THE DESIGN OF REGULATORY RULES

by

Ingo Vogelsang*
University of Bonn and MIT Energy Laboratory

Energy Laboratory Working Paper No. MIT-EL 79-027WP
May 1979

*I should like to thank Michael Crew, Paul Joskow, Sam Peltzman, Richard Schmalensee and Roger Sherman for helpful suggestions during earlier production stages of this paper.
The Design of Regulatory Rules
Ingo Vogelsang, University of Bonn and M.I.T. Energy Laboratory.

1. Introduction

The call for economists to design regulatory rules in order to improve regulation is an old one, but until now has only brought forward one practical breakthrough: deregulation. Otherwise, regulators still adhere to rules established during regulatory practice and mostly dating back many decades. Economists might want to explain their shortcoming by the nonapplicability of the exclusion principle. A regulatory rule is a public good. However, professional people hardly ever apply their principles to their own work. So economists try hard, even at no direct pecuniary reward. Hence one may conjecture that designing regulatory rules contains intrinsic difficulties. Schmalensee (1979) impressively illustrates this point. On the other hand he also shows that a variety of regulatory mechanisms have been proposed which for some reason or other are all unsatisfactory to the analyst. So far they did not convince the public decision makers either.

A regulatory rule can be viewed as part of the property rights structure imposed on an industry. In German the termuris "Ordnungspolitik" would be used to characterize the establishing of regulatory rules. That means they are a subset of the rules for feasible actions in the economy as a whole. But, whereas property rights defined on commodities are
exchangeable between economic agents, regulatory rules establish no direct claims. Implementing them involves regulatory agencies as mediators. This feature relates them to the allocation mechanisms discussed by Hurwicz (1973). Allocation mechanisms, however, are designed to govern all markets at the same time, while a regulatory rule deliberately applies to one market only. In doing so one takes the performance of other markets as given. The Second Best issue in this sense is assumed to be solved by choosing the scope of application of a regulatory rule.

In Section 2 we start by setting out a basic rationale for regulatory rules based on the theory of agency. Though far from being realistic it provides a takeoff point to accept the goals of regulation by rules as a given. Section 3 then introduces a number of postulates which regulatory rules should obey. Nonadherence to them would lead to regulatory failure in a sense that economists may agree upon. On the other hand, possibly not all of them can be fulfilled at the same time. Nonadherence to any single one does not necessarily make regulation infeasible. Then indeed, in a Second Best manner regulators could often do better by violating other postulates as well.\(^1\) Hence, we use the first set of seven postulates to distinguish regulatory rules from other types of government intervention and to indicate when such rules may be appropriate tools. In particular, the discussion of these postulates reveals the adverse

\(^1\) There is always a state contingent behavior that is at least as good as the rule, but it may go high in transaction costs.
consequences of discretionary power enjoyed by regulators. Postulates 8 and 9 set out performance standards of regulation, whether or not the rule approach outlined by postulates 1 to 7 is taken.

In Section 6 we deal with two rules aimed specifically at the pricing issue of regulation. They shall both illustrate the postulates laid out in Section 3, but the first example also casts some fresh light on rate of return regulation. We submit that under any realistic regulatory pricing rule an adjustment lag is inevitable. According to by now conventional wisdom such a lag stimulates the firm to lower costs during lag intervals, because it does not immediately have to pass on the additionally gained profits as lower prices. Against this we argue that the lag may as well drive the firm to increase costs now in order to receive a price increase later. A dynamic price adjustment mechanism for the multiproduct case suggested by Vogelsang and Finsinger (1979) is discussed as the second example because its development was motivated by some of the postulates discussed here. Therefore, the rule fares exceedingly well under these postulates which relate to informational requirements. The price to be paid for this is limited applicability.

In addition, low information requirements in combination with an adjustment lag could again induce misrepresentation through wasteful action by the regulated firm. The last section deals with rules not directly related to price. As we shall restrict ourselves throughout to natural monopoly regulation this section only scratches the surface of problems like the regulation of externalities.
2. The Rationale for Regulatory Rules

Regulatory rules fit well into Goldberg's (1975) notion of regulation as an administered long-term contract. In this framework regulators act as agents, whereas buyers and sellers act as potential principals. Regulatory rules would correspond to the clauses of such a long-term contract. Like contract clauses they constrain further behavior. There are two polar alternatives to regulatory rules. One is the absence of regulation. This "totally constrains the regulators," but adds no constraint on the regulated firm. It may be optimal e.g. if in a stationary environment a natural monopoly with decreasing average cost can just hope to break even at profit maximizing prices (in the multiproduct case the Ramsey number would be one). The other extreme is not at all to constrain the regulators and have them prescribe directly all actions the firm can take. This may be optimal in a rapidly changing environment where regulators serve only to put out fires. I shall reserve the term "regulatory rule" for something constraining both firm and regulators (and implicitly consumers). Ideally, for a given state of nature a regulatory rule shall completely determine the actions to be taken by the regulators, whereas the actions of the regulated firm shall be determined by the interaction between the regulatory rule, the market constraints and the self-interest of the firm. As indicated above, situations exist where the other extremes or combinations thereof are preferable to a regulatory rule. The general motivation for regulatory rules comes from the link between regulation as a permanent institution and market failure. If market failure is the basis for regulation then
to the extent that the failure itself is stable one would want to compensate for this by a stable mechanism. To wit, in a persistent natural monopoly situation one would want to have a regulatory rule that allows for cost savings by maintaining a monopoly position and at the same time balances market power between both sides of the market. Thus the rule shall help to achieve that the property rights structure with respect to this industry is known and well defined to everybody involved. In this sense no regulation comes closer to qualifying as a regulatory rule than does direct regulation of activities.

Regulatory rules are a normative construct. In the case of natural monopoly their purpose can be taken as serving consumers' interest. The main reason for this would be to compensate for the market power of the single supplier. However, if regulation is itself backed by the state's coercive power, there may be a case for regulation to serve the producer's interest as well. Clearly, the producer should be induced to stay in business as long as his service is valued highly enough by the consumers. This is a natural constraint on regulation. Its pendant ex ante is that the threat of regulation to an industry should not induce an efficient firm to leave it. Furthermore, the attainment of efficiency may require the firm to make a profit. Hence a possible goal of regulation may be to maximize the sum of consumers' and producers' surplus subject to a breakeven constraint for the firm.
We emphasize regulatory rules as policy tools and procedures to implement prespecified goals. So, goal setting is not the main topic of this article. Nevertheless it is important in order to understand the agency relationship. Regulators as agents cannot easily serve conflicting interests. Trying to maximize consumers' surplus\(^2\) subject to a break-even constraint seems to be more in line with this. Still, with a multitude of buyers as principals it becomes hard to believe the story of the principal-agent relationship. In fact, "the principal is unidentifiable" (Aharoni, 1979). However, we may have recourse to conceived voting procedures determining the principal and his goals.

Then the goal "maximization of consumers' surplus" appears as the result of a peculiar fictitious bidding process. If regulatory goals had to be auctioned off among consumers and if consumers would state their true willingness to pay for the imposition of a regulatory goal (the goal is a public good); consumers' surplus maximization would receive the highest bid under the usual partial equilibrium caveats. Other goals such as Pareto improvements correspond to other voting mechanisms, but probably consumers' surplus maximization comes closest to simulating economic evaluation.

We left open so far why the consumers would want to have regulators as their agents and not negotiate with the firm directly. The literature on competitive franchise bidding (e.g., Williamson, 1979) at least gives a

\(^2\)Maximizing consumers' or social surplus as an official goal may tremendously reduce subsequent conflicts if measurement problems could be overcome.
partial answer to this question. It centers around the observation that after the bidding process has been carried out for the first time the winning bidder has a first mover advantage in the next round. Excluding the currently supplying firm from the next auction leads to socially inefficient behavior at the end of the contract term. Von Weizsäcker (verbally) circumvents this problem by randomizing the exit of the currently successful bidder. Even this procedure will not be socially optimal, if the peculiar information the supplier has received during his tenure is of social value.

Optimizing supply in a market is an infinite horizon problem which franchise bidding splits up into consecutive finite horizon solutions. Regulation instead is generally stuck with the infinite horizon but brings in coercive power. Regulators are backed by the state and can thus save on transaction costs which would be caused by repeated direct renegotiation of contracts. Such costs indeed would be formidable between one seller (not always the same) and many buyers.

Now regulatory rules enter the picture as constraints on the behavior of regulators once the regulatory process has been established. Without such constraints regulators would face continuous pressure to alter their behavior from all groups affected. Hence the concept of regulation as a long-term contract would lose credibility. Instead it would give way to political processes of the sort described by Fiorina and Hall (1978 and 1979). Indeed, if these two authors are right in their view of political processes regulatory rules will largely remain to be a normative concept,
because they could only be enacted against the will of politicians.\footnote{3}

For then, in trying to gain or keep their constituency politicians continuously would have to interfere with issues concerning a narrowly specified group. Establishing general rules would make such politicians superfluous. Against this hypothesis two tendencies could help regulatory rules to gather momentum in the real world. First, politicians can diversify their strategies in order to be able to succeed. Because a regulatory rule supposedly establishes long run benefits, chances are nonnegligible that the constituency of a politician will prefer it to shortrun and uncertain interventions. Furthermore, a presidential candidate may have to carry through actions different from a member of parliament. Second, regulation may be influenced by court decisions to even a greater extent than by politicians.

3. A Set of Postulates for Regulatory Rules

By a regulatory rule we mean a verbal or mathematical statement concerning reactions of regulators on the behavior of firms or consumers in regulated markets that obeys the following postulates:

P1 (No discretion): It constrains the actions of the regulators.

P2 (Enforceability): It is enforceable on the regulated firms.

\footnote{3}Other groups such as regulatory commissions and their staff, who currently hold discretionary power, can also be expected to oppose their introduction.
P2 (Publicity): It is known to all parties affected (at low cost).
P3 (Duration): It is valid for some length of time.
P5 (Privacy): Its language concerns variables principally observable to all parties affected (at low cost).
P6 (Feasibility): It must allow the regulated firm at least to expect to break even.
P7 (Limitations): It must state observable conditions under which it is applicable, furthermore, who decides and which due process requirements pertain whenever it is not applicable.

I shall first discuss these postulates one by one and then say something about evaluating between regulatory rules.

P1 (No discretion): This postulate in its narrowest form means that the rule completely determines the actions of the regulator. In this case the agents in the market would accept the rule as a constraint very similar to the scope of property rights on physical assets. Thus in an institutional sense the rule would not interfere with the market mechanism. This has three more immediate consequences.

First, regulation would not introduce policy uncertainties, which seem to be severe consequences of all types of government intervention (see e.g. EPRI, 1978). Such uncertainties may lead to government induced market failure, because without rules regulatory agencies could behave
opportunistically. One may hold against this position that there is uncertainty about virtually any economic agent's behavior. However, ordinarily the market sufficiently constrains the freedom to behave opportunistically. Otherwise for this very reason market failure results (Williamson, 1975).

Second, there will be less of a wasteful competition to achieve a transfer of (quasi-) rents between actors in regulated markets. The rent transformation argument against regulation as brought forward by Posner (1975) could indeed lose its strength. Setting Posner's dubious welfare estimations aside, his main conjecture is that the possibility to capture rents induces competitive expenditures aimed at appropriating them.

Regulators command over the potential rents of a natural monopoly. Their decisions supposedly can be influenced by bribes, litigation, propaganda or future career opportunities. If so, (risk neutral) agents in this market will incur costs to change regulatory behavior in their favor as long as the expected rent transfer exceeds the expenditure. Regulatory rules could substantially reduce the success probabilities for a given expenditure of this kind. However, the rent transformation argument could still apply to the level of regulatory rulemaking. As long as different levels of policy making offer similar opportunities for firms to achieve rent transfers there will be a potentially infinite regress in this kind of reasoning. Against this I submit the hope that law makers and judges are harder to influence than regulatory commissioners. The transaction costs of changing a law then will be substantially higher than those of of changing vaguely prescribed regulatory behavior.
other hand a rule change can provide higher and longer lasting benefits. The result of this tradeoff in reality is not a priori clear. In a normative setting it leads to the metarule postulate that rule changes should be made difficult, e.g. require two-thirds majority in parliament.

Third, even if a regulator sincerely tries to implement a prespecified regulatory goal the very existence of his coercive power could weaken his ability to choose an optimal policy, provided that he and the regulated firm hold rational expectations on their mutual behavior. This paradox has been demonstrated by Kydland and Prescott (1977) as well as Maskin and Newbery (1978) in a related context. They show that a rational government may want to change its optimal plan over time, thus exploiting the outcomes of its previous action. If this change can be anticipated by the economic agents affected through such a policy they will accordingly alter their behavior ab initio. This again can be anticipated by the government leading to an iterative process. The resulting equilibrium is likely to be inferior compared to the situation where the government only had the possibility to design and implement its initial plan. This limitation is precisely the content of postulate PI.

As a very simple example for this argument, assume that the optimal policy for a regulatory agency in a stationary environment is to announce a price ceiling which over time declines to a steady state in which the firm can just hope to break even at minimum costs. Further, assume that the agency does not know the true demand and cost function of the firm and thus adds a risk factor to all ceiling prices. Now, if time reveals
the true cost to be lower than that for which the ceiling prices have been calculated, the agency could do better by decreasing the prespecified prices. Knowing this, the regulated firm lacks an incentive to incur cost lowering investments from the outset.

The more freedom of action the rule still gives to the regulator the less will the three arguments in favor of postulate P1 hold.

**P2 (Enforceability):** This postulate simply means that a regulatory rule is backed by the coercive power of the state and that it is a statement on the behavior of the regulated firm. A contract clause to which both parties of a sale have agreed is similar in these respects. However, regulation exists before explicit agreements of such kind.

**P3 (Publicity):** Fulfillment of this postulate ensures that the obligations and benefits of the rule reach all parties concerned. Just as it may be difficult to guess or calculate the next number in a sequence if the underlying formula is not known, it may be extremely hard to predict future regulatory actions simply from past decisions of the regulators. So a rule has to be published.

**P1 (Duration):** Ideally a regulatory rule should be valid indefinitely or at least coincide with the time horizon of the farthest forward looking decision maker affected by it. Whenever this is not so, a functioning of the regulated market with respect to investments is impeded by regulatory uncertainty and pressure to change the existing rule. On the other hand,
the continuation of inefficient behavior or the prevention of organizational innovation could pose severe problems overruling this postulate. Seen from a different perspective, a regulatory rule is a costly organizational investment, which in a metaphorical sense is profitable only if it is used long enough. After being instituted there may evolve situations calling for a change, but then costs of introducing the old rule have to be taken as sunk.

P5 (Privacy): This postulate refers to a diverse set of decision makers including consumers, regulators, the regulated firm and its dependents. All these persons at least in principle must be able to supervise if the rule was obeyed by the regulators and the regulated firm. Otherwise postulate P1 cannot really hold either.

The different access to information which the different groups quite naturally have, may necessitate auxiliary devices like the use of auditors. Still, taken together P3 and P5 mean that regulatory rules have to be simple and have to be based on simply assessable data. The main corollary of Postulate P5 is that the peculiar private information of the participants in the "regulation game" is either taken care of in the incentive structure imposed by the rule and thus irrelevant as a data input, or it must be capable of transformation into publicly controllable information. This is taken up in postulate P8.

Loeb and Magat (1970) propose an incentive mechanism where the regulated firm gets the total consumers' surplus as a subsidy. This will induce
the firm to produce and price efficiently. Their scheme does not qualify under P5. It gives rise to haggling, because no method to measure consumers' surplus exists, such that all agents would agree on its outcome. Furthermore, it depends on subsidies thus violating P6.

P5 (Feasibility): This postulate requires a regulatory rule to be feasible in a private market system. If it were not fulfilled, this could have three consequences. First, no firm would be willing to enter the market. Second, a firm already in the market would be expropriated. Such a consequence has to be distinguished from the possibility that even with no constraint the firm would not be able at least to break even, because e.g., the life cycle of the market has come to its end, as possibly true for the railroads, or the postal service. This is not necessarily the result of regulation. However, regulation could already have provided the wrong signals ex tunc. In either case, ex nunc a change of regulatory rules could come too late. Third, a violation of P6 may mean the necessity of outside subsidies. Through these, a firm generally loses its independence. The arguments on uncertainty and rent transformation brought forward in connection with postulate P1 would again hold. Furthermore, taxes to raise the subsidies tend to distort other parts of the economy.

P7 (Limitation): It seems inconceivable to formulate a regulatory rule obeying P1 to P6 and being applicable in all states of the world. This is analogous or may even be identical to the impossibility of a complete contingent claims contract covering any substantial length of time. The
The best that can be achieved under a regulatory rule is a specification of those states of the world in which it shall apply and an accompanying metarule on what is going to happen if it is not applicable. Such a metarule would have to specify two things in particular. First, through a mechanism to elect regulators, one may hope that regulation outside the scope of the regulatory rule is performed in the principal's interest. Second, procedural rules on how to arrive at decisions (due process requirements) may reduce the uncertainty of regulatory decisions (Owen and Graeutigan, 1972).

Summing up and simplifying the seven postulates we define a regulatory rule as a law or precedent constraining the decision of regulators. By making use of easily observable variables only it has self-enforcing properties as far as it is applicable.

Now a regulatory rule obeying the above seven postulates is not necessarily a good one. Such a rule could lead to a result departing substantially from the stated goal of regulation. Let us assume that the rule was installed and actuated with the goal formulation. Then actually the decision on the regulatory rule was meant to be the goal formulation. In this case such a departure can only happen if the effects of the rule differ from those anticipated by the decision makers. Leaving inconsistencies and false application aside, one main reason for this to occur is that the rule was based on information different from the information actually used by the decision makers in the market.

Postulate 23 defines the information feasible for applying a rule to be
observable data. This requirement is normally restrictive and important, because firms and consumers hold information which is not readily transferable to a regulatory agency. However, unless sufficiently motivated by the rule they will use such information to their own interest causing moral hazard and adverse selection problems. This leads to postulate P8 (No misrepresentation): A regulatory rule should give incentives for those to whom it concerns not to misrepresent their actions, preferences and possibility sets.

This postulate is weak in the sense that together with postulates P1 to P7 it does not guarantee an optimal outcome vis-a-vis the goal of regulation. Clearly, when installing a regulatory rule neither the legislator nor the regulator knows all preferences and possibility sets. If postulate P8 is obeyed, however, time could allow to learn more about them and make use of this additional information. In order ultimately to achieve the optimal goal fulfillment a stronger postulate is necessary. For then the rule must give the regulated firm an incentive to use its remaining freedom (only) to improve its actions vis-a-vis the regulatory goal. This is captured by postulate P9 (Optimality): The incentives provided by the regulatory rule on the free behavior of the regulated firm should lead the firm to maximize with respect to the regulatory goal.

In this postulate costs of instituting and implementing the regulatory rule are neglected. The postulates P1 to P7 provide formal procedural requirements the fulfillment of which could be observed by almost
anybody. Postulates PC and P2 are substantially more difficult to control. Usually, only sophisticated economic analysis can hope to establish if a rule conforms to them or not.

A given regulatory rule can be judged by its expected performance in terms of goal fulfillment. Relevant parameters are the expected value and the risk of the goal fulfillment. The expected value depends mainly on those factors that have a systematic impact on goal fulfillment. These are the items usually inquired into by students of regulation. Given the preferences of the regulated firms and consumers, how do regulatory constraints affect market performance? The postulates stated above concentrate more on reducing policy uncertainty introduced by regulation. Other things equal, a reduction in policy uncertainty will be achieved

(1) the more the regulator is constrained in his behavior,
(2) the simpler the language of the rule,
(3) the more states of the world the rule covers,
(4) the more difficult it is to abolish or change the rule.

A priori it is quite unclear what effects on uncertainty tighter constraints on the regulated firm's behavior have.

Whereas regulatory rules by themselves tend to reduce policy uncertainty they reduce the flexibility of government. Thus they also deprive government of a tool to reduce uncertainty coming from other actors of the environment. There seems to be no general answer to the question
which type of uncertainty is more severe. We have introduced the notion of regulatory rules especially to govern long run repeated transactions. If certain types of uncertain situations repeatedly come up during the course of this relationship, a regulatory rule could be appropriate to cope with them. This is definitely not true to resolve unique catastrophic events. Here, discretionary coercive power will be badly needed. However, because events hardly ever come as full surprises to those agents who are deeply involved, there remains a genuine tradeoff between incentives for precautionary measures initiated by rules and ex post remedies through discretionary coercive actions of government. This justifies postulate P7. The weight given ex ante to this escape clause determines if the regulatory rule approach is appropriate for a particular industry.

4. Examples for Regulatory Pricing Rules

4.1 Rate of Return Regulation

In reference to regulation the rate of return has been described as the small tail to wag a big dog (McKee, 1970). It certainly is the single most prominent yardstick applied in monopoly regulation. This is not restricted to the U.S. but becomes increasingly apparent in Western European countries, e.g., in electricity rate regulation in West Germany. Rate of return (ror) regulation means that the regulated firm is not allowed to earn a profit \( \pi \) higher than a prespecified net rate of return \( s - r \) on its total employed capital valued \( cK \), i.e., \( \pi \leq (s - r)cK \), where \( r \) is the interest rate and \( c \) the acquisition cost of capital \( K \) to
the firm. If effective \( r < s < \text{unconstrained gross rate of return} \) this constrains the firm in all aspects of its behavior but at the same time does not preclude freedom on all principal parameters with which according to Kahn (1970, p. 3) regulation should be concerned.

The firm can still choose

1. to leave the market, enter other markets, but others can enter its market as well,
2. product quality and conditions of service,
3. to serve whom it wants,
4. prices and profit.

In the following we assume this freedom of choice to be restricted to price and ask how ror-regulation complies with the postulates one by one:

**P1** (No discretion): The regulator can choose exactly one parameter, the allowed rate of return \( s \). He is constrained above by the profit rate the regulated firm could earn with no regulation and below by the firm's cost of capital. If he goes above one, or below the other, further changes of \( s \) have no effect on the behavior of the firm. Within this constraint interval, however, the regulator may have substantial discretion, which is hard to control by outsiders. In reality this can lead to describing ror-regulation as a fiction of economists (Joskow, 1974).

**P2** (Enforceability): In the pure sense that \( \pi_t \leq (s - r)cK_t \) for every time period \( t \) ror-regulation is not always enforceable unless retroactive corrections are allowed for. Because retroactive corrections interfere
with the market transactions between the regulated firm and its customers, we shall discard them. Then uncertainty of future events plus the time necessary to compute actual profits necessitate a regulatory adjustment lag. We shall therefore assume that such a lag exists. So in the following ror-regulation means that regulators allow prices for a time period \( t \) based on the actual rate of return in period \( t - 1 \):

\[
p_t \leq \frac{s c K_{t-1} + w L_{t-1}}{x_{t-1}}
\]

where \( w \) = wage rate, \( L_{t-1} \) = labor input in period \( t-1 \), \( x_{t-1} \) = quantity traded in \( t-1 \). In this modified form, ror-regulation can be enforced, though possibly in no period the firm earns less than the allowed rate of return \( s \).

\( P_3 \) (Publicity): This postulate generally can be satisfied. A change in \( s \) may need some publication effort.

\( P_4 \) (Duration): Formally this can be satisfied. However, the spirit of this postulate would be violated, if regulators changed \( s \) at short notice and in a capricious way. Thus, for ror-regulation \( P_4 \) is closely linked to \( P_1 \). Both postulates (and some others) are the topic of the Hope decision of the U.S. Supreme Court, which established American ror-regulation in its current form. The decision confirms that the ror-regulators have some discretion.

\[4\] In a puristic sense this should be \( t - 1 - \epsilon \), because it takes time \( \epsilon \) to compute profits and measured profits always refer to a period, not a point in time.
P5 (Privacy): Fulfillment of this postulate again necessitates the existence of a regulatory lag. Otherwise it can be satisfied by ror-regulation precisely because the unobservable cost of capital $r$ has been replaced by the allowed rate of return $s$.

P6 (Feasibility): In an uncertain environment the postulate to allow the firm a nonnegative profit expectation implies that there is a minimal feasible $s^* > r$ such that the allowed rate $s \geq s^*$. This holds if regulators truncate the upper tail of the probability distribution of future profits. Because regulators do not know cost and demand functions they cannot identify the effect of uncertainty. Hence if they want to limit excess profit they can only do that ex post, thus truncating the ex ante probability distribution of profits. Especially in times of inflation or scarcity such truncation could become highly important.

P7 (Limitation): After acknowledging discretion in the choice of $s$ (and the length of the regulatory lag) and with the proper choice made ror-regulation is feasible under almost any state of the world. Hence P7 can be assumed satisfied.

We may conclude that ror-regulation obeys the seven postulates defining a regulatory rule, but leaves room for discretion. Now what does one buy for that in terms of goal fulfillment?

In principle, administrative costs of ror-regulation are low. They certainly increase in times of quickly changing environment.
Furthermore, the regulator's discretion regarding s may induce firms and consumers to incur costs to influence the outcome of the regulatory process.

The properties of ror-regulation without regulatory lag have received considerable attention in the literature. As a sample, assuming a neoclassical production function for the regulated firm and a binding regulatory constraint without lag, models under certainty show that the profit-maximizing regulated firm will

(1) overcapitalize in producing a given output (Averch and Johnson, 1962),

(2) not waste any inputs (Bailey and Coleman, 1971),

(3) operate in the elastic region of the revenue curve (Bailey, 1973, p. 74).

(4) as long as s > r the capital usage is increased as s is decreased (Takayama, 1969).

As a general result, ror-regulation cannot hope to maximize either consumers' surplus or social surplus. However, Klevorick (1971) and Sheshinski (1971) have shown that it can hope to increase welfare vis-a-vis the unconstrained monopoly situation. Ror-regulation in an uncertain environment (see, e.g., Neufeld and Vogelsang, 1979, or for a survey Baron, 1978) or with different firm objectives (see e.g. Crew and Kleindorfer, 1979) may yield other results without establishing

---

5This terminology is somewhat misleading, because the very notion of ror-regulation presupposes that the regulator does not know cost and demand functions.
efficiency. Models on ror-regulation without lag tend to stress
distortions in terms of our strong postulate P9 (Optimality). Instead of
following them I shall focus on the effect of lagged ror-regulation by
discussing P8 (No misrepresentation).

Bailey and Coleman (1971) have shown that introducing a regulatory lag
may lessen the Averch-Johnson bias potentially introduced by
ror-regulation. If the regulator sets $s = r$, in the limit the bias could
even vanish altogether (Bailey, 1973). These results which were partly
foreseen by Baumol (1970) are plausible once one recalls that during the
regulatory lag the firm has an incentive to lower its costs.

In terms of prices it has to give away the advantages of lower costs only
in the next period. Hence, using the length of the regulatory lag as an
instrument to improve ror-regulation is suggestive. Contrary to this
notion we have above introduced the regulatory lag as a necessary evil.
I conjecture that under realistic conditions this comes nearly as close
to the truth as its potentially beneficial effect.

We maintain that the regulatory agency does not know the cost and demand
functions of the firm and that the data it gets about the performance of
the firm during the last period on costs, while accurate, do not
necessarily represent minimum costs. This means that the agency only
receives the information necessary to apply the regulatory rule and
provided by its outcome. Otherwise the whole approach of ror-regulation
would make little sense. On the other hand, we assume that the regulated
firm knows the allowed rate of return $s$, its cost and demand functions. Under these circumstances there may be not only incentives for the firm to reduce costs during lag intervals but also to increase costs from time to time. It will use these cost increases to improve its strategic position for future rate revisions. In order to do this the regulated firm may for some time even employ a capital that exceeds the level of capital used by the Averch-Johnson firm with no lag. To demonstrate this, assume that in periods $t - 1$ and $t$ the firm operates at the Averch-Johnson level of capital $K_{AJ}$. It employs labor $L_{AJ}$ and earns a profit $\Pi_{AJ}$. We assume a neoclassical production function with the possibility for the firm to change both its inputs instantaneously. There shall be constant returns to scale. Input prices are assumed given and stationary over time. The regulated price is specified by

$$p_{t+j} \leq \frac{scK_{t+j-1} + wL_{t+j-1}}{x_{t+j-1}}.$$ 

Thus the lag period is assumed to be $T = 1$.

Now we ask under what conditions the firm would want to employ more capital than $K_{AJ}$ in $t + 1$. The shortest and simplest calculation for the firm would involve two periods. Therefore, the firm does not want to lower average costs below the AJ level. Instead, in the first period the firm would increase capital by $\Delta K$. Staying on the same isoquant it can simultaneously reduce its labor input by $\Delta L$. Thus in $t + 1$ its profits would be reduced, because input inefficiency has been increased at $P_{t+1} = P_t$. In $t + 2$ the firm could make use of a higher price $P_{t+2}$ and employ less capital than $K_{AJ}$. It could make higher profits than in period $t$. In period $t + 3$ it goes back to the old AJ-point with $P_{t+3} = P_t$. The simplifying assumption that in $t + 2$ the firm keeps costs of
the AJ-level makes it unnecessary to investigate if the profit-maximizing firm would want to go back to $K_{AJ}$ and $L_{AJ}$ at all.

Figure 1 gives a simplified picture of the two steps. In the Averch-Johnson case the firm is assumed to produce at average cost $OA$ and receives a price of $OE$. Profits are $\Pi_{AJ} = ABCE$. Now in period $t+1$ the firm increases total costs by $rcaK - wAL = ABFG$. This results in an allowed price increase for period $t+2$ of $\Delta p = EK$. The potential profit increase in period $t+2$ is the difference between $\Delta p(x_{AJ} - \Delta x) = EHJK$ and $\frac{\Delta x}{x} \Pi_{AJ} = BCHL$. As this increase occurs in the period $t + 2$ only, it has to be discounted to the present. Thus, in present value terms the difference between the cost increasing loop and remaining at the AJ point for two periods is:

$$\Delta \hat{\Pi} = wAL - rcaK + \frac{1}{1+r} (\Delta p(x_{AJ} - \Delta x) - \frac{\Delta x}{x} \Pi_{AJ})$$

If this difference is positive the move will be profitable. The formula contains three interesting parts: $wAL - rcaK$ is the initial cost increase necessary to induce the price increase. This applies to the total quantity $x_{AJ}$. $\Delta p(x_{AJ} - \Delta x)$ is the per unit increase in profit, $P$, by virtue of the allowed price increase multiplied by the reduced output. $\frac{\Delta x}{x} \Pi_{AJ}$ is the AJ-profit lost through the output reduction. The interaction of these terms is not straightforward. In order for the total to be positive the allowed price increase for the second period must be substantially larger than the average cost increase in the first period. This is so because in the relevant range demand is elastic. A
cost increase of one unit of money per unit of output buys the firm a
price increase at least unit in the next period. Hence s has to be
large relative to r. Furthermore, the profit per unit of output at the
AJ-point should be small and demand not highly elastic. The extent of
the resulting additional overcapitalization is quite limited. x may not
become large relative to x_{AJ}.

It takes a rather special example to demonstrate in a stationary
environment that the regulatory lag actually gives the firm an invitation
to overcapitalize simply in order to misrepresent information. The
difficulty arises because the additional costs necessary to raise prices
for the next period accrue to a larger quantity than the benefits. In
the real world cost increases can most easily be pretended to occur in
business slumps. Here quantities are small, fully distributed costs of
capital intensive industries rise anyhow. Further overcapitalization
could have a large beneficial effect in the next boom period. The
following numerical example pictures this situation. The regulated firm
faces demand functions \( p_{t+1} = 100 - 2x_{t+1} \) and \( p_{t+2} = 100 - x_{t+2} \)
for the two periods \( t+1 \) and \( t+2 \). Thus \( R_{t+1} = 100x_{t+1} - 2x_{t+1}^2 \) and
\( R_{t+2} = 100x_{t+2} - x_{t+2}^2 \). The production function is \( x = K^{1/2}L^{1/2} \). Input
prices are \( w = 1, c = 20, r = 0.05 \). The allowed rate of return has been
set as \( s = 0.5 \). The firm knows all these functions and parameters.

The Averch-Johnson point for the two periods can be computed by using the
two first order conditions: \( \frac{\partial R}{\partial L} = w \) and \( R - wL - scK = 0 \) (Bailey, 1973,
This yields for \( t+1 \): \( K_{AJ} = 124.5 \); \( p_{AJ} = 50.2 \); \( x_{AJ} = 24.9 \); \( \pi_{AJ} = 1120.5 \). For \( t+2 \) the corresponding values are: \( K_{AJ} = 249 \); \( L_{AJ} = 9.96 \); \( x_{AJ} = 49.8 \); \( p_{AJ} = 50.2 \); \( AC_{AJ} = 5.2 \); \( \pi_{AJ} = 2241 \). Thus \( \pi_{AJ} = 1120.5 + \frac{1}{1.05} \cdot 2241 = 3254.786 \).

Now, assume a regulatory lag of one period. If the firm wants to increase its price to \( p = 51 \) for period \( t+2 \) this can be achieved by increasing \( K \) and decreasing \( L \) in \( t+1 \). This yields in \( t+1 \):

\[
K^* = 126.500; \quad L^* = 4.901; \quad \pi^* = 1118.579
\]

In period \( t+2 \) we find:

\[
p^* = 51; \quad x^* = 49; \quad AC^* = 5.2 \quad \text{and} \quad \pi^* = 2244.2
\]

Thus \( \pi^* = 1118.579 + \frac{1}{1.05} \cdot 2244.2 = 3255.9127 \). Hence the two period loop boosts the discounted present value of profits to the firm compared with the Averch-Johnson level. The example shows a deterioration beyond the Averch-Johnson point on account of a regulatory lag. Nevertheless, there could be an optimal lag period, under which the Averch-Johnson bias is reduced and welfare increased. My only reservation is that finding this lag requires knowledge which the regulatory agency ordinarily does not have. With the knowledge it would just as well or better prescribe actions to be taken directly by the firm, and not use lagged ror-regulation. On balance, I submit as unconscious lag is a priori neither beneficial nor detrimental to ror-regulation, just necessary.

The discussion in Bailey (1973, pp. 122-123) generates hope that lagged ror-regulation generally fares better with \( s = r \). However, there the efficiency property in the limit was established by using myopic arguments. The firm starts off at the AJ point on the constraint curve.
Thus profits are zero. By lowering costs the firm can now make a profit in the first period. This induces the regulator to set a new lower price, forcing the firm back onto the constraint curve. Again the firm improves its position by producing efficiently. "Eventually, by always moving to the minimum cost point on its current isoquant the firm will be drawn to the minimum cost, maximum output point..." "...no matter how small the length of the lag interval, so long as it is positive" (Bailey, 1973, p. 123). However, if the minimum average costs are increasing with output the firm will face a loss in the second period. Price then is just high enough to cover total costs at the previous output level but too small at the increased quantity demanded. In period three the firm could again make a profit, which would induce another loss in period four. At the outset it seems unclear if the summed present discounted value of these profits and losses is positive, which would be necessary for the process to start. A sufficient condition for the profitability is that average costs are non-increasing.\(^6\) Then, lagged regulation with \(s = r\) will eventually result in an efficient steady state, in which both postulates P8 and P9 are fulfilled. Nevertheless, as Sappington (1979) has demonstrated in a number of examples, strategic misrepresentation of costs could become profitable for the firm on the way to the steady state, unless it applies a very high internal discount rate to future profits. Here, like in the above numerical example, the result comes about, because the benefits of misrepresentation accrue to a larger quantity than the cost of such behavior. The welfare effect of

\(^6\)This follows from step 1 of the proof of Proposition 1 in Vogelsang and Finsinger (1979).
this could be worse than the ordinary Averch-Johnson effect with s r, because misrepresentation can take the form of pure input waste.

As a result of our discussion, ror-regulation fares rather well under postulates P1 to P7, although the regulator's discretion in choosing the allowed rate of return s may introduce precisely those three kinds of failures which led to postulate P1. More seriously, on both postulates, P8 and P9, ror-regulation is seen to fail. The severe criticism of ror-regulation by economists so far has rested on the nonfulfillment of postulate P9. Against this our discussion suggests that even if postulate P9 can ultimately be fulfilled in a steady state, nonfulfillment of postulate P8 on the way may pose problems. This can occur whenever the firm is not forced or motivated to reach the steady state immediately.

It would surprise any regulation economist, if efficient regulation by rules were feasible, meaning that P1 to P7 could hold at the same time with P8 and P9. Thus, I conjecture that quite generally regulatory rules obeying postulates P1 to P7 and P9 will be subject to potential misrepresentation (P8 not fulfilled). This is obvious if the regulated firm can misrepresent costlessly. Then it could always pretent to incur costs which under the rule would allow it to ask for the true monopoly price. Thus, in order to prevent this trivial type of impossibility the regulatory agency has to audit that reported costs do not contain profits.
4.2 A Rule for Multiproduct Monopoly Pricing

Regulated natural monopolies often are multiproduct firms. If the price structure for their products is not regulated directly one expects output distortions to go along with input distortions caused by regulation. For at least two reasons efficient direct rate structure regulation could be extremely difficult to achieve. First, the specific micro type of information on product specific demands and costs necessary to pursue this goal is especially hard to come by. If anyone has it at all it is held by the staff of the regulated firm and normally not controllable through regulatory auditing. Second, changes in the rate structure potentially hit the individual consumer or consumer groups harder than changes in the rate level. Their relative position (competitiveness) is altered; structure is a better lever; and changes in price structure seem less evitable. Thus pressure groups may both have more incentive and means to influence rate structure regulation in their favor. A regulatory rule aimed at the price structure could help overcome these tendencies and improve on information requirements. Vogelsang and Finsinger (1979) have recently suggested such a regulatory price adjustment process for multiproduct monopolies. Its core is a constraint which the regulatory agency applies iteratively and which for time period $t$ constrains the firm to choose prices $p \in \mathbb{R}^n$ in the set $R_t = \{ p | px_{t-1} - C(x_{t-1}) \leq 0 \}$ and to serve all demand at those prices.

How does this procedure fare under the above postulates?
P1 (No discretion): When applied, the process constrains the regulators to the utmost degree.

P2 (Enforceability) and P3 (Publicity): The mechanism could be easily enforced and made known to the public.

P4 (Duration): This postulate is a requirement for the process to develop its properties. So it can be assumed satisfied.

P5 (Privacy), and P9 (Optimality): The main motive to construct this process was to induce the regulated firm to approach prices satisfying the Ramsey condition for a constrained consumers' surplus maximum without requiring regulators or the public to know items which only firm insiders have access to. This is proved in Vogelsang and Finsinger (1979). The data required by the process are the quantities, prices and the total cost figures of the firm for the last period. All these are rather easily available if there are no intertemporal cost effects.7

P6 (Feasibility): Under the mechanism the firm's possibility to avoid losses depends on a stationary (nondeteriorating) environment and nonincreasing ray average costs ($rC(x) \geq C(rx)$ for all $r \geq 1$, $r \in \mathbb{R}^1$, $x \in \mathbb{R}^n$). With increasing ray average costs periods of losses will alternate with periods of profits but a positive discounted present value

7See, however, footnote 4, page 20.
cannot be guaranteed. This restricts the procedure in its present form to be applied to decreasing cost industries only.  

87 (Limitation): The procedure takes time to evolve. During this time conditions are likely to change. The most likely changes refer to a) input prices, b) demand, and c) production technology.

a) Input price changes: They can affect the feasibility of the process in terms of the viability of the regulated firm and in terms of the convergence property. Feasibility is only endangered by input price rises. Amending the rule by an automatic input price adjustment clause does not necessarily help and creates new problems (Baron and DeBondt, 1978). Especially feasibility may be hampered if outlays for the inputs used in the previous period are increased by more than the previous profit. On the other hand, input price decreases may overtake the speed of convergence of the process.

b) Demand changes: Aside from the magnitudes involved, these have similar effects to input price changes. Decreasing demand corresponds to increasing input prices. However, due to the problem of measuring demand there exists no equivalent to input price adjustment clauses on the demand side. Inflation may be viewed as a combination of a) and b).

8 If, however, the firm does not have to serve all demand at current prices losses can be avoided. Under the further assumption of myopic profit maximization and $x_{t+1} \geq x_t$ if demand was not satisfied in period t the process will still converge. But the efficiency loss on the way to the optimum is carried by the consumers.
c) **Technology changes:** These may result in cost decreases affecting the existence and the speed of convergence. More important, if through technological advance the decreasing cost condition no longer persists both entry and overshooting (losses) may pose problems.

All the changes described raise problems of uncertainty. If they do not occur the firm may be perfectly well informed. Only through strategic misrepresentation could it prevent the regulators to become the same. In a changing environment both regulators and the firm will lack some information. Hence, even if under perfect information the firm would incur no losses and prices would converge to a constrained welfare optimum this in general, will not be true under uncertainty. Recent work by Bawa and Sibley (forthcoming) suggests that randomizing the regulatory constraint may be a way out.⁹

Summing up, the process described by Vogelsang and Finsinger is defined under specialized conditions. If these are not obtained other mechanisms have to replace it to cope with postulate P7.

**P8 (No misrepresentation):** If the assumed conditions for the mechanism hold the regulated firm has no incentive to misrepresent data to such an extent that the process does not eventually converge to the constrained welfare optimum. We simply require the regulator not to relax the

---

⁹In this sense, regulatory policy uncertainty can be advantageous under well specified conditions.
constraint if the firm shows a loss for a period. As Sappington (1979) has shown, however, misrepresentation can occur on the way to the optimum. This again hints at the validity of the general conjecture on regulatory rules stated above in Section 4.1. Misrepresentation here can take the special form of a waste of inputs. This is likely to follow from the lump sum effect the constraint $R_t$ has on the possibility of cost manipulation. On the other hand, I submit that the incentive for pure waste crucially depends on the shapes of the demand and cost functions. The constraint $R_t$ gives the firm enough degrees of freedom not to be forced to misrepresent information in the form of waste.

The adjustment process described by $R_t$ involves no systematic bias of the Averch-Johnson type because there is no asymmetric treatment of any cost component which the firm could influence. However, when applied under realistic conditions the assumption of no intertemporal cost effects will be violated. It is this feature that originally gave rise to ror-regulation and to all the problems involved in measuring the cost of capital to the firm. Once intertemporal cost effects are present, cost of capital will play a dominant role also in this process.
5. **Extensions and Conclusion**

We have tried to motivate some postulates for regulatory rules by demonstrating their meaning in two concrete examples for pricing rules. Regulatory rules may also be defined especially on other components of regulation such as entry and the quality of service. A rule heavily discussed in the recent literature on natural monopoly is whether one should allow entry where the incumbent and the intruder are treated differently by the regulator. It has been shown that the affirmative may be suboptimal (e.g. Panzar and Willig, 1977). So another rule would be to treat them similarly. To wit, if both firms have an obligation to serve all demand at current prices, inefficient entry will be itself hard to sustain. If nothing more an entry rule is simple as compared to the direct regulation of entry, which takes recourse to proving that a particular entry is efficient or not.

Similarly, an obligation to serve may be a clear rule only if the consequences for not serving are specified. Obliging the regulated firm to pay damages to unserved consumers may be one way to turn an obligation to serve into a regulatory rule.

One consequence of discussing postulates not directly defined on efficiency was to draw attention to procedural and informational problems inherent in regulation. We conjecture that it may be impossible to avoid strategic misrepresentation under regulatory rules that depend on data observable to regulators external to the firm. As a consequence, it
would be desirable to compare the extent and form this takes for different regulatory rules. Furthermore, implementation costs as well as the uncertainty caused by the limit in scope of applicability should be compared before a judgement on a particular rule is made. Viewed from these perspectives the Averch-Johnson effect may be a low price to pay for a rule that is applicable under practically all circumstances.
References


Figure 1: Profit Increasing Misrepresentation