to commit to a mechanism, or collection, of regulatory policies with a policy for each possible cost level that the firm might have. The regulator then delegates to the firm the choice of the policy to be implemented. That policy will be chosen as a function of the firm's knowledge of its costs, and the firm will earn rents on it information. The resulting pricing policies are prospectively cost-based and depend on the (prior) information of the regulator and on the observed performance. Prices will be fully responsive to costs if a regularity condition is satisfied.
2. In a setting with incomplete information and identifiable events, commitment to long-term policies by the regulator improves *ex ante* efficiency but will generally be *ex post* inefficient given the information learned by the regulator through performance.
3. In a multiperiod regulatory setting in which the costs of the firm are known to be perfectly correlated over time and in which credible commitments are possible, the regulator prefers not to exploit fully the information obtained from the firm.
4. If costs are anticipated to change over time according to a known stochastic process and if credible commitments are possible, the optimal regulatory mechanism adjusts prices prospectively in every period in response to the changes in costs.
5. When the regulator cannot commit credibly to multiperiod policies, the set of policies the regulator is able to implement is restricted by its opportunism and the consequent opportunism of the firm. That opportunism may be reduced by regulatory institutions,

such as fairness (defined below), or by capital recovery rules that provide a means of deferred compensation.

6. With an inability to commit to multiperiod policies but with either fairness regulation or capital recovery rules that limit the opportunism of the firm, the regulator fully exploits the information obtained from the firm. This allows implementation of policies that are *ex post* efficient yet *ex ante* inefficient.
7. With fairness or capital recovery rules the regulator will choose a regulatory mechanism in which prices are only coarsely-responsive to costs. The purpose of such a policy is to limit the opportunism of the regulator.
8. With an inability to commit credibly to multiperiod policies but with either fairness or capital recovery rules that limit opportunism, the incentive to invest is diminished because the regulator will be expected to confiscate the rents it earns on its information.
9. With incomplete information about the costs of the firm it is generally impossible to hold the firm to a specific *ex post* rate of return. Instead, a range of returns should be tolerated.

The next section addresses the commitment issue, and Section III presents optimal, multiperiod regulatory mechanisms for the cases in which commitment is and is not possible. The impact on investment decisions is addressed in Section IV, and an example of regulation with the monitoring of performance is presented in Section V. Conclusions are offered in the final section.

II. Commitment, Opportunism, and Delegation

The incentive problems identified in Section I take on added complexity in a multiperiod setting. As Joskow and Schmalensee (1986) argue,

> the nature of the game played by the regulator and the firm changes dramatically when both make decisions over time. In principle, the commission can use repeated observations of firm performance to improve its information, and use that information to fine tune rewards and penalties. Knowing this, the firm has an incentive to try to fool the regulator, perhaps even raising costs and sacrificing profits today in order to make tomorrow's reward/penalty structure more favorable. Since public utility commissioners cannot sign contracts that prevent themselves or their successors - not to mention current and future legislatures - from changing policies, they cannot solve this problem by promising not to use what they learn. Such a head-in-the-sand policy would be plainly irresponsible even if it were credible.

> When incentives to deceive are taken into account, the problem of designing an optimal dynamic regulatory regime moves to a new level of complexity. (p. 24)

The relation between this commitment problem and politics has been addressed by Noll (1987) in his review of the politics of regulation.

> One key issue is whether political agents can credibly commit to durable, long-term arrangements with utilities which, even if optimal *ex ante*, could produce supra-competitive profits *ex post*. Such an outcome would leave the architects of a bidding or cost-revelation mechanism vulnerable to attack by political entrepreneurs seeking elective office. But even if this problem could be solved, interest group theory suggests that such mechanisms are extremely unlikely to be politically acceptable because they reduce to formula the politically relevant act of creating and distributing rents. Only upon the collapse of an economic regulatory process when too many interests are being cut in, combined with natural monopoly, would the political process be likely to consider such a mechanism. These circumstances have taken place in railroads, and may be under way in electricity and local telephone networks. (page 39)[8]

The regulatory policies optimal when information is incomplete and commitment is limited are analogous to private long-term contracts but differ in their incompleteness. In private contracting, parties can conclude an agreement that takes into account all available information and any future events that are anticipated. As long as the variables on which the contracts are based are jointly observable and are verifiable to a court, the parties can be confident that the contract will be implemented as anticipated. Government agencies, however, may have more difficulty in making credible long-term commitments than do private parties because, as Noll indicates, political forces can cause changes in policies and procedures.[9]

Commitment is a practical problem because one government cannot bind a future government to a specific policy. More fundamentally, citizens cannot bind themselves to act politically in a particular manner; for example, to how they will vote in the future. Consequently, regulatory commissions cannot be bound to long-term policies even if they want

[8] With respect to the empirical studies of the politics of regulation, Noll states, "Organized interests not only seem to succeed, but usually they do so at the cost of economic efficiency, at least as far as the data can tell us. Yet the evidence is still far from conclusive. (page 40)" Noll, however, warns not to blindly attribute causality to the gainers from regulatory policy. "It is impossible to imagine that regulation could be imposed without redistributing income. Hence, a look for winners in the process - and organizations that represent them - is virtually certain to succeed." (p. 40)

[9] The experience with incentive regulatory systems is reviewed by Schmidt (1984) and Joskow and Schmalensee (1986).

to bind themselves (and to bind future commissions). In addition, regulatory commissions have difficulty making credible commitments because all parties recognize that their membership can change as a consequence of a direct election or an appointment by an executive officer. Furthermore, regulatory commissions may choose to alter policies in response to political pressure or political opportunities. When credible commitments cannot be made, efficiency is reduced by the opportunism of the regulator and the regulated firm. As Baron and Besanko (1987c, p. 413) argue,

> This opportunism may be more characteristic of the policies of public agencies than of private parties because although courts will prohibit inefficient breach by private parties they generally will not proscribe revisions of policies by regulatory or administrative agencies. Instead courts tend to restrict their review to procedure, process, and consistency. Perhaps the greatest impediment to establishing commitment in governmental and regulatory settings arises from electoral competition. Presidential candidates and parties can pledge to preserve or to rescind laws or to force regulatory agencies to alter policies either through the appointment process, executive orders, or the authorization and appropriations process. Similarly, Congress can alter policies as well as initiate new ones. The political incentive to respond to an *ex post* opportunity, even though that opportunity results from an event anticipated under an *ex ante* efficient policy, seems unavoidable in many settings.

Even apparently conclusive legislation establishing regulatory policy may not suvive subsequent electoral politics. As Kahn (1987, p. 18) states "What of the possibility, finally, of total deregulation of local service? Nebraska has in fact passed a law – which is, incidently, the subject of strenuous Constitutional challenge – imposing a ceiling of ten percent per year on rate increases for local telephone service, subject to regulatory review if specified percentages of subscribers complain – and providing for total deregulation at the end of five years. I find it unsurprising that the bill's sponsor was just defeated in his bid for reelection, obtaining only 38 percent of the vote."

The politics of regulation ultimately spelled the end of the cost-of-service indexing policy for the Public Service Company of New Mexico. As hearing examiner Helman (1984, p. 152-3) stated, "The political atmosphere is such that the consumer and public view with suspicion any automatic rate relief to a utility even when there is no question of 'the appearance of the eye'; therefore, how much more so when suspicions are strengthened."[10]

[10] Helman refers to three incidents that, although their "overall revenue effects were small," attracted public, staff, and commission attention. One was an incentive compensa-

Even in the absence of party of political competition, a legislature is likely to prefer to leave open an option to review the policies of a regulator. As State Senator Robert C. Jubelirer of Pennsylvania stated regarding deregulation of intrastate telecommunications services, "It is not altogether clear whether or not deregulation could be achieved solely through state regulatory process. However, even if public utility commissioners have the authority, I do not think they want to take such a step without legislative sanction. And quite frankly, I do not believe most legislators would want them to do so." (1987, p. 4) Even when a legislature does establish policy, it may specifically choose to limit commitments to long-term policies. For example, Section 115 of the Public Utility Regulatory Policies Act limits the extent to which state regulatory commissions can commit to automatic adjustment clauses for electric utilities. The Act limits such clauses to four years and requires that they be reviewed at least every two years.[11]

There are a variety of obvious reasons why a regulatory policy may be altered when one political party succeeds another or when different constituencies prevail at different points in time. The concern here is with a more fundamental reason for an inability to make credible long-term commitments: the incentive to act opportunistically by taking advantage of *ex post* inefficiencies associated with *ex ante* efficient policies. That is, in a setting in which incomplete contracts are a fact of life, events may occur that provide an opportunity to revise policies in order to capture efficiency and/or distributive benefits. This opportunity may result from a desire to reduce "excess" profits, to capture quasi-rents associated with sunk investments, or to revise policies in light of information revealed about the capabilities and costs of the firm. In the setting considered here, the firm has private information about its costs and hence is able to earn rents on that information. Performance of the firm reveals information about those costs, and the regulator has an *ex post* incentive to take advantage of that information by revising its policy. Since the firm recognizes that the regulator will have this incentive and is unable to commit credibly not to take advantage of it, the firm will anticipate the regulator's behavior and will act strategically at the time the initial policy is formulated. This prevents the regulator from implementing policies that *ex ante* it would

tion system for management, and another was the purchase of a Lear Jet that management claimed would reduce transportation expenses. The third was the decision by management not to file a product liability suit to recover a $250,000 insurance deductible.

[11] See Schmidt (1984).

prefer to implement. The resulting regulatory policy is then inefficient.

One reason a regulator might behave opportunistically stems from the objectives that political office holders and commission members have to seek short-run benefits when they recognize that they may not be around to bear the long-run costs. To the extent that regulators, or their reputations, do not bear the long-run consequences of their actions, they may have an incentive to act opportunistically to their own advantage. Particularly when the opportunistic behavior appears on the surface to be promoting *ex post* efficiency, resisting the temptation may be difficult. The inability to give credible assurances that opportunistic behavior will be avoided then generates the inefficiency.[12]

The incentive of a regulator to act opportunistically to confiscate rents or quasi-rents in order to serve political or constituent interests is constrained both by the law and by features of the political system. For example, a firm with non-fungible assets is potentially subject to the risk of regulatory "confiscation" of the quasi-rents generated by those assets through prices or mandated service that are not compensatory. In *Smith v. Ames*, however, the Supreme Court concluded that the Constitution requires a fair return on assets employed in regulated service. What constitutes a fair return, however, is subject to a range of interpretation that allows considerable variation over time and across jurisdictions. That there remains considerable leeway for state regulators is evidenced by the change many states made during the 1970's from a "fair value" system of rate base measurement to an "original cost" method with the objective of holding down rates during a period of rapid inflation. The theory addressed here respects the fair return requirement. The opportunism based on information addressed here is not governed by an equivalent to a fair return requirement, however, so it is more difficult to limit.

The structure of political institutions can itself impede changes in regulatory policies. Legislative changes in regulatory mandates and procedures must command a majority in committee and on the floor of both chambers of the legislature and must be signed by the executive. Failure at any point in the process preserves the status quo and makes legislative modification of regulatory policies difficult.

The procedural due process requirements of administrative law also limit regulatory op-

[12] In practice, regulatory behavior responds to unanticipated events, but the theory addressed here cannot accommodate such events.

portunism by requiring that changes in policies be supported by the record. This, however, is a procedural test and as such does not constrain substantive changes in policy for which a basis can be established in the record. That is, the courts will generally review regulatory decisions for procedural correctness and not for substantive content such as whether the policy change is desirable.

Regulatory opportunism is also restricted by administrative rules that are difficult to change. For example, capital recovery rules limit opportunism by requiring that an asset's cost be recovered from consumer revenue. Unless the regulator determines that the asset is not "used or useful," its cost and return must be included in the revenue requirement. Consequently, regulatory opportunism associated with the confiscation of quasi-rents on long-term investments is restricted if the assets continue to be used. The analysis presented here is intended to be consistent with the requirements of administrative law and the protection of sunk assets. In particular, in the Appendix a capital recovery rule will be shown to be important in limiting the opportunism of the firm, which allows the regulator to implement an expanded class of policies when it is unable to commit credibly not to act opportunistically.

III. Optimal Regulatory Policies

A. The Model

The model presented here is intended to provide a basis for the development of intuition, underlying the design of regulatory mechanisms and for the presentation of results most of which are developed in the source papers referenced herein.[13] The model incorporates private information about costs with that information evolving over time based on past costs and investments. The cost C_t incurred in period t when the firm produces a quantity q_t is[14]

$$C_t = \theta_t q_t + k_t + B(x_t), \tag{1}$$

where θ_t is marginal cost, k_t is a fixed cost (e.g., overhead), x_t is the investment made in period t, and $B(x_t)$ is the cost of that investment with $B(0) = 0, B'(x_t)>0$, and $B''(x_t) \geq 0$.

[13] The static theory is presented in Baron and Myerson (1982), Guesnerie and Laffont (1984), Laffont and Tirole (1986), and Sappington (1982). The dynamic theory is presented in Baron and Besanko (1984) (1987) and Laffont and Tirole (1986).

[14] The model presented here pertains to a single good. Sappington (1983) presents a single-period, multi-product theory in a related model.

The cost θ_t is observed by the firm at the beginning of period t but is unobservable to the regulator. The cost thus is the private information of the firm and represents its "type." The private information could correspond to information about the firm's technology, about costs common to regulated and unregulated segments of the firm's business, to opportunity costs of fixed assets, or to characteristics of technological change.

The marginal cost θ_t evolves according to a stochastic process with transition function

$$\theta_t = \theta_t(\theta_{t-1}, x_{t-1}, \xi_t), \tag{2}$$

where $\theta_t \in \Theta_t$ and ξ_t is a random variable representing uncertain components of costs that are observable only to the firm. Investment x_{t-1} in period $t-1$ thus affects marginal costs in period t. The assumptions on the transition functions are

$$\frac{\partial \theta_t}{\partial x_{t-1}} < 0; \quad \frac{\partial \theta_t}{\partial \theta_{t-1}} > 0.$$

The marginal cost θ_t is increasing in θ_{t-1}, so higher costs in one period imply (stochastically) higher costs in the subsequent period. Investment reduces cost in the subsequent period.

The optimal regulatory mechanism in this setting depends importantly on whether the firm has private information prior to the regulator's choice of a mechanism or obtains private information after the mechanism has been chosen. The former case seems more descriptive of the current state of the telecommunications industry, so attention will be restricted to it. The firm thus knows θ_1 at the beginning of the regulatory horizon, and the regulator's prior information is represented by the distribution function $F_1(\theta_1) \equiv F_1(\theta_1 \mid \theta_0, x_0)$. A regularity assumption to be employed is that $(\theta_1 + \frac{F(\theta_1)}{f(\theta_1)})$ is a nondecreasing function of θ_1, where $f(\theta_1)$ is the density function. The random variable ξ_t in the transition function (2) induces a distribution function $F_t(\theta_t \mid \theta_{t-1}, x_{t-1})$ on the marginal cost.

The regulator would like to base its pricing policy on the marginal cost of the firm, but it does not know which cost the firm actually has. The next best alternative is to design a mechanism that includes a collection of pricing policies and delegate to the firm the choice of one of them. That choice will be based on its true cost, so the pricing policy can be made responsive to the costs of the firm through its selection of a policy. The task of the regulator is thus to choose a mechanism such that the firm's choice of a pricing policy serves the interests of the regulator. A two-part price structure will be employed where p_t denotes

the price and T_t is a fixed (or monthly) charge.[15] For each period a policy thus specifies the price $p_t(\hat{\theta}_t, \hat{\theta}_{t-1}, \ldots, \hat{\theta}_1)$ and a transfer or fixed charge $T_t(\hat{\theta}_t, \hat{\theta}_{t-1}, \ldots, \hat{\theta}_1)$, where the arguments denote selection variables by which the firm chooses a policy. A strategy of the firm in period t is thus a function $\hat{\theta}_t(\theta_t) : \Theta_t \to \Theta_t$.

In this setting, the firm has a natural incentive to choose a policy intended for a type of the firm with a higher marginal cost. For example, if the regulator were to attempt to implement a marginal cost pricing policy $p_t(\hat{\theta}_t) = \hat{\theta}_t$ with $T_t(\hat{\theta}_t) = k_t$ when $x_t = 0$, the firm has an incentive to choose a policy intended for a firm with a higher marginal cost. To illustrate this, consider a single-period model. The profit $\pi(\hat{\theta}_1; \theta_1)$ for a firm with marginal cost θ_1 is

$$\Pi(\hat{\theta}_1; \theta_1) = (\hat{\theta}_1 - \theta_1)Q(\hat{\theta}_1),$$

which has a maximum at $\hat{\theta}_1 > \theta_1$. The regulator thus prefers to design the mechanism to counter this incentive to overstate costs. The revelation principle implies that an optimal regulatory mechanism can be found in the class of policies such that the firm prefers to choose the policy designed for its marginal cost, i.e., $\hat{\theta}_t(\theta_t) = \theta_t$ for all θ_t.

The sequence of moves by the regulator and the firm depends on whether the regulator can credibly commit to a policy for the entire length of the horizon.[16] If the regulator can make such commitments, then it chooses a multiperiod mechanism M that specifies pricing policies for every period. The mechanism M thus is a collection

$$M = \{p_t(\hat{\theta}_t, \hat{\theta}_{t-1}, \ldots, \hat{\theta}_1), T_t(\hat{\theta}_t, \hat{\theta}_{t-1}, \ldots, \hat{\theta}_1), t = 1, \ldots, \tau\}, \tag{3}$$

where τ is the number of periods in the horizon.[17] Then, at the beginning of each period, the firm chooses from the policies for that period by selecting $\hat{\theta}_t$. Thus, the regulator moves first and chooses a mechanism M. At the beginning of period one the firm chooses a particular pricing policy $(p_1(\hat{\theta}_1), T_1(\hat{\theta}_1))$ by selecting $\hat{\theta}_1 = \hat{\theta}_1(\theta_1) \in \Theta_1$.[18] At the beginning of period two the firm observes θ_2 and chooses a pricing policy $(p_2(\hat{\theta}_2, \hat{\theta}_1), T_2(\hat{\theta}_2, \hat{\theta}_1))$ by

[15] The fixed charge is assumed not to affect demand.

[16] The regulatory setting involves no unanticipated events, so the regulator can in principle choose a policy at time zero and have it govern subsequent decisions for the entire horizon.

[17] The policy selected is the only variable observable to the regulator.

[18] Since θ_t is observed at the beginning of period t, prices are prospectively based on costs for the coming period.

selecting $\hat{\theta}_2 = \hat{\theta}_2(\theta_2) \in \Theta_2$. The subsequent periods are analogous. The equilibrium sought is a Bayesian Nash equilibrium.[19]

If, as discussed in Section II, the regulator is unable to make credible commitments to future policies, the regulator will act optimally in every period, conditional on the information available. The regulator thus will choose its policies for period t at the beginning of that period. The regulator still prefers to rely on self-selection by the firm and therefore will choose a menu $M_t = \{p_t(\hat{\theta}_t, \hat{\theta}_{t-1}, \ldots, \hat{\theta}_1), T_t(\hat{\theta}_t, \hat{\theta}_{t-1}, \ldots, \hat{\theta}_1)\}$ at the beginning of each period. This choice will be made optimally given whatever information the regulator has about θ_t, so the equilibrium sought is a subgame perfect, Bayesian Nash equilibrium. In the absence of commitment the regulator is unable to avoid exploiting whatever opportunities are available in period t. In particular, the regulator is unable to commit not to exploit any information that becomes available regarding marginal cost. The regulator thus cannot avoid acting opportunistically, and recognizing this the firm will act strategically by anticipating the behavior of the regulator. As shown in Section III.C, this opportunism results in inefficiency.

Commitment means than the regulator, if it finds it desirable to do so, can credibly pledge not to act opportunistically when it receives information relevant to the cost of the firm. If the regulator is capable of making such credible commitments, the firm need not take into account the future behavior of the regulator but can instead rely on the announced mechanism. The agreements reached between the NYPSC and the New York Telephone Company may be interpreted as attempts to establish credible commitments for the periods of the agreements. This is weaker than full commitment, however, because at the conclusion of the period covered by the agreement the regulator would presumably establish a new policy based on the information available at that time. The regulator's choice of the new policy thus will affect the firm's behavior which in turn affects the choice of the initial agreement. This is the source of the inefficiency addressed by Vickers and Yarrow in their analysis of the price cap system used to regulate British Telecom.

The profit π_t of the firm in period t is

$$\pi_t = p_t Q(p_t) + T_t - \theta_t Q(p_t) - k_t - B(x_t), \tag{4}$$

[19] The optimal regulatory mechanism for this case is characterized in Baron and Besanko (1984).

where $Q(p_t)$ is the demand function. The objective of the firm is to maximize the (expected) discounted sum of its profit over the τ-period horizon. The regulated firm is assumed to be privately-owned and to be assured a fair return in each period. The regulatory policy is thus chosen subject to the constraints

$$\pi_t \geq 0, \quad t = 1, \ldots, \tau. \tag{5}$$

The regulator is assumed to maximize the expectation of the discounted sum of consumer surplus where the expectation is taken with respect to the regulator's information about θ_t conditional on the history to that point.[20] Consumer surplus $S(Q(p_t))$ in period t is given by

$$S(Q(p_t)) = Y(Q(p_t)) - p_t Q(p_t) - T_t, \tag{6}$$

where Y is the aggregate willingness to pay of consumers.

B. Optimal Regulatory Mechanisms with Commitment

The basic tradeoff facing the regulator is between consumer surplus and the profit (or information rent) of the firm, since substituting (4) into (6) yields

$$S(Q(p_t)) = Y(Q(p_t)) - \theta_t Q(p_t) - k_t - B(x_t) - \pi_t. \tag{7}$$

If the regulator knew θ_t, it could choose a pricing policy that held profit π_t to zero. With incomplete information, however, the firm earns profit, or more correctly rents, on its private information. In a one-period model those rents $\pi(\theta)$ are[21]

$$\pi_1(\theta_1) = \int_{\theta_1}^{\theta^+} Q(p_1(\theta_1^o))d\theta_1^o, \tag{8}$$

as shown in Baron and Myerson (1982). Consequently, the lower are the costs of the firm the higher are the rents it earns. The natural incentive of the firm is to select $(\hat{\theta}_1 > \theta_1)$ a

[20] The model can be extended directly to the case in which the regulator maximizes a weighted average of consumer and producer surplus. The optimal mechanism has the same qualitative properties as that considered here.

[21] An important issue is whether in a regulatory context a firm has a property right to information about its costs and demand. If so, the firm may be viewed as having a right to reveal its information only in exchange for adequate compensation. The information rents thus may be viewed as the return extracted by the firm for its information. In a regulatory setting the regulator has the authority to minimize those rents by making a take-it-or-leave-it offer, but the regulator has to trade-off rent reduction against the efficiency of the regulatory policy.

pricing policy intended for a higher cost firm in order to obtain higher profits. The rents represent the incentive payment to the firm with cost θ_1 required to offset the incentive of the firm to select a pricing policy intended for a firm with a higher cost.

The objective of the regulator is to maximize *ex ante* consumer surplus W which from (7) is given by[22]

$$W = \int [Y(Q(p_1(\theta_1))) - \theta_1 Q(p_1(\theta_1)) - k_1 - \pi_1(\theta_1)] f_1(\theta_1) d\theta_1. \tag{7a}$$

In this setting, the regulator is able to implement a marginal-cost-pricing policy. From (8), however, the higher is the price the lower are the information rents, so the regulator has an incentive to distort price above marginal cost to reduce the rents even though that reduces the surplus $[Y(q(p_1)) - \theta_1 Q(p_1) - k_1]$ in (7). The price $p_1(\theta_1)$ that optimally trades off rents and surplus is[23]

$$p_1(\theta_1) = \theta_1 + \frac{F_1(\theta_1)}{f_1(\theta_1)} = \theta_1 \Big(1 + \frac{F_1(\theta_1)}{\theta_1 f_1(\theta_1)}\Big), \tag{9}$$

where $\frac{F_1(\theta_1)}{f_1(\theta_1)}$ has the interpretation as the marginal information costs to the regulator (or rents to the firm). Note that as long as $\theta_1 + \frac{F(\theta_1)}{f(\theta_1)}$ is strictly increasing in θ_1, the price is "fully responsive" to costs.[24] The fixed charges $T_1(\theta_1)$ are then chosen to induce the firm to select $(\hat{\theta}_1(\theta_1) = \theta_1)$ the policy corresponding to its marginal costs. These fixed charges are given by equating (4) and (8) which yields

$$T_1(\theta_1) = \theta_1 Q(p_1(\theta_1)) + k - p_1(\theta_1) Q(p_1(\theta_1)) + \int_{\theta_1}^{\theta^+} Q(p_1(\theta_1^o)) d\theta_1^o. \tag{10}$$

The incentive for political opportunism is evident from the price in (9). The *ex ante* efficient price is greater than marginal cost, and hence there is an *ex post* incentive to reduce it in the next period to capture efficiency gains. For example, if $f_1(\theta_1)$ is uniform, then price $p_1(\theta_1)$ is twice the marginal cost. The difference between price and marginal cost increases with θ_1, so the incentive is greater for high costs than for low costs. It is this incentive that is at the heart of the commitment problem studied in Sections III.C and IV.B. The optimal

[22] The investment x_1 is assumed to be zero here.

[23] This is dervied by substituting (8) into (7*a*), integrating by parts, and maximizing pointwise with respect to $p_1(\theta_1)$.

[24] See Baron and Myerson (1982) or Baron (1987) for a demonstration of the optimality of this mechanism.

multiperiod regulatory mechanism with commitment is characterized next to identify the properties of the *ex ante* efficient mechanism and to provide a benchmark for the evaluation of mechanisms in the absence of commitment.

To indicate the nature of efficient regulation in a multiperiod model, consider the case in which the costs of the firm are characterized by technological change that reduces marginal cost over time. This is intended to be representative of the long-term decline in the real costs of telecommunications services. Marginal cost θ_t will initially be assumed to be given by the deterministic transition function

$$\theta_t = \gamma\theta_{t-1} = \gamma^{t-1}\theta_1, \tag{11}$$

where the parameter γ is less than or equal to one. Marginal costs are thus perfectly correlated and decrease over time according to a known function. In this case, the regulator need only choose a mechanism M of the form $M = \{p_1(\hat{\theta}_1), p_2(\hat{\theta}_1), \ldots, p_\tau(\hat{\theta}_1), T_1(\hat{\theta}_1), T_2(\hat{\theta}_1), \ldots, T_\tau(\hat{\theta}_1)\}$, since costs do not change over time. In the context of the price cap systems discussed in Section II, the cap in each period is represented by the pair $(p_t(\hat{\theta}_1), T_t(\hat{\theta}_1))$, which may be interpreted as caps on both the usage charge and the monthly charge. With these caps the firm, when delegated the choice of prices, will choose the usage charge and the monthly charge at their caps.[25]

The optimal prices, expressed here as a function of θ_1, can be derived from the char-

[25] This conclusion is based on the condition that prices are below the monopoly price for all θ. This condition seems reasonable, but it is possible from (9) that the regulator would prefer a price above the monopoly price as a means of economizing on the information rents.

acterization in Baron and Besanko (1984) as[26]

$$\begin{aligned} p_t(\theta_1) &= \theta_t + \gamma^{t-1}\frac{F(\theta_1)}{f(\theta_1)} \\ &= \gamma^{t-1}\theta_1\Big(1 + \frac{1}{\theta_1}\frac{F(\theta_1)}{f(\theta_1)}\Big) \\ &= \gamma^{t-1}p_1(\theta_1). \end{aligned} \tag{12}$$

The price cap thus decreases over time at the same rate at which marginal costs decrease, but in all periods the price is above marginal cost by the same percent.

For example, suppose θ_1 has a triangular distribution with $f_1(\theta_1) = 2\theta_1$ for $\theta_1 \in [0,1]$, so that high costs are more likely than low costs. Then the prices are given by[27]

$$p_t(\theta_1) = \gamma^{t-1}\frac{3}{2}\theta_1.$$

Price is thus fifty percent greater than marginal cost in every period but declines at the same rate as does marginal cost.

Viewed from time zero when the regulator designs the mechanism of regulatory policies, the time path of prices is deterministic once the firm has selected the particular policy based on its true θ_1. Prices decline over time at the same rate as marginal costs, but in every

[26] If the costs were independent across periods, the optimal price cap would be the price in (9) for the first period and in each subsequent period a cap equal to marginal cost. The optimal static price is implemented in the first period because the firm has the same informational advantage in the first period as in the single-period model. At the beginning of the regulatory horizon, however, the firm has no informational advantage relative to the regulator about the cost in any future period. The firm thus can earn no information rents associated with its operations in future periods, so the regulator need not tradeoff efficiency against rent control. The regulator thus implements a marginal cost price cap in every period after the first. Since the firm observes its cost at the beginning of each period and the regulator does not observe it, the regulator still must implement marginal cost pricing by employing in every period a self-selection mechanism. The firm earns rents in that period, but the regulator can costlessly eliminate those rents by charging in the previous period a "franchise fee" equal to the discounted expectation of the future rents. Thus, other than for the first period the rents do not affect the pricing policy. If the firm did not learn θ_1 until after the menu of regulatory policies has been announced, a marginal cost price cap could be implemented in the first period as well.

[27] The optimality of the mechanisms characterized here requires verification of global incentive compatibility conditions. See Baron (1987) and Guesnerie and Laffont (1984) for an analysis of these conditions. With commitment and perfectly correlated marginal costs the condition that $\theta + \frac{F(\theta)}{f(\theta)}$ be nondecreasing in θ is sufficient for global incentive compatibility. In other cases the specification of sufficient conditions is more complicated.

period price incorporates a mark-up above marginal cost equal to the marginal information costs resulting from the firm's private information. An optimal regulatory mechanism thus specifies price caps that decrease over time if $\gamma < 1$ and are constant over time if $\gamma = 1$. Since the transition function in (11) is known to the regulator, the regulator is able to specify the caps in advance. In this respect, this mechanism corresponds to the price cap system used for British Telecom. That system is not cost-based, however.

To indicate the significance of commitment for the design of price cap mechanism, consider the case in which $\gamma = 1$. The optimal caps are then to commit to the repetition in each period of the optimal price in (9) for a one-period, or static, model with the fixed charges given in (10). Even though the regulator observes the selection of a policy in the first period which completely reveals the marginal cost θ_1, the regulator prefers to commit not to use that information in subsequent periods. The regulator prefers to implement a policy in which there is no learning because that provides the optimal tradeoff between consumer surplus and the rents the firm earns on its private information. Since the firm has at the beginning of the horizon the same private information about its costs in each future period, it earns rents on that information in each period. The tradeoff between rent and consumer surplus is thus the same for each period.

Of course, after the selection of a regulatory policy in the first period, the regulator knows θ_1 and hence knows what costs will be in every subsequent period. The regulator thus could implement marginal cost pricing in every period other than the first. The regulator prefers not to exploit that information, however, because without credible commitment the firm would anticipate that its rents would be exploited and would act strategically in its initial selection of a policy. This would then reduce the efficiency of the regulatory mechanism.

In a multiperiod model, the rents $\Pi(\theta_1)$ earned by the firm as a consequence of its private information about θ_1 are analogous to (8)

$$\Pi(\theta_1) = \int_{\theta_1}^{\theta^+} \left[\sum_{t=1}^{\tau} \beta^{t-1}\gamma^{t-1} Q(p_t(\theta_1^o))\right] d\theta_1^o, \tag{8a}$$

where β is the discount factor. The firm thus earns more than its cost of capital, but it does so because of its private information about costs. Because of that private information it is impossible for the regulator to eliminate these "excess" profits because eliminating profits

for a firm with marginal cost θ would cause a firm with a higher cost not to recover its capital costs. This firm would then be unable to raise capital.

Next consider a stochastic transition function where γ is the realization of a random variable $\tilde{\gamma}$ that is uniformly distributed on the interval $[0,1]$. Costs are thus imperfectly correlated over time. Initially, consider a two-period horizon so that γ is realized once at the beginning of period two. The price cap for the first period is unaffected and is given in (9). Viewed from time zero when the regulator designs the mechanism of regulatory policies, the distribution $F_2(\theta_2 \mid \theta_1)$ is given by

$$F_2(\theta_2 \mid \theta_1) = \frac{\theta_2}{\theta_1} \quad \text{if } \theta_2 \in [0,\theta_1]. \tag{13}$$

At the beginning of period two the regulator does not know θ_2 and thus designs a price cap $p_2^*(\theta_2,\theta_1)$. That cap is[28]

$$\begin{aligned} p_2^*(\theta_2,\theta_1) &= \theta_2 - \frac{\frac{\partial F_2(\theta_2|\theta_1)}{\partial \theta_1}}{f_2(\theta_2 \mid \theta_1)} \frac{F_1(\theta_1)}{f_1(\theta_1)} \\ &= \theta_2\Big(1 + \frac{1}{\theta_1}\frac{F_1(\theta_1)}{f_1(\theta_1)}\Big). \end{aligned} \tag{14}$$

The second term on the right side of (14) is the marginal information cost resulting from the firm's private information about θ_1. It is important to note that this regulatory mechanism is "nested" in that in each period the firm makes a selection but that selection is conditioned on the selections in earlier periods.

It is important to note that the price caps are set prospectively but are based on costs in prior periods. That is, in (14) for example, the price cap for period two is proportional to the costs that the firm will have in period two. The margin above that marginal cost, however, depends on the costs in period one. That dependence is due to the information that period-one costs provide for period-two costs. That is, in the first line of (14) the term $\frac{\partial F_2(\theta_2|\theta_1)}{\partial \theta_1}$ represents the impact of θ_1 on the information about θ_2. For the case of perfect correlation, $\theta_2 = \theta$ and this derivative equals -1. The expression in (14) then reduces to (9). If θ_1 provides no information about period-two marginal costs so that θ_1 and θ_2 are statistically independent, then $\frac{\partial F_2(\theta_2|\theta_1)}{\partial \theta_1} = 0$ and $p_2^*(\theta_2,\theta_1) = \theta_2$. Thus, it is the information that costs in earlier periods provides for costs in the future periods that determines how

[28] This follows from Baron and Besanko (1984).

the price caps depend on past costs. In all cases, however, the cap for the next period is prospectively based on the costs anticipated for that period.

If θ_1 has a triangular density $f_1(\theta_1) = 2\theta_1$, this becomes

$$p_2^*(\theta_2, \theta_1) = \frac{3}{2}\theta_2. \tag{15}$$

Again, price is marked up above marginal cost by fifty percent of the marginal information costs. Although the markup is present in every period, the path of the price caps is stochastic when viewed from the time at which the regulator chooses the mechanism. In general, the price $p_t^*(\theta_t, \theta_{t-1}, \ldots, \theta_1)$ in period t is given by

$$p_t^*(\theta_t, \theta_{t-1}, \ldots, \theta_1) = \theta_t\Big(1 + \frac{1}{\theta_1}\frac{F_1(\theta_1)}{f(\theta_1)}\Big) = \frac{\theta_t}{\theta_1}p_1(\theta_1). \tag{16}$$

The cap on the fixed charges $T_t^*(\theta_t, \theta_{t-1}, \ldots, \theta_1)$ is then determined in the manner used to obtain (10.

The price patterns when costs evolve deterministically and stochastically can be compared by examining the case in which $\gamma = \frac{1}{2}$ and $f_1(\theta_1)$ is triangular. The expected prices, where the expectation is taken at the beginning of period one, are then equal.

$$\begin{aligned} Ep_t = Ep_t^* &= 1 \qquad \text{if } t = 1 \\ &= \frac{3}{4}\gamma^{t-1} \qquad \text{if } t>1. \end{aligned} \tag{17}$$

Although the expected time paths are the same, the path for the case of perfect correlation is deterministic once the firm selects the regulatory policy in the first period, whereas the path is stochastic in the case of imperfect correlation.

The mechanism in this case involves a price cap that is revised at the beginning of each period based on the costs that the firm realizes. Since the realized cost is exogenously determined, this is consistent with the FCC's statement that adjustments in the cap should be based on exogenous factors. Even though the price cap is established in each period and viewed from time zero the cap is a random variable, the formula governing the adjustment is specified in advance. Commitment means that formula cannot be adjusted. The FCC proposal for price-cap regulation would, however, allow adjustments in the cap. The theory presented here indicates how the price cap should be adjusted if commitment to the mechanism can be made credible, but the formula for the adjustment should not be subject to

change. If the rate at which costs are anticipated to decline is known with certainty, the rate at which the cap will decrease can be specified in advance. In the more realistic case in which the rate of change is not known in advance, the mechanism should specify the formula by which information will be used to adjust the cap.[29]

C. Optimal Regulation With No Commitment

If the regulator is unable to commit credibly to long-term policies, it will at the end of the first period base the mechanism M_2 for the second period on whatever information was revealed by the firm's selection from the mechanism M_1 in the first period. The policies that the regulator is able to implement in this case depend importantly on the extent of the opportunism that is possible in the regulatory relationship. In the perfect correlation case specified in (11), Laffont and Tirole (1986a) demonstrate that if the regulator can fully exploit the cost information it obtains in the first period, and thus would implement a marginal-cost price cap in period two if it learned θ_1 in the first period, the firm has an incentive in the first period to select a pricing policy designed for a lower marginal cost and then not to produce in the second period. The regulator thus cannot implement in period one pricing policies that are fully responsive to costs and must resort to "coarse" policies that specify the same price for many different costs. This result is analyzed in the Appendix.

Two forces serve to limit this opportunism. First, if the firm has substantial sunk investment costs that through regulatory rules are recoverable over several years provided they are usefsul and used, the incentive of the firm to act strategically is limited as demonstrated in the Appendix. This capital recovery rule coupled with substantial sunk costs allow the

[29] If the firm had an unobservable effort decision a_t to make where marginal cost was $c(\theta_t, a_t)$, the firm will choose its effort $a_1(\theta_1)$ to satisfy

$$-c_a(\theta_1, a_1(\theta_1))Q(p_1(\theta_1)) - V'(a_1(\theta_1)) = 0,$$

where $V'(a_1(\theta_1))$ is the marginal disutility of effort. This is the same effort level the regulator prefers, so with commitment the effort decision can be delegated to the firm. The same is true for the second period. See Baron (1987), Section IV.F, for an analysis of this case. This result would differ if the model were formulated with an *ex post* monitor z, so that the fixed charges could be based on both $\hat{\theta}$ and z. Laffont and Tirole (1986a) consider such a case and show that the regulator will base prices on the monitor as a means of reducing the information rents. Baron and Besanko (1987a)(1987b) consider a similar model with a risk-averse manager.

regulator to implement pricing policies that are continuously responsive to costs. Second, Baron and Besanko (1987c) consider a regulatory relationship, characterized by what they label as "fairness," under which the firm agrees not to quit the regulatory relationship as long as the regulator provides the firm with a fair return given the information it reveals through its selection of a pricing policy in the first period.[30] This limited form of commitment allows the regulator to implement pricing policies that are continuously responsive to costs.[31] In this section the regulatory relationship is assumed to be characterized either by fairness or by substantial sunk costs and a capital recovery rule.

The regulatory policies that would be implemented in the absence of commitment correspond to those that would be implemented at the conclusion of a regulatory mechanism such as that for British Telecom or that implemented in New York. Two aspects of this regulatory setting are of particular interest. First, what mechanism will be implemented at the conclusion of the duration of the first mechanism? Second, what is the impact of the choice of the second-period mechanism on the choice of the mechanism for the prior period? That is, since the firm will anticipate the regulator's choice of a second-period mechanism and will take that into account in making its selection from the mechanism in the first period, the regulators will find it optimal to anticipate the firm's strategic choice. This results in a reduction in *ex ante* efficiency.

To investigate these issues, suppose that in the first period the regulator were to implement a mechanism M_1 that is continuously responsive to costs. At the end of the first period the regulator would then know θ_1 from the policy selected by the firm in period one. Since the regulator cannot resist exploiting this information, it will base the price for the second period on the posterior distribution $F_2(\theta_2 \mid \theta_1)$. For the case of perfectly correlated costs, the posterior distribution places mass one on $\theta_2 = \gamma\theta_1$. The price in the second pe-

[30] Fairness is directed to the incentive the parties have to revise regulatory policies once information has been revealed either by self-selection or by performance. An inability to ignore this incentive prevents the implementation of efficient policies in prior periods. For an example, Baron and Besanko demonstrate that both the firm and the regulator prefer to abide by a regulatory relationship characterized by fairness than to participate in one in which commitment is not possible. Administrative rules and the courts, however, are required to make this agreement credible.

[31] Even though continuously responsive prices can be implemented, the regulator may not find them to be optimal but instead may prefer a coarse pricing policy in which different price caps are specified for different sets of possible costs.

riod, and in each subsequent period, is then equal to the marginal cost the regulator knows that the firm will have. The fixed charges T_2 then equal the fixed cost k_2 and the firm earns no rent after the first period. As demonstrated in Baron and Besanko (1987c), this results in a welfare loss compared to the mechansims characterized in the previous section because too large an incentive payment is required in the first period to implement those price caps.

For the case of imperfect correlation, the price $p_2^o(\theta_2, \theta_1)$ in period two is analogous to (9) but is based on the posterior distribution $F_2(\theta_2|\theta_1)$ or

$$p_2^o(\theta_2, \theta_1) = \theta_2 + \frac{F_2(\theta_2 \mid \theta_1)}{f_2(\theta_2 \mid \theta_1)}. \tag{18}$$

For the example with θ_2 uniformly distributed on $[0, \theta_1]$ the price is

$$p_2^o(\theta_2, \theta_1) = 2\theta_2 \quad \text{if } \theta_2 \in [0, \theta_1]. \tag{18a}$$

In this case, the period-two price is higher when the regulator exploits the information obtained during the first period. A higher price is not necessarily an indication of inefficiency, but the *ex ante* welfare is strictly lower in the absence of commitment because the regulator prefers to implement the price p_2^* in (15).[32] The *ex ante* welfare loss when long-term commitments cannot be made and the regulatory relationship is characterized by fairness is characterized in Baron and Besanko (1987c).

Because an *ex ante* welfare loss results when the regulator cannot credibly commit to long-term policies and thus cannot avoid acting opportunistically when it observes the policy selection in the first period, the regulator would be expected to seek means of restricting its own opportunism. One means of doing so is not to learn θ_1 in the first period. The regulator can accomplish this by choosing a menu M_1 that contains a single pricing policy; i.e., a single price cap rather than a cap as a function of $\hat{\theta}_1$. In the case of perfectly correlated marginal costs this allows the optimal price cap with commitment to be implemented in the second period, but such a mechanism may not be optimal. For an example, Baron and Besanko (1987c) characterize the optimal policy and show that the first-period mechanism

[32] The price with commitment may be higher than the price with fairness. Consider a triangular distribution $f_1(\theta_1) = 2(1 - \theta_1)$ if $\theta_1 \in [0, 1]$. Then, the period-two price cap with commitment is $p_2^*(\theta_2, \theta_1) = \left(\frac{\theta_2}{1-\theta_1}\right)\left(2 - \frac{3}{2}\theta_1\right)$, which is greater than the price with fairness for all θ_1 and θ_2 such that $\theta_1\theta_2 > 0$.

contains a countable number of policies. This mechanism is coarsely responsive to costs and is more responsive for low costs than for high costs. There is thus a trade-off between coarse pricing in the first period as a means of limiting opportunism in the second period and the welfare loss in the first period that results when prices are not continuously responsive to costs.

This suggests that regulatory mechanisms that are coarse and prescribe the same prices for sets of different possible costs may not be inefficient when long-term commitments cannot be made. Coarse price cap mechanisms thus may contribute to informational efficiency when the regulator is able to make only limited commitments to long-term policies.[33]

IV. Investment

A. Investment with Commitment

The purpose of this section is to examine the importance of commitment for investment by the regulated firm. As a benchmark the optimal investment with commitment is characterized, and the investment resulting in the absence of commitment is then examined.

If the regulator were able to make credible commitments, the incentives for investment would be second-best efficient given the informational asymmetry. The resulting investment depends on whether the regulator can observe investment and force the investment it prefers. The first case considered is that in which the regulator can dictate investment, and the second case considered is that in which investment is unobservable to the regulator in which case the investment decision is delegated to the firm. To illustrate these cases, the perfect correlation case in a two-period model will be used with $\theta_2 = \frac{\theta_1}{1+x_1}$. The welfare W maximized by the regulator can be written as

$$W = \int_{\theta^-}^{\theta^+} \Big[Y(Q(p_1)) - \theta_1 Q(p_1) - k_1 + \beta\Big(Y(Q(p_2)) - \frac{\theta_1}{1+x_1} Q(p_2) - k_2 \Big) - B(x_1) \\ - \Big(Q(p_1) + \frac{\beta Q(p_2)}{1+x_1} \Big) \frac{F_1(\theta_1)}{f_1(\theta_1)} \Big] f_1(\theta_1) d\theta_1. \tag{19}$$

The optimal (second-best) investment $x_1(\theta_1)$ satisfies

$$\beta \frac{\Big(\theta_1 + \frac{F_1(\theta_1)}{f_1(\theta_1)}\Big)}{(1+x_1(\theta_1))^2} Q(p_2(\theta_1)) - B'(x_1(\theta_1)) = 0. \tag{20}$$

[33] With an inability to commit, the effort chosen by the firm will again be second-best efficient. That is, in each period the firm simultaneously chooses the pricing policy and effort a_t given the mechanism M_t. The effort choice is efficient given the pricing policy p_t^o.

Given the quantity, the marginal product of investment is greater with the informational asymmetry than without it because, with this specification of the transition function, investment decreases the information rents earned by the firm.[34] The investment decision thus involves a rent-reduction externality.

With the specifications $B(x_1) = x_1$ and $F_1(\theta_1) = \theta_1$, the optimal investment is given by

$$x_1(\theta_1) = \begin{cases} 0 & \text{if} \quad 2\beta\theta_1 Q(p_2(\theta_1)) \leq 1 \\ (2\beta\theta_1 Q(p_2(\theta_1)))^{\frac{1}{2}} - 1 & \text{if} \quad 2\beta\theta_1 Q(p_2(\theta_1)) \geq 1. \end{cases} \tag{21}$$

Given the quantity, the investment with incomplete information is greater than the efficient investment for the same quantity by a factor of $2^{\frac{1}{2}}$. The price cap then satisfies, for $f_1(\theta_1)$ triangular,

$$p_2(\theta_1) = \frac{\theta_1}{1 + x_1(\theta_1)}\Big(1 + \frac{F_1(\theta_1)}{\theta_1 f_1(\theta_1)}\Big) = \frac{\frac{3}{2}\theta_1}{1 + x_1(\theta_1)} = \frac{\frac{3}{2}\theta_1}{(2\beta\theta_1 Q(p_2(\theta_1)))^{\frac{1}{2}}}.$$

If the investment were unobservable, then the firm would choose its investment to maximize its profits $\pi(\hat{\theta}, x_1; \theta)$ which, using the $T_1(\theta_1)$ and $T_2(\theta_1)$ that implement $p_1(\theta_1)$ and $p_2(\theta_1)$, are given by

$$\begin{aligned} \pi(\hat{\theta}, x_1; \theta) = (\hat{\theta}_1 - \theta_1)Q(p_1(\hat{\theta}_1)) + \beta\Big(\frac{\hat{\theta}_1}{1 + x_1(\hat{\theta}_1)} - \frac{\theta_1}{1 + x_1}\Big)Q(p_2(\hat{\theta}_1)) \\ + B(x_1(\hat{\theta}_1)) - B(x_1) + \int_{\hat{\theta}_1}^{\theta^+} \Big(Q(p_1(\theta_1^o)) + \frac{\beta Q(p_2(\theta_1^o))}{1 + x_1(\theta_1^o)}\Big)d\theta_1^o. \end{aligned} \tag{22}$$

The firm will choose its investment to satisfy, given $\hat{\theta}_1 = \theta_1$,

$$\beta\frac{\theta_1}{(1 + x_1(\theta_1))^2}Q(p_2(\theta_1)) - B'(x_1(\theta_1)) = 0. \tag{23}$$

For the same quantity, the firm thus invests less than the regulator prefers as given in (20).[35] The regulator will, however, take into account the firm's choice and alter the quantity

[34] The information rents are decreased by the investment because it reduces the range of costs to which the regulator must respond. The rents $\Pi(\theta_1)$ are

$$\Pi(\theta_1) = \int_{\theta_1}^{\theta^+} \Big[Q(p_1(\theta_1^o)) + \frac{\beta Q(p_2(\theta_1^o))}{1 + x_1(\theta_1^o)}\Big]d\theta_1^o.$$

[35] The investment level is given by the expressions in (21) with the 2 replaced by 1.

produced accordingly. Since the regulator prefers a greater investment than does the firm, the regulator will choose a lower price, and hence higher quantity, in the second period to increase the marginal product of investment.

B. Investment in the Absence of Commitment

If the regulator were unable to make credible commitments to future policies, it would at the end of the first period choose a mechanism M_2 that is optimal given its information. That information includes what can be inferred from the policy selected in the first period and from what is known about equilibrium strategies. If, for example, the mechanism M_1 implemented in the first period were continuously responsive to costs (completely separating), the regulator would know θ_1 from the firm's choice of policy in the first period. From the equilibrium strategies the regulator then could infer the investment chosen by the firm, so in the case of perfectly correlated marginal costs the regulator would know the marginal cost θ_2 the firm would have in the second period. The regulator could then fully exploit that information by instituting a marginal cost price cap. The firm would then only recover its initial investment and would have no incentive to invest.[36] To provide some incentive to invest, the regulator would choose a coarse mechanism for period one that would prevent it from learning θ_1.

If the costs were not perfectly correlated, the firm would earn rents in the second period on the information it privately observes at the beginning of period 2, and this could provide an incentive to invest. The expected rents $E\pi_2(\theta_1, x_1)$ are, after integrating by parts,

$$E\pi_2(\theta_1, x_1) = \int_{\theta^-}^{\theta^+} Q(p_2(\theta_2^0))F_2(\theta_2^o \mid \theta_1, x_1)d\theta_2^o. \tag{24}$$

The firm will choose its investment to maximize $E\pi_2(\theta_1, x_1) - B(x_1)$, and the marginal (value) product of investment is thus

$$\frac{dER_2(\theta_1, x_1)}{dx_1} = \int_{\theta^-}^{\theta^+} Q(p_2(\theta_2^o))\frac{\partial F_2(\theta_2^o \mid \theta_1, x_1)}{\partial x_1}d\theta_2^o. \tag{25}$$

If the investment reduces the range of marginal costs, the derivative in (25) is negative, so the firm will not invest when it recognizes that the regulator will fully exploit any information it

[36] The opportunism of the regulator here pertains only to information and because of the fair return requirement does not involve confiscation of second-period quasi-rents on the (sunk) investment made in the first period.

has at the end of the first period.[37] If the investment were to increase the range of marginal costs, then the firm might have an incentive to invest. One would expect, however, that that incentive would be weak and that the investment would be considerably lower than that preferred by the regulator.

As suggested by this example, when the regulator cannot make credible commitments to multiperiod policies, and thus can be expected to act opportunistically in response to information obtained, the incentive for investment is nonexistent or weak at best. Since public utilities make considerable capital investments, the question is what generates the incentive to invest. Several explanations seem plausible. First, partial commitments of the type agreed to by the NYPSC and the New York Telephone Company provide the firm an opportunity to capture returns on investments with rapid recovery rates or which are fungible and can be used to provide unregulated services. This explanation is analogous to regulatory lag and allows some profits to be earned before rates are adjusted. Second, the capital recovery rule may be sufficiently protective that regulated firms are confident that the return on non-fungible assets will be forthcoming. This would be characteristic of revenue requirements regulation that provides revenue as a function of original investment. Third, the prospect of deregulation and the returns that potentially can be earned under competition can provide an incentive to invest.

Fourth, an equilibrium may result in which the regulator and the firm give and honor trust.[38] The regulator has an incentive not to act opportunistically because it wants the firm to continue to invest to provide the capacity needed to serve a growing demand and to replace inefficient equipment and facilities. The firm has an incentive to invest because of the expectation that the regulator will forego the opportunity to take advantage of the information to confiscate profits. When the regulator does act opportunistically, the firm can punish the regulator by not investing and threatening that there will be inadequate capacity to meet demand. The regulator may then find it desirable to return to the strategy of honoring trust by not taking advantage of the firm and its non-fungible assets. This equilibrium, however, is subject to the short-run interest of politically ambitious regulators

[37] It is important to note that the firm would still be provided a fair return in this case through the amortization of $B(x_1)$. If a capital recovery rule is in effect, then $B(x_1)$ would be recovered over some set of periods.

[38] This requires an infinite horizon.

and legislators if they do not have to bear the political consequences of their opportunism. The possibility of such opportunism reduces the likelihood that such an equilibrium would be supportable. Even if such an equilibrium were attainable, the investment would likely be lower than that preferred by the regulator.[39]

V. Monitoring: An Example

The theory presented in the previous sections is based on the assumption that the regulator is only able to observe the policy chosen or, equivalently, the price, quantity, and the fixed charges. All regulators, of course, monitor closely the accounting profits of the firms they regulate. Accounting profits are not, however, the same as the economic profits that motivate the firm, and thus accounting profits are at best a noisy monitor of true profits. When commitment is possible, the availability of a monitor may not affect the prices specified in the regulatory mechanism, but in general the fixed charges will depend on the monitor. In the absence of commitment, the monitor will also be used to update the regulator's information about the costs of the firm.

The example presented here is intended to illustrate the role of an observable monitor of performance in regulating the firm. The example is "non-optimal" in the sense that the pricing policy in the first period is assumed to be constant over an interval, so the optimal intervals are not characterized as they are in Baron and Besanko (1987c). The example focuses solely on the second period and hence does not address how the observation of performance might affect the first-period policy.

The example has two periods and perfectly-correlated marginal costs $\theta_1 = \theta_2 = \theta$, which are uniformly distributed on the interval $[0,1]$. The cost C_1 incurred in the first period is assumed to be a function of θ and of a random variable that is not observed until the end of the period. The regulator is able to observe C_1, and hence it can use that observation to update its information in period two. Regulation is assumed to be governed by fairness, so the regulator is able to utilize fully this information in formulating the policy for the second period. The regulator must, however, offer a mechanism of price caps for

[39] In certain regulated industries, such as electric power, firms may have little incentive to invest in new capacity in those jurisdictions in which the regulatory environment is characterized by opportunism steming from political action. Capacity, however, appears to be sufficient that this is not a problem for the short-run.

the second period that allows the type of the firm revealed in the first period to earn non-negative profits. The price cap p_1 in the first period is set at the beginning of the period and thus cannot be based on C_1, which is not observed until the end of the period.[40]

The first-period cost C_1 specified as given by

$$C_1 = (\theta + \epsilon_1)q_1 + k_1, \tag{26}$$

where q_1 is the quantity and ϵ_1 is the realization of a randon variable $\tilde{\epsilon}_1$ uniformly distributed on the interval $[0,1]$.[41] The marginal cost thus depends on a random component and on the private information of the firm. Conditional on θ the density $g(C_1 \mid \theta)$ is

$$g(C_1 \mid \theta) = \frac{1}{q_1} \quad \text{if} \quad C_1 \in [\theta q_1 + k, (\theta+1)q_1 + k]. \tag{27}$$

The unconditional density function $g(C_1)$ is

$$\begin{aligned} g(C_1) &= \frac{C_1 - k}{q_1^2} \quad \text{if} \quad C_1 \in [k, q_1 + k] \\ &= \frac{2q_1 + k - C_1}{q_1^2} \quad \text{if} \quad C_1 \in [q_1 + k, 2q_1 + k]. \end{aligned} \tag{28}$$

If the regulatory policy in the first period pools over the interval $[\theta^-, \theta^+]$, the posterior density $f_2(\theta \mid C_1)$ at the beginning of the second period is uniform and given by

$$\begin{aligned} f_2(\theta \mid C_1) &= \frac{q_1}{C_1 - k} \quad \text{if } C_1 \in [k, q_1 + k] \\ &= \frac{q_1}{2q_1 + k - C_1} \quad \text{if } C_1 \in [q_1 + k, 2q_1 + k]. \end{aligned} \tag{29}$$

The support of θ however, depends on C_1, since if $C_1 \leq q_1 + k$ then $\theta \leq \theta^*(C_1) \equiv \frac{C_1 - k}{q_1}$, and if $C_1 > q_1 + k$ then $\theta > \theta^{**}(C_1) \equiv \frac{C_1 - q_1 - k}{q_1}$.

The marginal cost in the second period is assumed to be $\theta + \epsilon_2$, where ϵ_2 is the realization of a random variable $\tilde{\epsilon}_2$ uniformly distributed on $[0,1]$. Since the posterior density is uniform for any realization C_1, the price $p_2(\theta \mid C_1)$ set in the second period is[42]

$$p_2(\theta \mid C_1) = \theta + \frac{1}{2} + \frac{F_2(\theta \mid C_1)}{f_2(\theta \mid C_1)} = 2\theta + \frac{1}{2} \quad \text{if } \theta \geq \theta^*(C_1) \text{ and } C_1 \in [k, q_1 + k], \tag{30}$$

[40] The fixed charges T_1 can be based on the monitor.

[41] The mechanism M_1 is the pair $(q_1 = Q(p_1), T_1 = T_1^*)$, where T_1^* is a constant, so there is no self-selection in the first period.

[42] The term $\theta_1 + \frac{1}{2}$ is the expectation of marginal cost.

and

$$p_2(\theta \mid C_1) = 2\theta + \tfrac{1}{2} - \frac{C_1 - q_1 - k}{q_1} \quad \text{for } \theta \geq \theta^{**}(C_1) \quad \text{and } C_1 \in [q_1 + k, 2q_1 + k]. \qquad (31)$$

Comparing (30) with (12) evaluated at $\gamma = 1$ indicates that the price in period two is higher than in the absence of a monitor if $\theta + \epsilon_1 \leq 1$. If $\theta + \epsilon_1 > 1$, the price from (31) can be lower than that in (12). For high realizations of cost C_1, lower prices thus result because the support of the posterior distribution is $[\theta^{**}(C_1), \theta^+]$. This reduces the *ex ante* information costs to the regulator, and the regulator responds by reducing the price.

The rents $\pi_2(\theta \mid C_1)$ earned by the firm in period two then are given by

$$\pi_2(\theta \mid C_1) = \int_{\theta}^{\theta^*(C_1)} Q(2\theta^o + \frac{1}{2}) d\theta^o \quad \text{if } C_1 \in [k, q_1 + k],$$

and

$$\pi_2(\theta \mid C_1) = \int_{\theta}^{1} Q(2\theta^o + \frac{1}{2} - \frac{C_1 - q_1 - k}{q_1}) d\theta^o \quad \text{for } \theta \geq \theta^{**}(C_1) \quad \text{and } C_1 \in [q_1 + k, 2q_1 + k].$$

The rents $\pi_2(\theta)$ in the absence of the observable performance are given by

$$\pi_2(\theta) = \int_{\theta}^{1} Q(2\theta^o) d\theta^o \quad \forall \theta \in [0, 1].$$

The monitor of performance thus reduces the rents for low realizations of cost and may increase them for high realizations. The expected second-period rents may be greater or less than in the absence of the monitor, depending on the value of θ. Taking the expectation over θ indicates that the *ex ante* rents are reduced by the monitor.

Since the rents are affected, the period-one policy will be affected, but the effect is difficult to evaluate. Since the reason for pooling in the first period is that pooling reduces the information rents, intuition suggests that the availability of a monitor may reduce the incentive to pool in the first period. This would result in prices in the first period that are more responsive to first-period costs than is pricing in the absence of a monitor of performance.

The model analyzed here pertains to one regulatory jurisdiction and one regulated firm, but a commission may have authority over several firms in the same industry. In this case, monitors can be based on performance of all the firms. Since the private information of

the firms is likely to be correlated, such monitors should sharpen the regulator's posterior distribution. Similarly, monitors based on firms in other jurisdictions should be employed.

Monitors of performance can improve the efficiency of regulatory mechanisms in several ways. First, a monitor can reduce the rents of the firm by "tightening" the posterior distribution of marginal cost. This benefit is not present when commitment is possible, so it is the interaction of monitoring and policy revision that produces the benefit.[43] Second, a monitor can affect pricing by altering the marginal information costs. Third, if the firm were risk-averse, a monitor could be used to relieve the firm of some risk.

These benefits suggest that monitoring can be an important function of a regulator when a price cap or delegation mechanism is employed. From a positive perspective more active monitoring by a regulator would be expected to be correlated with an inability to make credible commitments. That is, monitoring is likely to be more valuable the more likely the regulator is to act opportunistically in response to information obtained through observation of performance.

VI. Conclusions

A variety of factors complicate the application of the principles addressed here. The major complicating factor is incompleteness of the policies due to unanticipated events and the costs of writing complex contingent policies. The theory presented above is based on the assumption that all possible events are known to the regulator and the firm and their likelihood of occurrence is representable by a probability distribution. In addition, if complexity and its associated costs preclude writing policies that are conditioned on each possible event or possible value of a cost parameter, the regulatory policies will be incomplete. In either the case of incompleteness or when there are unanticipated events, *ex ante* efficiency may be enhanced by allowing revisions in policies. What is needed are practical means to permit changes in regulatory policies when those changes promote *ex ante* efficiency and to preclude changes in policies when that would reduce *ex ante* efficiency.

In terms of performance evaluation the analysis presented here implies that it is the regulatory mechanism that should be the subject of the evaluation. The optimality of a regulatory mechanism is a function of 1) the information available to the regulator, 2) the

[43] If commitment were possible, the regulator would implement prices $p_t(\theta) = 2\theta + \frac{1}{2}, t = 1, 2$.

transition function that governs how the costs of the firm are anticipated to evolve over time, 3) and the extent to which commitments can be made credible. If credible commitments can be made, the framework presented in Sections III.B and IV.A. provides the basis for the evaluation. An important conclusion from this theory is that a search for *ex post* efficiency is likely to be misleading. A second conclusion is that even under delegation, administrative rules that provide a degree of commitment can improve efficiency. A capital recovery rule, for example, can reduce strategic behavior in addition to providing a fair return.

If credible commitments cannot be made, the cause of that inability must be assessed. If it is due to the policies of the regulator, then the consequences are attributable to the efficiency of regulation. If, however, the source of the inability to make credible commitments is due to other factors, such as opportunism by a legislature, then its consequences should be evaluated separately.

Appendix

Sunk Costs, the Inability to Commit, and Implementable Policies

For the case in which the marginal costs of the firm are perfectly correlated ($\theta_1 = \theta_2 = \theta$), Laffont and Tirole (1986*a*) have demonstrated that the regulator is unable to implement any regulatory policy that is continuously responsive to costs over any interval of possible costs that has positive probability. This results because of a conjunction of the conditions required for responsive pricing and the expanded strategy set of the regulated firm when it can make its participation decision in each period. Since the firm has a natural incentive to overstate its costs in an attempt to obtain a more profitable regulatory policy (i.e., to choose a policy intended for a higher cost type), it is necessary that the regulator choose the fixed charges to offset that incentive. To do so in a multiperiod model in which the regulator cannot commit credibly to future policies, the policy in the first period must include a payment that offsets for both periods the incentive to select a policy designed for a higher cost type. The rent $\pi(\theta)$ required to implement any prices $(p_1(\theta), p_2(\theta))$ is thus

$$\pi(\theta) = \int_{\theta}^{\theta^+} [Q(p_1(\theta^o)) + \beta Q(p_2(\theta^o))] d\theta^o.$$

If the firm were to select the policy defined for its marginal cost, the regulator in the second period would, knowing the marginal cost θ, choose a policy of marginal cost pricing (i.e., $(p_2(\theta) = \theta,\ T_2(\theta) = k_2)$), which yields the firm a zero profit. If the firm selected $\hat{\theta} \neq \theta$ and produced in the second period, the rent $\pi_2(\hat{\theta}; \theta)$ in the second period would thus be

$$\pi_2(\hat{\theta}; \theta) = (\hat{\theta} - \theta) Q(\hat{\theta}),$$

which is negative if the firm selected in the first period a pricing policy intended for a $\hat{\theta}$ less than θ.

Because the firm can always limit its period-two profit to zero by refusing to participate in the second period, its two-period profit $\Pi(\hat{\theta}; \theta)$ is

$$\Pi(\hat{\theta}; \theta) = (\hat{\theta} - \theta) Q(p_1(\theta)) + \int_{\hat{\theta}}^{\theta^+} [Q(p_1(\theta^o)) + \beta Q(p_2(\theta^o))] d\theta^o + \max\{0, (\hat{\theta} - \theta) Q(p_1(\hat{\theta}))\}.$$

Given this profit function, the firm finds it optimal to underreport ($\hat{\theta} < \theta$) its costs in the first period and then not to participate in the second period in order to avoid having to

produce at a loss.[44] This conclusion holds for all intervals of marginal cost, so no policy that is continuously responsive to costs on any interval can be implemented.

In the case of a regulated firm with long-lived assets that are sunk and non-fungible, the capital recovery rules used by regulatory commissions may eliminate the problem identified by Laffont and Tirole. A capital recovery rule is taken here to be a rule enforceable under administrative law that entitles the firm to recover an asset's cost according to a prescribed schedule as long as the firm continues to produce the quantity specified in the regulatory policy. Suppose that the firm has a sunk investment $B(x_0)$ prior to period one, and suppose that an enforceable recovery rule allows the firm to recover that cost at a constant rate over the n-period life of the asset. In a two-period model ($n = 2$) the cash flow in the second period is now $\frac{1}{2}B(x_0)$ if the firm continues to participate under a marginal cost pricing policy. The firm's two-period profit $\Pi^*(\hat{\theta};\theta)$ is thus

$$\Pi^*(\hat{\theta};\theta) = (\hat{\theta}-\theta)Q(p_1(\hat{\theta})) + \int_{\hat{\theta}}^{\theta^+} [Q(p_1(\theta^o)) + \beta Q(p_2(\theta^o))]d\theta^o$$
$$+ \max\{0,\ \tfrac{1}{2}B(x_0) + (\hat{\theta}-\theta)Q(p_1(\hat{\theta}))\}.$$

If the sunk assets are sufficiently great, the incentive to underreport costs in the first period and not participate in the second period may be outweighed, allowing the regulator to implement a policy even if it cannot commit to its pricing policy in the second period. This results because a capital recovery rule provides a limited form of commitment that allows compensation to be deferred.[45,46]

The significance of sunk, non-fungible assets and a capital recovery rule is that even in the absence of commitment the regulator may be able to implement a regulatory policy that is continuously responsive to costs. In particular, it may allow the regulator to implement a policy that fully exploits the information revealed in the first period. Such a policy is not generally optimal, however, as indicated in Baron and Besanko (1987*c*).

---

[44] Technically, the incentive compatibility constraints bind both upwards and downwards.

[45] This effect is offset if the firm has financed the asset with debt. For example, suppose that an η share of $B(x_0)$ is financed with debt with a repayement schedule with half repaid in each period. The cash flow of the owners of the firm is then only $\frac{1}{2}(1-\eta)B(x_0)$, which provides a diminished incentive to participate in the second period.

[46] If the asset were fungible so that the firm could earn $\frac{1}{2}B(x_0)$ on the asset employed elsewhere, the capital recovery rule would be ineffective.

References

Baron, David P. (1987), "The Design of Regulatory Mechanisms and Institutions," in *The Handbook of Industrial Organization*, R. Schmalansee and R. Willig, eds., forthcoming.

Baron, David P. and David Besanko (1984), "Regulation and Information in a Continuing Relationship," *Information and Economics and Policy*, Vol. 1, pp. 267-301.

Baron, David P. and David Besanko (1987a), "Monitoring, Moral Hazard, Asymmetric Information, and Risk Sharing in Procurement Contracting," *Rand Journal of Economics*, (forthcoming).

Baron, David P. and David Besanko (1987b), "Monitoring of Performance in Organizational Contracting: The Case of Defense Procurement," Working Paper, Stanford University.

Baron, David P. and David Besanko (1987c), "Commitment and Fairness in a Dynamic Regulatory Relationship," *Review of Economic Studies*, Vol. 54, pp. 413-436.

Baron, David P. and Roger B. Myerson (1982), "Regulating a Monopolist with Unknown Costs," *Econometrica*, Vol. 50, pp. 911-930.

Besanko, David and David E. M. Sappington (1986), "Designing Regulatory Policy with Limited Information." In *Fundamentals of Pure and Applied Economics*, (forthcoming).

Caillaud, B. (1986), "Regulation, Competition and Asymmetric Information," Working Paper, Massachusetts Institute of Technology.

Caillaud, B., R. Guesnerie, P. Rey, and J. Tirole (1985), "The Normative Economics of Government Intervention in Production in the Light of Incentives Theory: A Review of Recent Contributions," Working Paper , Massachusetts Institute of Technology.

Demski, J.S., Sappington, D.E.M., and Spiller P.T. (1985), "Entry in Regulated Industries: An Information-Based Perspective," Working Paper , Bell Communications Research.

Federal Communications Commission (1987), "FCC Proposes Replacing Rate-of-Return Regulation for Dominant Carriers with "Price Caps" Regulation for Interstate Services (CC Docket 87-313)," Washington, D.C., August 4.

Guesnerie, R. and J.J. Laffont (1984), "A Complete Solution to a Class of Principal-Agent Problems with an Application to the Control of a Self-Managed Firm," *Journal of Public Economics*, Vol. 25, pp. 239-369.

Helman, Leonard A. (1984), "Rulemaking in Lieu of Ad Hoc Case-by-Case Adjudication: Automatic Adjustment Clauses, the New Mexico Experiment," in *Regulatory Reform*, J.R. Foster, G.R. Hall, S.R. Holmberg, C.F. Phillips, Jr., and R.L. Wallace, eds., Institute for Study of Regulation, Washington, D.C., pp. 148-153.

Joskow, Paul R. and Schmalensee, Richard (1986), "Incentive Regulation for Electric Utilities," *Yale Journal of Regulation* Vol. 4 (Fall), pp. 1-49.

Jubelier, Robert C. (1987), "Comment," in Discussion Paper #22, Fishman-Davidson Center, Wharton, University of Pennsylvania, June.

Kahn, Alfred E. (1987), "The Future of Local Telephone Service: Technology and Public Policy," Discussion Paper #22 Fishman-Davidson Center, Wharton, University of Pennsylvania, June.

Laffont, J.J. and J. Tirole (1986a), "The Dynamics of Incentive Contracts," Working Paper, Massachusetts Institute of Technology, Cambridge, Mass.

Laffont, J.J. and Tirole, J. (1986b), "Using Cost Observation to Regulate Firms," *Journal of Political Economy*, Vol. 94, pp. 614-641.

National Telecommunications and Information Administration (1987), "Regulatory Alternatives Report," Department of Commerce, July.

New York Public Service Commission (1987), "Opinion and Order Extending Rate Case Moratorium," Opinion No. 85-178 (D), May 11.

Noll, Roger E. (1987), "Economic Perspectives on the Politics of Regulation," in *Handbook of Industrial Organization*, Richard Schmalensee and Robert Willig (eds.), North-Holland, (forthcoming).

Noll, Roger G. and Owen, Bruce M. (1987), "United States v AT&T: An Interim Assessment," Working Paper, Stanford University, June.

NTIA (1987), "NTIA Regulatory Alternatives Report," Department of Commerce, Washington, D.C., July.

Sappington, D.E.M. (1982), "Limited Liability Contracts Between Principal and Agent," *Journal of Economic Theory*, Vol. 29, pp. 1-21.

Sappington, D.E.M. and J. E. Stiglitz (1986), "Information and Regulation," Working Paper, Bellcore.

Sappington, David E.M. (1983), "Optimal Regulation of a Multiproduct Monopoly with Unknown Technological Capabilities," *Bell Journal of Economics*, Vol. 14, pp. 453-463.

Schmidt, Michael (1984), "Rulemaking in Lieu of Case-by-Case Adjudiction: An Overview of Automatic Adjustment," in *Regulatory Reform*, J.R. Foster, G.R. Hall, S.R. Holmberg, C.F. Phillips, Jr., and R. L. Wallace, eds., Institute for Study of Regulation, Washington, D.C., pp. 136-147.

Vickers, John and Yarrow, George (1985), *Privatization and the Natural Monopolies*, Public Policy Centre, London.

Vickers, John and Yarrow, George (1988), *Privatization in Britain*, MIT Press, Cambridge, MA, (forthcoming).